Surjit B. Dixit
Felicia Pitici
D. L. Beveridge
Chemistry Department and Molecular Biophysics Program, Wesleyan University, Middletown, CT 06459

Structure and Axis Curvature in Two $dA_6 \cdot dT_6$ DNA Oligonucleotides: Comparison of Molecular Dynamics Simulations with Results from Crystallography and NMR Spectroscopy

Received 20 July 2004; accepted August 20, 2004
Published online 3 November 2004 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bip.20157

Abstract: *Molecular dynamics (MD) simulations have been performed on the A6 containing DNA dodecamers d(GGCAAAAAACGG) solved by NMR and d(CGCAAAAAAGCG) solved by crystallography. The experimental structures differ in the direction of axis bending and in other small but important aspects relevant to the DNA curvature problem. Five nanosecond MD simulations of each sequence have been performed, beginning with both the NMR and crystal forms as well as canonical B-form DNA. The results show that all simulations converge to a common form in close proximity to the observed NMR structure, indicating that the structure obtained in the crystal is likely a strained form due to packing effects. A-tracts in the MD model are essentially straight. The origin of axis curvature is found at pyrimidine–purine steps in the flanking sequences.* © 2004 Wiley Periodicals, Inc. Biopolymers 75: 468–479, 2004

Keywords: *A-tracts; DNA curvature; molecular dynamics simulations; crystal packing effects; pyrimidine-purine steps*

INTRODUCTION

The structure of "A-tracts," oligonucleotide sequences of adenine–thymine base pairs, has attracted considerable interest in nucleic acids research as a consequence of the linkage between A-tracts and the sequence dependent curvature of DNA.[1,2] DNA sequences with tandem repeats of A-tracts phased by 10 bp, a full turn of a B-form helix, generally exhibit anomalously slow gel migration and increased circularization rates compared with random sequence controls. These results are considered indicative of enhanced axis curvature in phased A-tracts, but do not uniquely specify the origin of the curvature. Opinion is divided over whether the origin of curvature in A-tract sequences with helix phasing is within A-tracts (the wedge model),[3] at the point of articulation of A-tracts and flanking sequences (the junction model),[4] or in sequence elements other than A-tracts (the non-A-tract model).[5]

A set of recent studies has focused on dodecamer sequences containing the dA_6 motif, in order to characterize the structure of a prototype A-tract. In particular, structure determinations of dA_6 oligonucleotides have been reported by both high resolution x-ray crystallography[6] and NMR spectroscopy[7] using the

Correspondence to: S. B. Dixit; email: sdixit@wesleyan.edu
Biopolymers, Vol. 75, 468–479 (2004)

new information obtained from residual dipolar coupling (RDC) experiments. In particular, the NMR results show axis curvature at or near both the 5′ and 3′ junctions of dA_6 with flanking sequences, resulting from bp axis base inclination and appreciable roll at bp steps. The results on dA_6 dodecamers from x-ray and NMR differ notably, particularly in the direction of axis curvature. The NMR model was the first to show overall axis curvature toward the minor groove viewed from opposite the A-tract, a property attributed to phased A-tracts sequences in earlier experimental work.[8]

The structures of oligonucleotides in solution have become accessible to detailed theoretical study by MD simulations including water and counterions. A comparison between MD calculated and observed NMR results provides an assessment of the quality of the MD models. For the prototype case of d(CGCGAATTCGCG), a recent article reports results that are in close accord,[9] but more cases must be examined to determine whether results are generally valid. A specific comparison of MD and crystallographic results on A-tract oligonucleotides has been reported, with generally close accord with the B′ form of DNA (relatively narrow minor groove, high propeller twist) characteristic of A-tracts.[10] Since the structure of the dA_6 sequences report an interesting and significant divergence between x-ray and NMR results, a study of this system provides a sensitive next step in the validation of MD models of DNA as well as the potential for additional insight into the DNA axis curvature problem.

In this article we discuss results from four new MD simulations on d(GGCAAAAAACGG) and d(CGCAAAAAAGCG) sequences, the NMR and crystallographic sequences with PDB designations of 1FZX and 1D98, respectively. MD simulations were initiated starting from structures obtained from NMR spectroscopy and x-ray crystallography, respectively, as well as the canonical B-form DNA of the same sequences. The objective is to provide a comparison of the MD results with those from crystallography and NMR on dA_6 dodecamers and to investigate further what new knowledge the MD simulations reveal about the nature of A-tract–induced axis curvature.

BACKGROUND

The presence of tandem repeats of A-tracts with helical phasing in DNA sequences that exhibit anomalously slow migratory behavior in gel electrophoresis experiments has been noted in an number of recent research studies.[4,11–14] Circular DNA containing phased A-tracts exhibit an overall curvature that has been reported to be 11 to 28° per A-tract[14–17] and exhibit a strong temperature[18,19] and salt dependence.[13] The consensus value close to room temperature is about 18°.[20] Zinkel and Crothers[8] studied the direction of curvature in A-tract DNA relative to the bend induced in DNA when bound to the catabolite activator protein and concluded that the effective bend direction is equivalent to minor groove compression at the center of the A-tract.

The first single crystal structure of an A-tract–containing DNA dodecamer sequence d(CGCAAAAAAGCG) indicated the structure of the A-tract to be straight and rigid, raising a question about the role of A-tracts in A-tract–induced curvature. The crystal structure exhibited deformations in bp steps flanking the A-tract.[6] This observation was corroborated by subsequent crystallographic studies of other A-tract–containing DNA sequences, i.e., essentially straight A-tracts. In one notable case, the direction of bend in the helix axis in two different structures of the same sequence were different, indicative of packing effects and raising the question of whether crystal structures correspond to distinct stable substates of the DNA or are packing induced strained forms of the sequence.[21] The bend in each of these structures originates mainly as a result of kinks at the CpA and CpG steps in the regions outside the A-tracts. Analysis of dinucleotide step properties in the crystal structures of DNA both with and without A-tract reveals that such pyrimidine–purine (YpR) steps have an intrinsic deformability.[22] These steps appear with either negligible roll (i.e., straight) or with a positive roll value, i.e., into the major groove along the long axis of the bp.[22,23] The question of whether A-tract–induced axis bending occurs within or outside the A-tracts per se for DNA in solution has been vigorously debated.[5,8,24]

Recent developments in the use of residual dipolar coupling (RDC) data in NMR spectroscopy permit higher resolution structure determination of DNA in solution.[25] Lu et al.[7] solved the structure of the A-tract–containing dodecamer DNA d(GGCAAAAAACGG) sequence by NMR using RDC data, which is close but not identical to the sequence solved by x-ray crystallography.[6] The results indicated an overall helix axis bend of 19°, of which 14° occurs in the G,C region flanking the A-tract and the remaining 5° bend occurs evenly within the A-tract region. Notably, the direction of curvature in the NMR structure is toward the minor groove, consistent with solution state data but orthogonal to the helical axis direction in the crystal structure.

In a recent article, we compared a MD model with NMR and x-ray diffraction derived structures of the d(CGCGAATTCGCG) EcoRI endonuclease DNA se-

quence and it was established that a strong agreement between the NMR and MD model existed.[9] The approach taken was based on two-dimensional (2D) NOSEY spectra back-calculated from MD aimed at overcoming the deficiencies introduced in the refinement procedures employed in NMR, which depend on empirical restraints and conformational energy calculations based on empirical potential energy functions. To this end, the observable 2D NOSEY volumes and dihedrals were back-calculated from the MD trajectories for the DNA sequence in solution and compared with the equivalent NMR spectral data. Close agreement was observed between the back-calculated 2D NOSEY spectra for the RDC NMR structure solved by Tjandra et al.[26] and the MD model. Comparison of the DNA helicoidal parameters for the solution state MD model of the EcoRI endonuclease sequence with the values in the NMR RDC structure of this molecule indicate that the two models agree within the statistical uncertainty of the MD model in most of the cases.

MD simulations on A-tract DNA including explicit water and ions have yielded models that compare favorably with known structural properties. Essential feature of B′-DNA such as high propeller twist and narrowing of a minor groove in the 5′ to 3′ direction have been consistently observed in simulations on A-tracts.[10,27–32] Recent comparison studies of the MD model of A-tract DNA using AMBER,[33] CHARMM,[34] and the BMS[35] force fields by Lankas and coworkers[31] have revealed that these force fields behave in a similar manner on this point. A current perspective on DNA curvature and flexibility on the basis of MD simulations has been presented recently.[36]

All-atom MD models present a perspective on the structures of oligonucleotides complementary to x-ray and NMR experiments and have the advantage of a detailed microscopic evolution of structure with time. A property such as axis curvature can be readily calculated and analyzed with respect to sequence-dependent features. Here, we apply MD simulations to study the structural details of A-tract–containing DNA dodecamers that have been solved by x-ray crystallography and RDC NMR spectroscopy in order to further validate the capabilities and limitations of MD models of DNA against experimental data. In particular we address the issues of the degree and direction of bend and its origin and comment on the nature of the apparent differences in crystal and solution structure of dA_6 dodecamers.

METHODS

MD simulations were performed on the d(GGCAAAAAACGG) structure solved by MacDonald et al.[7] (PDB ID: 1FZX) using RDC NMR spectroscopic techniques. The first structure in the family of NMR refined structures, being the most representative of the solved NMR structures, was employed. The second simulation was on the d(GCGAAAAAACGC) structure solved using x-ray crystallography by Nelson et al.[6] (PDB ID: 1D98). Two other simulations on each of these sequences starting from canonical B-form DNA double helix[37] were also carried out. Exactly similar simulation conditions were employed in all of the four calculations.

Each system was hydrated in a rectangular box of TIP3P[38] water molecules extending to a minimum distance of 12 Å from all DNA atoms in the system. The resulting box sizes are approximately 52 Å × 52 Å × 75 Å, solvating the DNA with approximately 4800 water molecules. In each case a sufficient number of Na^+ ions were included to neutralize the net charge of the system being simulated and their position was randomized while placing them at least 5 Å away from the DNA atoms.

Molecular potentials are described by the parm94 force field reported by Cornell et al.[33] Periodic boundary conditions are employed with long-range interactions calculated via the Particle Mesh Ewald (PME) method.[39,40] All the MD simulations were performed using AMBER 8.0[41] and PMEMD version[42] of the sander module. The PME calculations were carried out with a 9 Å cutoff for direct space nonbonded calculations and a 10^{-6} Ewald convergence limit for the long-range electrostatic interactions. Each MD was initiated with a 500 steps (100 steps steepest descent and 400 steps conjugate gradient) restrained minimization to optimize the ion and solvent positions followed by another 500 steps of unrestrained minimization on the complete system. MD heating and initial system equilibration was performed in a semirestrained fashion as follows. First, 20 ps of heating was performed on the constant volume system with a 1-fs time-step while restraining the DNA heavy atoms with a 25 cal/mol force constant followed by a 10-ps run during which the restraints were slowly reduced. Unrestrained constant volume dynamics were continued for another 10 ps, at which time constant temperature, constant pressure MD (isothermal–isobaric ensemble) was initiated, utilizing the Berendsen algorithm for temperature and pressure coupling. At this point a 2-fs time step was introduced and all the covalent bonds involving hydrogen atoms were held rigid using the SHAKE procedure[43] with a tolerance of 0.00001 Å. The system energy was stable at this point and the density stabilized close to 1 g/cm^3. To maintain the alignment of the DNA within the center of the rectangular box, translations of the solute center of mass were removed every 1000 steps of MD.

Each of the four systems was simulated for a length of 5 ns on a PC linux cluster at the Texas Advanced Computing Center. The properties of the simulation were stable as a function of time in the length of trajectory. In a previous study,[44] we reported relaxation times of 0.5 ns for DNA parameters and 5 ns was chosen as an order of magnitude longer than the relaxation time. The issue of convergence was further confirmed by performing a principal component

analysis wherein the displacements of the essential dynamical modes were studied. The net fluctuations appear to be converged as a function of the simulation time, confirming sufficient conformational sampling.

All of the comparisons of simulated and the experimental structures were performed after aligning the central dA_6-dT_6 section of the dodecamer structures. We report detailed analysis of the origin and direction of DNA curvature using the normal vector plots (NVP).[45] To generate the NVP, the DNA helical axis is first aligned to the *Z*-direction, followed by translation of bp normal vectors to a common origin, while maintaining their orientation with respect to the *Z*-axis. The sequential connection of the *XY* components of the unit normal vector tips of each bp plane constitutes the NVP. The distance between any two vector end points is the arc sin of the angle between the mean planes through each bp. These plots measure the magnitude and direction of curvature at the individual bp step level. While a straight DNA helical axis with a small writhe such as the canonical B-form DNA would exhibit a NVP with the vector tips lying on the circumference of a small circle around the origin in the plot, any variation of the individual bp step geometry from the canonical structure appears as a discontinuity in the plot.

RESULTS

MD is reported for dA_6-tract–containing dodecamer DNA structures solved by x-ray diffraction (PDB ID: 1D98)[6] and NMR spectroscopy (PDB ID: 1FZX).[7] Hereafter the simulation trajectories initiated from the experimentally derived structures will be referred to as MD-1D98 and MD-1FZX. The stereo image in Figure 1(A) shows the original 1D98 and 1FZX structures colored cyan and green, respectively. The structures are positioned least square fit with reference to the central A_6-T_6 region (red–blue pairs), along with their helical axis calculated using the program CURVES.[46] Superimposed on these structures are the corresponding MD-1D98 and MD-1FZX structures based on the average of the 5-ns MD simulation, shown in orange and white, respectively, all aligned with respect to the central dA_6-dT_6. Figure 1(B) shows the same set of structures from a view that is rotated by 90° with respect to Figure 1(A). As noted by MacDonald et al.,[7] the direction of axis curvature in the 1D98 crystal structure is noticeably different from the rest, with the direction of curvature essentially orthogonal to the 1FZX NMR form. After the 5 ns of simulation in solution environment, the MD-1D98 structure is almost indistinguishable from the MD-1FZX, and both of these structures appear to be close to the 1FZX structure solved by NMR spectroscopy. The structures obtained from the two other simulations starting from canonical B-form DNA

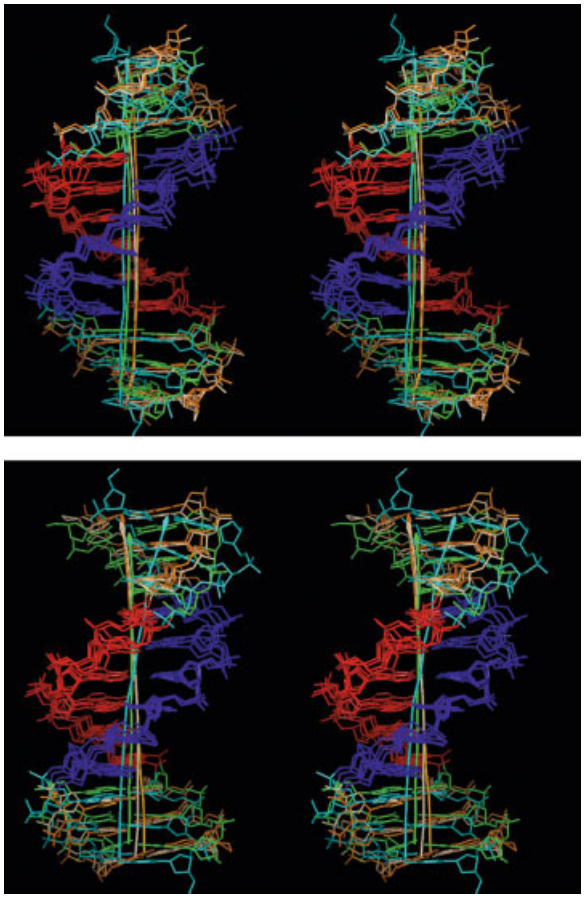

FIGURE 1 Stereo image (top) of the original 1D98 and 1FZX structures colored cyan and green, respectively, along with their helical axes calculated using the program CURVES.[46] Overlaid on these structures in orange and white are their respective average structure obtained from 5 ns of MD simulation. All the structures have been least square fit with reference to the central dA_6-dT_6 (red–blue pairs). Bottom image shows the same set of structures from a view that is 90° to that at the top. Image generated with PyMOL.[58]

(MD-CB-1D98 and MD-CB-1FZX) confirm that, irrespective of the starting configuration, the MD model of the A-tract–containing dodecamers is similar to the direction of curvature in the NMR-derived experimental structure. In each of these simulated structures, the DNA helical axis is curved toward the minor groove of the central A-tract region, in accord with the 1FZX NMR structure.

During the simulations, an equilibrated state is achieved within the first 250–500 ps of the simulation for the MD-1FZX, MD-CB-1FZX, and MD-CB-1D98 while the MD-1D98 takes a slightly longer time. All of the MD simulations exhibit structural stability with fluctuations in the range of 1–1.5 Å mass weighed all atom root mean squared deviations (RMSD) with

reference to the average structure. The equilibrated DNA structures in these simulations exhibit mean RMSD of about 2.3 Å with respect to the original 1FZX solution structure and about 2.8 Å with respect to the 1D98 crystal structure. On the other hand, the simulated structures differ by about 3 Å RMSD with reference to canonical B-form DNA structures. Figure 2 is a two-dimensional RMS plot that presents a matrix comparing the RMS difference between every snapshot in each of the four trajectories to every other snapshot. The light shades of grey indicate a small RMS difference and the grey scale code is shown to the right of Figure 2. Comparing individual snapshots in the simulations starting from the canonical B-form DNA structures (last two blocks) and the NMR structure (first block), we notice small RMS differences from almost the very beginning of the simulation. On the other hand, the MD trajectory of the crystal structure achieves the same low RMS difference after close to 2 ns of dynamics. This indicates that the structure of the A-tract in the 1D98 crystal structure is a strained form stabilized by crystal packing that ultimately adopts a structure that is morphologically similar to the rest of the simulated structures.

This observation is further confirmed by the 1D-RMS plot shown in Figure 3. The lines in black and red presents the RMS difference between the structures in the MD-1D98 trajectory with reference to the original 1D98 and 1FZX structures, respectively. The

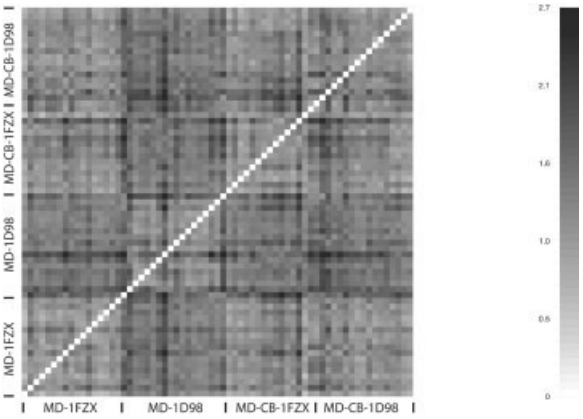

FIGURE 2 2D RMS comparison of the dA_6-dT_6 structures in the MD trajectories starting from the solution structure (MD-1FZX), crystal structure (MD-1D98), and their respective canonical B-form DNA (MD-CB-1FZX and MD-CB-1D98). The two axes represent the time series of structures from the four trajectories, each 5 ns long. The four quarters along each axis corresponds to structures in the MD-1FZX, MD-1D98, MD-CB-1FZX, and MD-CB-1D98 trajectories, respectively. The graph is symmetrical along the central diagonal.

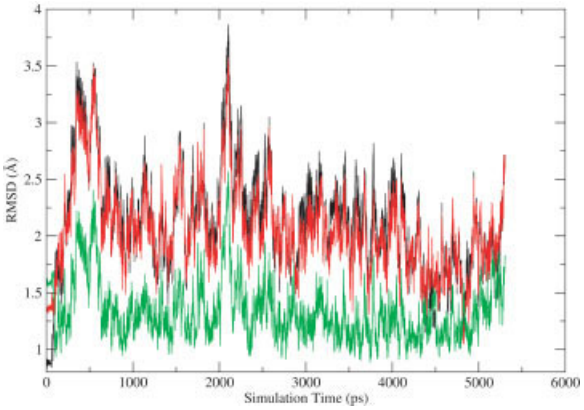

FIGURE 3 1D RMS plot comparing the dA_6-dT_6 structures in the trajectory of the 1D98 structures (MD-1D98) to the original 1D98 crystal structure (black), the original 1FZX NMR structure (red), and the average structure of the MD-1FZX trajectory (green).

considerably lower RMS difference between the structures in MD-1D98 and the average structure obtained from the simulation of the 1FZX structure shown in green confirms that the simulated structures in MD-1D98 and MD-1FZX are structurally similar.

The net curvatures of the DNA helical axis in the original 1D98 crystal and 1FZX solution structures are 19.2 and 19.0°, respectively, estimated using the program CURVES.[46] The net curvature is measured by fitting the global curved axes of the DNA structure. The average curvature angle for the structures in the MD-1D98 trajectory is 18.9 ± 9.5° while the average angle in the case of MD-1FZX structures is 20.0 ± 9.9°. The large standard deviations in the calculated curvature for the structures in the MD trajectory point to the degree of intrinsic flexibility in these structures. Another approach, as implemented in the program Madbend,[30] uses the bp step roll, tilt, and twist values and the specification of a reference plane to produce the global tilt and global roll and the bend magnitude. The corresponding values of curvature calculated using the program Madbend are 16.3 and 16.2° in the MD-1FZX and MD-1D98 structures, respectively. All of these calculated values are in reasonable accord with the experimental observations[20] that indicate a curvature close to 18° per A-tract under normal conditions.

Normal vector plots (NVP) were used to examine the nature of axis bending with respect to sequence. The first graph on the top left corner of Figure 4 shows the NVP of the original 1FZX solution structure in black and that of the 1D98 crystal structure in red. The normal vector tips of the AT base pairs are highlighted in yellow and cyan in the 1D98 and 1FZX

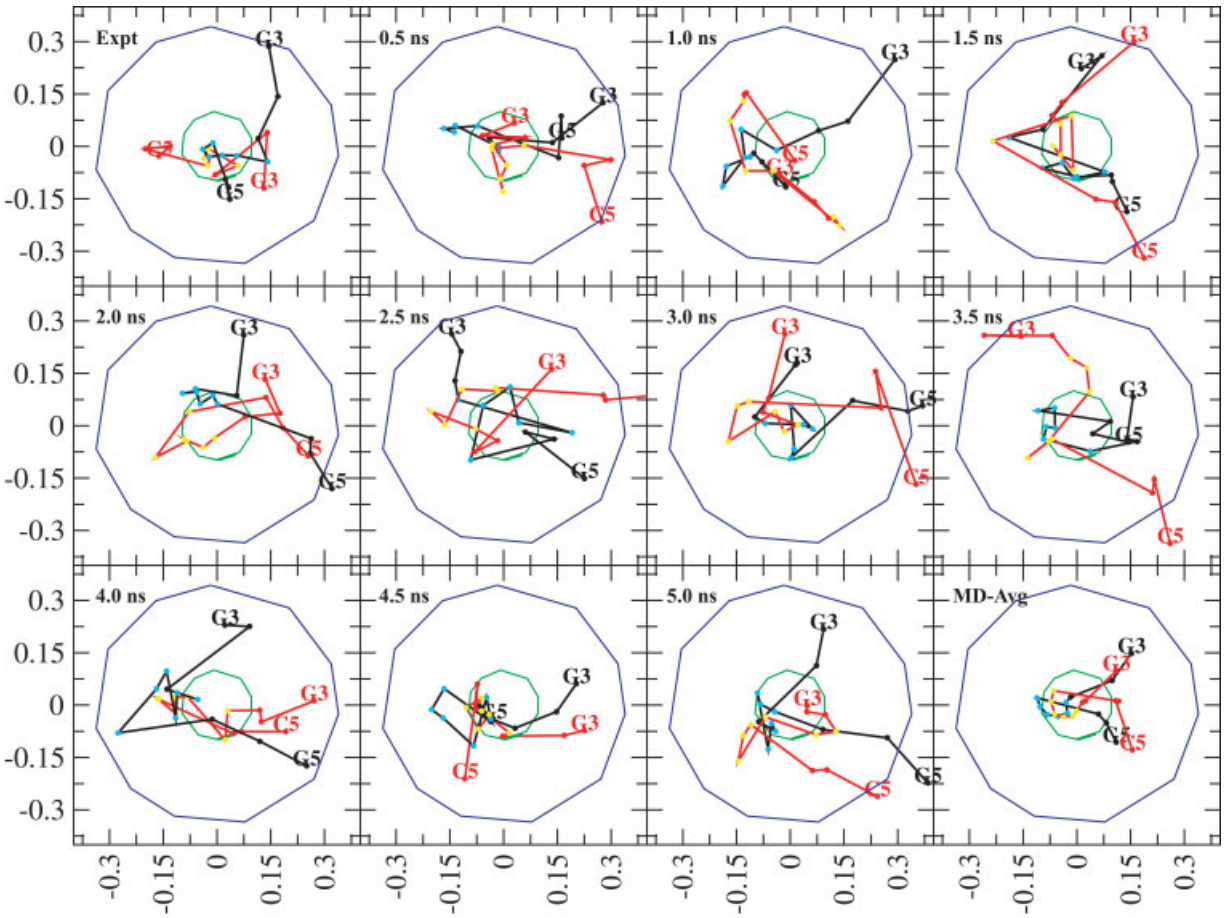

FIGURE 4 Normal vector plots (NVP) comparing the curvature of 1FZX (black) and 1D98 (red) structures in the original experimentally derived structures (top right) and at intervals of 0.5 ns in the MD trajectory. The bottom left image corresponds to average MD structure. All of the structures were aligned with reference to the dA_6-dT_6 of the original NMR structure before creating the plot. The normal vector tips of the A-tract base pairs are colored yellow in the 1D98 structures and cyan in the 1FZX structures. Normal vector plots of the canonical B and A form DNA[37] are shown in green and blue and correspond to writhes of 5 and 21°, respectively.

structures, respectively. The NVP analysis of canonical B-DNA in green and canonical A-DNA in blue are included for reference. Canonical A and B form DNA structures both have overall straight helical structures but exhibit writhes of 5° and 21°, respectively.

In comparing structural forms, the two DNA structures are aligned at the central A-tract region and a common orientation is maintained before generating all of the NVPs to interpret the differences in helix axis orientation. The DNA dodecamer is oriented along the *Z*-axis such that the minor groove of the central A-tract is facing the left. A similar alignment has been performed before generating all the NVPs in this figure to make them comparable. The south to north orientation of the vector tips in the 5′ to 3′ direction in the case of 1FZX indicates that the DNA is curved into the minor groove with respect to the central A-tract section of the dodecamer and the NVP of 1D98 crystal structure is almost orthogonal to the NVP of 1FZX. In both of the structures the central six points which correspond to the normal vectors of the dA_6–dT_6 base pairs are clustered in a small region while the sequence in the flanks exhibit much larger displacement, implying their larger contribution to the observed curvature of the DNA structure.

The set of graphs in Figure 4 shows the NVP of structures every 500 ps in the simulated trajectory. Plenty of flexibility at each of the individual base steps is observed in these structures. Noticeably, a good fraction of these structures have the 5′ to 3′ NVP roughly falling in the south to north direction, which, as noted above, implies a curvature similar to the 1FZX structure. The last graph (bottom right) of

Figure 3 shows the normal vector plots for the average structures obtained from the equilibrated section of MD-1FZX and MD-1D98 trajectories. The normal vector plots are essentially the same for both the average structures, implying that both the trajectories sample structures of similar axis curvature in their course. Further, the orientation of normal vectors in the 5′ to 3′ direction matches that of the original 1FZX structure and the large difference in the NVP direction of the original 1FZX and 1D98 structures is absent in the simulated solution state models of these sequences. The normal vectors of the central dA_6-dT_6 falls on the circumference of a smooth semicircle around the origin, indicating that the ApA in these trajectories are straight in comparison to the other bp steps flanking them in the DNA dodecamer.

The contribution of each bp step to the net curvature can be estimated from the component of each bp vector to the global curvature vector defined by the vector connecting the tips of the first and last bp normal vectors. In Figure 5(A), the curvature in the original 1FZX structure is compared to the average value along with one standard deviation, estimated from the normal vector plots for the ensemble of structures in the MD trajectory. The first bar in the histogram is a measure of the net curvature in the DNA structure and the component of each bp step is provided in the subsequent bars. The scale on the *Y*-axis provides a measure of the DNA curvature. The variance in the histogram of the NMR-derived structures corresponds to the distribution in the family of 10 model structures in the 1FZX PDB file. In Figure 5(B) the corresponding data for the 1D98 crystal structure and the ensemble average of MD-1D98 trajectory snapshots is presented. The ensemble average of net curvature from the MD structures is comparable to value in the experimental structures. The net curvature and contribution of the individual bp steps in the two simulations (MD-1FZX and MD-1D98) is comparable, indicating that essentially similar structural tendencies lead to curvature of both the sequences in solution state. Notably, in all the cases, the central ApA steps contribute little to the net curvature compared to the CpA and CpG steps flanking the A-tract. This property is observed in both the original experimental structures and the simulation results. The major exception is observed in the case of the CpG step at the 5′ end and the ApA, ApG, and GpC steps at the 3′ end of the A-tract in the crystal structure. The MD results for these steps are in accord with the properties of these steps in the other simulations and the 1FZX solution structure data. Interestingly, these are the nucleotides in the crystal structure that make contact as a result of crystal packing.

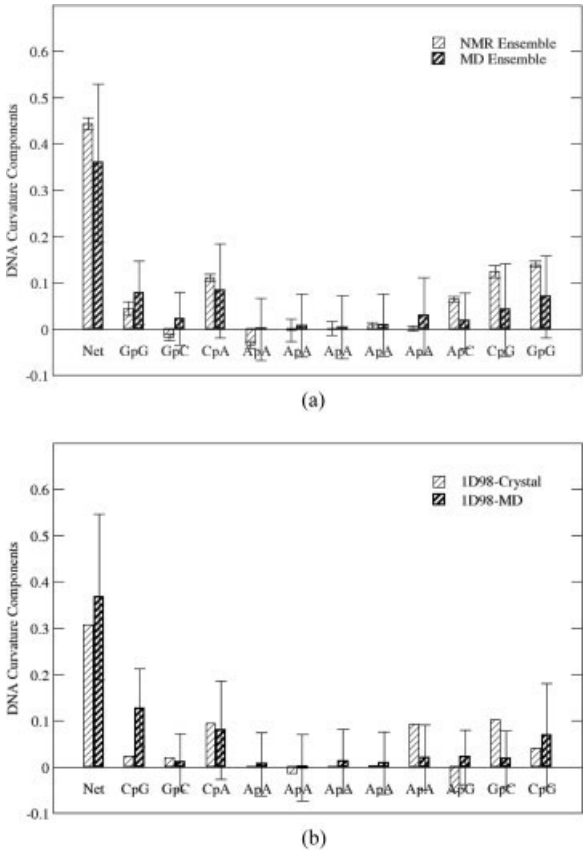

FIGURE 5 Components of the constituent base steps to the net curvature of the A-tract DNA sequence. (A) The curvature in MD-1FZX and the original family of structures of 1FZX and (B) the curvature in MD-1D98 and the 1D98 crystal structure.

From the plot of roll in the original NMR structure and MD results in Figure 6(A), we observe that the central A6 section maintains a low roll state averaging close to zero, while the C3pA4 step at the 5′ end of the A-tract and the C10pG11 step at the 3′ end of the dodecamers exhibit prominently large roll values close to 8° at each of these step. In the case of the crystal structure, while C3pA4 at the 5′ end exhibits a similar positive roll value, the G10pC11 step at the 3′ end exhibits a negative roll close to –13°, almost twice as large as the roll value at the corresponding CpG step in the solution structure but in an opposite direction. The property of a GpC step to exhibit negative roll values in contrast to the positive roll at CpG step is a familiar sequence effect.[47] The A8pA9 and A9pG10 steps in the crystal structure exhibit significant roll values that would contribute to orienting the direction of the helical axis, which are missing in the simulation of these structures in solution state. As noted in earlier simulations of A-tract sequences,[10]

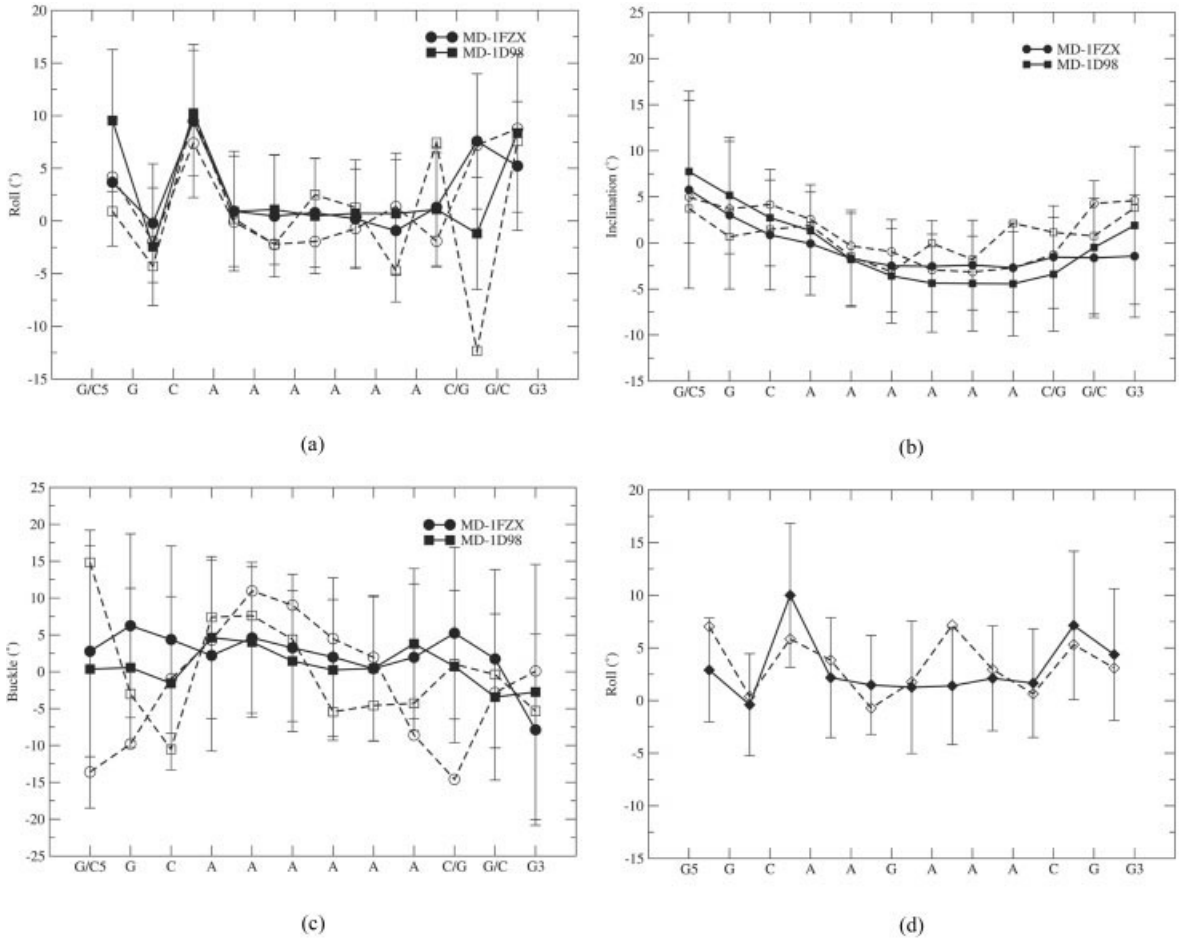

FIGURE 6 The (A) roll, (B) inclination, and (C) buckle angles along the DNA sequence in the original 1FZX (circle) and 1D98 (square) structures shown as dotted lines. The average values computed from all the snapshots in the MD trajectory of these structures are shown as solid lines. The second set of nucleotide names at certain positions on the *X*-axis indicates the difference in the 1D98 sequence. (D) The observed roll in the non-A-tract control sequence 1G14. The NMR structure data are shown with a dotted line and the average from MD trajectory is the solid line.

the structures exhibit the general features of high propeller twisting and narrow minor groove width associated with the B′-DNA. The value of tilt in the simulated trajectories lie close to the tilt observed in structures solved both by NMR/RDC and x-ray diffraction techniques.

As seen in Figure 6(B), the MD simulations are able to reproduce quite accurately the pattern of inclination of the NMR and x-ray diffraction derived structures. MacDonald et al.[7] report an increased negative inclination for the four central thymine bases compared to the corresponding adenine bases. As shown in Figure 6(C), this results in buckling of the AT base pairs within the A-tract, which causes a change of the buckle angle by almost 25° between the 5′ and 3′ end of the A-tract. The 1D98 crystal structure shows a change of about 15° between the two ends of the A-tract. The average value of buckle in the MD trajectory is significantly small compared to these reported values. The large standard deviation observed in the simulated value indicates that the experimentally observed buckle is accessible during the dynamics, although these are not the equilibrium values.

MD simulations have also been carried out on the non-A-tract control d(GCGAAGAAACGG) structure reported by MacDonald et al.[7] (PDB ID 1G14), which exhibits a reduced bend close to 9° in the NMR study. The plot of roll as a function of the DNA sequence in this structure is presented in Figure 6(D). The NMR structure presents an anomalous large positive roll at the A7pA8 step that is out of phase with the curvature observed in other portions of this DNA sequence, resulting in a significant decrease in the net curvature.

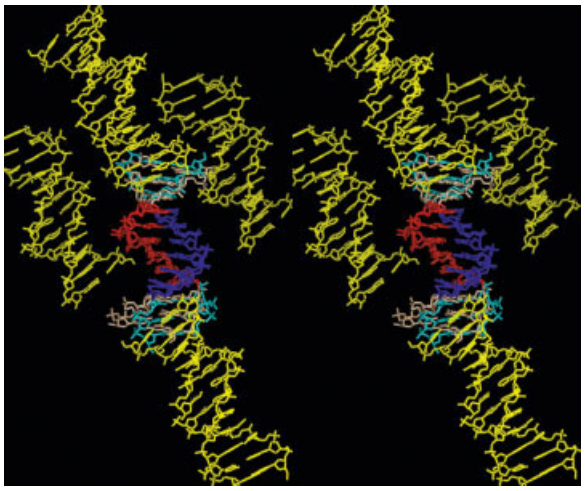

FIGURE 7 Stereoimage of the central dA_6-dT_6 (red and blue) of the 1FZX structure (wheat) least square fit onto the 1D98 structure (cyan) in its crystal lattice environment (yellow). Image generated with PyMOL.[58]

Simulations indicate that the presence of GC bp within the A tract does not affect the structure of the sequence significantly, which continues to exhibit close to a 18° bend in the DNA, similar to the regular A-tract sequence. These differences observed in the MD and original NMR structures remain to be resolved.

DISCUSSION

Detecting Crystal Packing Effects Using MD Simulation

The results presented here verify that all atom MD simulations of A-tract containing DNA dodecamer sequences provide a consistent model of the DNA structure in close accord with the experimental solution state structure derived from NMR spectroscopy, invariant of the starting configuration. The conformational change observed in the simulation of the original 1D98 crystal structure toward a state that resembles the structure in solution indicates that crystal packing likely contributes to the axis curvature in the original crystal form. The stereoimage in Figure 7 shows a superimposition of the 1FZX structure (wheat) onto the 1D98 crystal structure (cyan). The location of neighboring DNA molecules, making contact (within 4 Å) in the crystal lattice, is shown in yellow. Strong correlation is observed between the base steps involved in packing contacts in the crystal lattice and the value of helicoidal parameters observed at these positions. The position of the C1pG2 step at the 5′ end and, more significantly, the A8pA9, A9pG10, and G10pC11 steps at the 3′ end of the dodecamer exhibit effects of crystal contacts. With reference to the superimposed 1FZX solution structure, deformation of the base pairs at the 3′ end of the DNA plays a dominant role in determining the direction of axis bend. Figure 6(A) shows that the values of bp roll observed at these steps in the original crystal structure are distinctly large compared to the NMR and MD derived solution structures.

The contacts being made (Figure 7) appear to have little effect on the A-tract structure at the center of the dodecamer. In the various crystal structures of the EcoRI dodecamer sequence CGCGAATTGCGC that crystallize in the $P2_12_12_1$ space group isomorphous to the 1D98 structure,[48] the central ApT step exhibits varying roll values. On the other hand, the central ApA step in 1D98 is straight and remains largely unchanged during the simulation. This indicates that packing-induced deformations largely occur at only those bp steps of the DNA that are intrinsically deformable,[22] such as the ApT step, but are not strong enough to distort the structure within the A-tract.

DNA Bending Model from MD Simulations

In accord with crystallographic data, MD simulations support the "non-A-tract" model of DNA curvature according to which the A-tracts are rigid units with an essentially straight helical axis that phase the curvature introduced by the highly flexible pyrimidine–purine steps in the flanking section of the DNA which act as the loci of curvature. This is consistent with other MD simulations of A-tract sequences.[36] While the pure ApA steps in the A-tract maintain a unique structure that could potentially impact the curvature in the flanking sequences, their direct contribution to the curvature in the MD model as well as the experimentally derived structures from NMR RDC data and x-ray diffraction of crystals is small compared to the prominent roll angles exhibited at the pyrimidine–purine (YpR) steps. Detailed analysis indicates that the direct contribution of the A-tract region is less than 25% of the net curvature in the DNA. It must be noted that a small writhe is characteristic of the helical DNA structure, to the extent of 5° in canonical B form and going up to 21° in A form.[49]

Role of ApT and TpA Steps in A-Tract Curvature

Recent structural studies[50–52] on A-tract–containing DNA converge on a delocalized bend model of curva-

ture[53] that ascribes small bending contribution toward the minor groove within the A-tract and more drastic changes at the pyrimidine–purine and purine–pyrimidine steps, often at the junction of the A-tract and non-Atract region. Significant support for this model comes from the structure of A_nT_n sequences such as A_2T_2 and A_4T_4 that also exhibit DNA curvature properties similar to pure A-tract sequences. In a RDC NMR spectroscopic study on the origin of curvature in A_4T_4 compared to the lack of curvature in the T_4A_4 sequence, Stefl et al.[52] observe a large negative roll of about –12° at the ApT junction and an equivalent but positive roll at the TpA junctions. The presence of the ApT step within an A-tract enhances the negative roll observed at this step. It has been shown that MD simulations on similar phased A-tract sequences[27] comparing curvature in A_4T_4 versus T_4A_4 reproduce the experimentally observed differences in the curvature of these two sequences. While the MD simulations reproduce many of the trends characteristic of the A-tract sequence, the effect of roll at ApT step, which averaged about –1° in the MD, is significantly smaller than the approximately +12° roll at the TpA steps.

MD studies on the E2 protein-binding sequence ACCGAATTCGGT reveals a similar character of the MD simulations with regard to the ApT step. In a recent structural study of this sequence by crystallography,[50] three unique structures were found in the crystal lattice with the roll values at the central ApT step being 0, −1, and −7°, respectively, signifying the propensity of this step to roll into the minor groove, while the range of values indicates the possible impact of crystal packing. Independent MD simulations on these three structures in an aqueous environment[54] result in the trajectories converging to a common dynamical form with a roll value averaging about –1° at the central ApT step. Analysis of crystallographic structures has revealed that TpA steps are more deformable in comparison to the ApT step.[55] Based on theoretical conformational mapping of the dinucleotide steps,[56] an optimized roll value of −8.1° has been noted for the ApT step, pointing to the tendency of the MD procedure to underestimate the roll at this step.

Dynamical Information from NMR versus MD Simulation

Solution of DNA structure based on NMR data is also prone to the effects of force field and methodology used for energy refinement.[9,57] The refinement protocol based on the selection of lowest energy structures confirming to the NMR data does not necessarily constitute a Boltzmann sampling of structures nor provide accurate information regarding the dynamical tendencies of the molecule. Notably, the average RMS differences within the family of 10 structures solved for two different A-tract DNA using RDC NMR data have yielded different structural properties. In the case of the 1FZX[7] structure, the RMS difference within the family of solved structures is 0.3 Å, while the average difference between the 10 structures in the A-tract sequence d(GGCAAAACGG) solved by Barbic et al.[51] is 0.9 Å. This difference could be attributed to either a lesser stiffening effect of the shorter A_4 sequence in the Barbic et al. structure compared to the A_6 or to differences in the refinement protocol.

CONCLUSION

MD model of dA_6 tract DNA structure in solution provides a description of the structure that reconciles with structural information derived from both x-ray diffraction of crystals and solution state NMR–RDC data. In the MD simulations the A-tract behaves as fairly rigid unit and exhibits very small roll or tilt angles that could contribute directly to curvature. The simulated structure resembles the structure obtained on the basis of NMR–RDC data and matches the degree and direction of curvature observed experimentally. MD simulations lend support to the model of DNA curvature that is consistent with a number of sequence-dependent structural characteristics of DNA observed in a wide range of crystallographic structures. The results point to the significant role of the flexibility of YpR steps in determining the curvature of DNA. MD simulations endorse the idea that the general sequence-directed structural deformation tendencies of DNA observed in the crystal also hold true for DNA in solution except where crystal contacts perturb the structures. The MD simulations of DNA structure in an aqueous environment are able to distinguish between attributes of the structure in crystal that appear as a result of packing effects and the intrinsic sequence-dependent structural characteristics of the DNA.

Financial support for this research was provided by NIGMS Grant GM 37909. Supercomputer time for our calculations was generously provided under the auspices of the PACI program on the facilities of NCSA at the University of Illinois at Champaign/Urbana and the Texas Advanced Computing Center at the University of Texas, Austin.

REFERENCES

1. Hagerman, P. J. Annu Rev Biochem 1990, 59, 755–781.
2. Marini, J. C.; Effron, P. N.; Goodman, T. C.; Singleton, C. K.; Wells, R. D.; Wartell, R. M.; Englund, P. T. J Biol Chem 1984, 259, 8974–8979.
3. Ulanovsky, L. E.; Trifonov, E. N. Nature 1987, 326, 720–722.
4. Koo, H. S.; Wu, H. M.; Crothers, D. M. Nature 1986, 320, 501–506.
5. Goodsell, D. S.; Kaczor Grzeskowiak, M.; Dickerson, R. E. J Mol Biol 1994, 239, 79–96 issn: 0022–2836.
6. Nelson, C. M. H.; Finch, J. T.; Luisi, B. F.; Klug, A. Nature 1987, 330, 221–226.
7. MacDonald, D.; Herbert, K.; Zhang, X. ; T., P.; Lu, P. J Mol Biol 2001, 307, 1081–1098.
8. Zinkel, S. S.; Crothers, D. M. Nature 1987, 328, 178–181.
9. Arthanari, H.; McConnell, K. J.; Beger, R.; Young, M. A.; Beveridge, D. L.; Bolton, P. H. Biopolymers 2003, 68, 3–15.
10. McConnell, K. J.; Beveridge, D. L. J Mol Biol 2001, 314, 23–40.
11. Simpson, L. Proc Natl Acad Sci U S A 1979, 76, 1585–1588.
12. Marini, J. C.; Levene, S. D.; Crothers, D. M.; Englund, P. T. Proc Natl Acad Sci U S A 1982, 79, 7664–7668.
13. Diekmann, S.; Wang, J. C. J Mol Biol 1985, 186, 1–11.
14. Ulanovsky, L.; Bodner, M.; Trifonov, E. N.; Choder, M. Proc Natl Acad Sci U S A 1986, 83, 862–866.
15. Levene, S. D.; Wu, H. M.; Crothers, D. M. Biochemistry 1986, 25, 3988–3995.
16. Griffith, J.; Bleyman, M.; Rauch, C. A.; Kitchin, P. A.; Englund, P. T. Cell 1986, 46, 717–724.
17. Rivetti, C.; Walker, C.; Bustamante, C. J Mol Biol 1998, 280, 41–59.
18. Chan, S. S.; Breslauer, K. J.; Austin, R. H.; Hogan, M. E. Biochemistry 1993, 32, 11776–11784.
19. Jerkovic, B.; Bolton, P. H. Biochemistry 2000, 39, 12121–12127.
20. Tchernaenko, V.; Radlinska, M.; Drabik, C.; Bujnicki, J.; Halvorson, H. R.; Lutter, L. C. J Mol Biol 2003, 326, 737–749.
21. DiGabriele, A. D.; Steitz, T. A. J Mol Biol 1993, 231, 1024–1039 issn: 0022–2836.
22. Dickerson, R. E.; Goodsell, D. S.; Neidle, S. Proc Natl Acad Sci U S A 1994, 911, 3579–3583.
23. Young, M. A.; Ravishanker, G.; Beveridge, D. L.; Berman, H. M. Biophys J 1995, 68, 2454–2468.
24. Haran, T. E.; Kahn, J. D.; Crothers, D. M. J Mol Biol 1994, 244, 135–143.
25. Vermulen, A.; Zhou, H.; Pardi, A. J Am Chem Soc 2000, 122, 9638–9647.
26. Tjandra, N.; Tate, S.-I.; Ono, A.; Kainosho, M.; Bax, A. J Am Chem Soc 2000, 122, 6190–6200.
27. Sprous, D.; Young, M. A.; Beveridge, D. L. J Mol Biol 1999, 285, 1623–1632.
28. Sherer, E. C.; Harris, S. A.; Soliva, R.; Orozco, M.; Laughton, C. A. J Am Chem Soc 1999, 121, 5981–5991.
29. Pastor, N.; MacKerell, A. D., Jr.; Weinstein, H. J Biomol Struct Dyn 1999, 16, 787–810.
30. Strahs, D.; Schlick, T. J Mol Biol 2000, 301, 643–663.
31. Lankas, F.; Cheatham, T. E., 3rd; Spackova, N.; Hobza, P.; Langowski, J.; Sponer, J. Biophys J 2002, 82, 2592–2609.
32. Madhumalar, A.; Bansal, M. Biophys J 2003, 85, 1805–1816.
33. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, Jr., K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. J Am Chem Soc 1995, 117, 5179–5197.
34. Foloppe, N.; MacKerell, J. A.D. J Comput Chem 2000, 21, 86–104.
35. Langley, D. R. J Biomol Struct Dyn 1998, 16, 487–509.
36. Beveridge, D. L.; Dixit, S. B.; Barreiro, G.; Thayer, K. M. Biopolymers Nucleic Acid Sci 2004, 73, 380–403.
37. Arnott, S. Polynucleotide Secondary Structures: An Historical Perspective; Oxford University Press: London, 1999.
38. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. J Chem Phys 1983, 79, 926–936.
39. Darden, T. A.; York, D. M.; Pedersen, L. G. J Chem Phys 1993, 98, 10089–10092.
40. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. J Chem Phys 1995, 103, 8577–8593.
41. Case, D. A.; Darden, T. A.; Cheatham, III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. AMBER8, University of California: San Francisco, 2004.
42. Duke, R. E.; Pedersen, L. G., PMEMD 3, University of North Carolina-Chapel Hill, 2003.
43. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. J Comput Phys 1977, 23, 327–336.
44. Ponomarev, S.; Thayer, K. M.; Beveridge, D. L. Proc Natl Acad Sci U S A (in press; 10.1073/pnas.0406435101).
45. Dickerson, R. E.; Kopka, M. L.; Pjura, P. Proc Natl Acad Sci U S A 1983, 80, 7099–7103.
46. Lavery, R.; Sklenar, H. J Biomol Struct Dyn 1989, 6, 655–667.
47. Gorin, A. A.; Zhurkin, V. B.; Olson, W. K. J Mol Biol 1995, 247, 34–48.
48. Dickerson, R. E.; Drew, H. R. J Mol Biol 1981, 149, 761–786.
49. Dickerson, R. E. Nucleic Acids Res 1998, 26, 1906–1926.

50. Hizver, J.; Rozenberg, H.; Frolow, F.; Rabinovich, D.; Shakked, Z. Proc Natl Acad Sci U S A 2001, 98, 8490–8495.
51. Barbic, A.; Zimmer, D. P.; Crothers, D. M. Proc Natl Acad Sci U S A 2003, 100, 2369–2373.
52. Stefl, R.; Wu, H.; Ravindranathan, S.; Sklenar, V.; Feigon, J. Proc Natl Acad Sci U S A 2004, 101, 1177–1182.
53. Crothers, D. M.; Shakked, Z. In: Oxford Handbook of Nucleic Acid Structure; Neidle, S., Ed.; Oxford University Press: Oxford, New York, 1999; pp 455–469.
54. Byun, S. K.; Beveridge, D. L. Biopolymers 2003, 73, 369–379.
55. Mack, D. R.; Chiu, T. K.; Dickerson, R. E. J Mol Biol 2001, 312, 1037–1049.
56. Packer, M. J.; Dauncey, M. P.; Hunter, C. A. J Mol Biol 2000, 295, 71–83.
57. McAteer, K.; Kennedy, M. A. J Biomol Struct Dyn 2003, 20, 487–506.
58. DeLano, W. L. The PyMOL Molecular Graphics System. DeLano Scientific: San Carlos, CA, 2002.

Reviewing Editor: David Case