A measure of bending in nucleic acids structures applied to A-tract DNA

F. Lankaš[1],*, N. Špačková[2], M. Moakher[3], P. Enkhbayar[4] and J. Šponer[2]

[1]Centre for Complex Molecular Systems and Biomolecules, Institute of Organic Chemistry and Biochemistry, Prague, Czech Republic, [2]Institute of Biophysics, Brno, Czech Republic, [3]Laboratory for Mathematical and Numerical Modeling in Engineering Science, National Engineering School at Tunis, Tunisia and [4]Department of Biophysics, Faculty of Biology, National University of Mongolia, Ulaanbaatar

ABSTRACT

A method is proposed to measure global bending in DNA and RNA structures. It relies on a properly defined averaging of base-fixed coordinate frames, computes mean frames of suitably chosen groups of bases and uses these mean frames to evaluate bending. The method is applied to DNA A-tracts, known to induce considerable bend to the double helix. We performed atomistic molecular dynamics simulations of sequences containing the A_4T_4 and T_4A_4 tracts, in a single copy and in two copies phased with the helical repeat. Various temperature and salt conditions were investigated. Our simulations indicate bending by roughly $10°$ per A_4T_4 tract into the minor groove, and an essentially straight structure containing T_4A_4, in agreement with electrophoretic mobility data. In contrast, we show that the published NMR structures of analogous sequences containing A_4T_4 and T_4A_4 tracts are significantly bent into the minor groove for both sequences, although bending is less pronounced for the T_4A_4 containing sequence. The bending magnitudes obtained by frame averaging are confirmed by the analysis of superhelices composed of repeated tract monomers.

INTRODUCTION

Since the first single-crystal structure of a DNA oligomer was resolved 30 years ago (1), it has become clear that the conformation of double-stranded DNA departs from a straight, uniform double helix characteristic for idealized B or A forms. DNA conformation is commonly characterized by intra-base pair local conformational parameters (buckle, propeller, opening, shear, stretch and stagger) a by inter-base pair, or step parameters (tilt, roll, twist, shift, slide and rise) (2). One of the most important structural irregularities is DNA bending. At the level of base pair steps, bending has two components, roll and tilt. Of considerable interest, however, is also global bending, that is, the bending magnitude and direction of a longer DNA fragment as a whole.

Many authors define global bending magnitude as the angle between normal vectors of base pairs at the two ends of the DNA fragment. Dickerson and co-workers (3) proposed to plot the projections of base pair normals onto a reference plane perpendicular to a suitably chosen helical axis, yielding the normal vector plot. Bending magnitude and direction are deduced from vectors connecting the projected normals. While the method provides a quick and easy-to-understand picture of the oligomer bending, its precise quantitative evaluation is made more difficult by various approximations involved in the scheme.

If the base pair normals are not available directly, they can be inferred from the local conformational parameters. The *3DNA* code by Lu and Olson (4) includes a rebuilding functionality which enables one to reconstruct the whole oligomer from the values of inter- and intra-base pair parameters. The subsequent analysis by *3DNA* of such a rebuilt structure then yields, among other things, all the base pair normals.

A related approach has been adopted in the *Madbend* code by Strahs and Schlick (5). Here, local roll, tilt and twist are used to compute quantities called global roll and global tilt, which describe both the bending magnitude and direction. The precise meaning of these quantities depends on the definition of local roll, tilt and twist used to calculate them.

While the use of base pair normals provides a rather straightforward relationship between local and global bending, it also has disadvantages. In particular, it emphasizes the influence of local bends, especially at the ends of the oligomer. Thus, methods to smooth out local bending variations have been proposed.

---

*To whom correspondence should be addressed. Tel: +420 220 410 319; Fax: +420 220 410 320; Email: filip.lankas@uochb.cas.cz

The program *Curves* developed by Lavery and Sklenar (6) yields a global curvilinear helical axis computed by a least squares optimization scheme. In the recently released *Curves +* (7), the global curvilinear axis is obtained from the average local screw axes by polynomial smoothing. The curvilinear axis as computed by *Curves* has been widely used, not only to evaluate shape but also to estimate global elastic properties of DNA (8,9). The disadvantage of such approaches is a rather complicated relationship between the local structure and global bending defined in this way.

Other approaches have been published. Goodsell and Dickerson (10) proposed a simple smoothing method based on computing arithmetic mean of base pair normals over a window of varying length. Lankas *et al.* (11) used two segments of a straight helical axis, each associated with a short fragment of the DNA double helix, to characterize the magnitude and direction of kinks in simulated minicircles. Curuksu *et al.* (12) calculated arithmetic mean of a set of screw axes, each connecting intra-strand neighbouring bases.

The most prominent base sequences that cause the DNA double helix to bend globally are A-tracts, stretches of at least four A·T base pairs without a TpA step (13–15). The bend occurs when A-tracts are inserted into a non-A-tract sequence. If more A-tracts are phased with DNA helical repeat, the individual bends add constructively. The precise magnitude of bending and details of how it is achieved at the local level differ somehow from one structure to another (16,17).

Gel electrophoretic mobility experiments performed by Hagerman (18) and Crothers and co-workers (19,20) showed that sequences containing A_{5-6} (19) and A_4T_4 (20) tracts are bent towards the minor groove at the tract centre; in contrast, T_4A_4 sequences were found to induce no appreciable bending (18). Koo *et al.* (21) found that an A_6 tract bends the DNA helix by 17–21°. The value was deduced by fitting Monte Carlo simulations of a DNA model to cyclization data. Lutter and co-workers (22) proposed a topological method to measure A-tract bending. The A_6 tract bend angle they found is in agreement with the previous estimate (21).

Several studies investigated the dependence of the A-tract induced bending on temperature and ionic environment. Koo *et al.* (23) found a decrease in bending magnitude when passing from 4°C to room temperature, and Diekmann (24) revealed that the decrease with temperature is monotonic (exceptional cases could be explained by a loss of planarity of the construct). The decrease of bending at elevated temperatures has been attributed to a premelting transition of the A-tract, well described by a two-state model (25). Diekmann (24) also found that bending decreases with increasing concentration of NaCl.

At the atomic resolution level, A-tract DNA has been extensively studied by X-ray crystallography (14,17). The observed global bending direction of A-tract sequences differs from the direction suggested by solution experiments, which may be due to crystal packing effects (26). The analysis of crystal packing effects, on the other hand, can provide valuable information on DNA structure and deformability (27).

NMR spectroscopy including residual dipolar couplings (RDC) represents a unique tool to study atomistic structure of DNA oligomers in solution. In particular, the use of RDC is essential to capture global bending (28,29). Several groups used the NMR technique with RDC to study A-tracts (16,28–30). The structure of a DNA dodecamer containing the A_6 tract obtained by MacDonald *et al.* (28) shows a global bend of 19° in a geometry consistent with earlier experiments. Barbic *et al.* (16) investigated a decamer containing an A_4 tract and found agreement with electrophoretic data and with a cyclization assay. Stefl *et al.* (30) obtained NMR structures of sequences containing the A_4T_4 or T_4A_4 tracts. They are discussed in detail below.

A-tract sequences have also been investigated by atomic resolution computer modelling. An early study based on energy minimization (31) was followed by extensive investigation using explicit-solvent molecular dynamics (MD) simulations. Beveridge and co-workers (32–36) conducted a series of pioneering studies aimed at investigating various A-tract properties. Their results on A-tract induced bending, as well as those by Strahs and Schlick (5) suggest general agreement with solution experiments. The work of Lankas *et al.* (37) indicated that the N2 amino group in the minor groove has a decisive effect on structure and mechanical properties of polypurine tracts. Other studies focused on revealing the origins of A-tract induced bending by investigating the effect of single nucleotide substitutions within an A-tract (38) and the role of ions (39).

These studies reflect the state of the art in DNA modelling of their time. The length of the simulated trajectories was in the nanosecond range, which in some cases limited the statistical quality of the collected data. The force fields used in these studies [`parm94` (40) or `parm99` (41)] were later shown to exhibit serious defects in longer B-DNA simulations, which were largely eliminated in the newly developed `parmbsc0` force field (42).

Recently, Curuksu *et al.* investigated restrained bending of A-tract containing sequences by energy minimization and by MD with umbrella sampling, using the `parm94` (12) and `parmbsc0` (43) force fields. They computed energy profiles as a function of the bending angle and found a minimum corresponding roughly to the experimental bending magnitude.

The present study develops the subject in two directions. First, we propose a new method for global bending measurement, which uses a properly defined averaging of two or more base-fixed coordinate frames. Second, we use the method to study sequences containing either A_4T_4 or T_4A_4 tracts. For this purpose, we carried out an extended set of new reference MD simulations. The length of the simulated trajectories is increased by almost two orders of magnitude compared with the early studies. We use the `parmbsc0` force field and a range of simulation conditions. We also analyse published experimental data, in particular the NMR and electrophoretic mobility data of closely related sequences obtained by Stefl *et al.* (30), and compare the structural information provided by the different methods.

METHODS

We performed explicit solvent, atomic resolution MD simulations of a set of DNA oligomers containing either A_4T_4 or T_4A_4 tracts using various temperature and ionic conditions. The simulations are summarized in Table 1. Our oligomers contain the sequences GCA_4T_4GC and CGT_4A_4CG studied by NMR by Stefl *et al.* (30) (PDB codes 1rvh and 1rvi, respectively).

The simulations were carried out using the AMBER 9 program package with the `parmbsc0` force field (42). The ions were parametrized according to Dang (44) and the SPC/E water model was used. A truncated octahedral periodic box, particle mesh Ewald technique, a 2 fs time step and the Berendsen thermostat and barostat with time constants of 5 ps were applied. The system preparation and equilibration protocol closely follows the one used in a recent extensive DNA simulation study (45), where more details can be found.

Each MD snapshot was analysed by the *3DNA* program (4). The snapshots were then filtered using two criteria. First, we checked the hydrogen bonds (H-bonds) defining a Watson–Crick base pair. We considered an H-bond broken if the donor–acceptor distance was >4 Å. Notice that any cutoff distance is artificial, since the free energy profile is in fact continuous. We found that the 4 Å cutoff allows for enough H-bond length fluctuations while capturing all the important breaking events. Second, we monitored the torsion angle γ in the backbone fragments connecting adjacent bases and considered a backbone fragment non-canonical if the value of γ was >110°. We found that excluding non-canonical γ also eliminates non-canonical torsion angles α and β. From each trajectory, only snapshots with no broken H-bond and no non-canonical base-connecting backbone fragment anywhere in the oligomer (including terminal base pairs and steps) were further analysed. Additionally, the first nanosecond of each trajectory was removed. We refer to the resulting data sets as filtered trajectories.

For each filtered trajectory, arithmetic means of the intra- and inter-base pair conformational parameters were computed. Based on these mean parameters, the average structure was rebuilt using *3DNA*. We also analysed the NMR structures of AT and TA tracts reported by Stefl *et al.* (30) (see above), denoted, respectively, by AT_nmr and TA_nmr. The average structures were obtained analogously to those from MD, using the ensemble of NMR models. All the NMR models exhibit intact H-bonds and canonical γ torsions, so that no filtering was performed.

The average structures were re-analysed by *3DNA* to obtain base-fixed frames (46), which were then used to measure global bending. We propose the following method for global bending measurement. We choose three groups of bases (called initial, middle and final group) and compute the average of the base-fixed frames in each group. To properly define the average frame, we employ the theory of averaging in the group of rotations developed by Moakher (47). The bending magnitude and direction are defined using these average frames. The method can be applied to a broad range of DNA and RNA structures. It is described in more detail in Section S1 of the Supplementary Data, and illustrated in Figure 1.

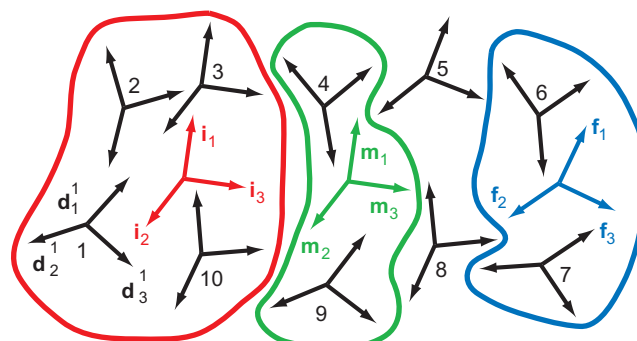The reported minor groove widths are averages of the widths from the snapshots of the filtered trajectories or

Figure 1. Illustration of the proposed bending measurement method. A hypothetical nucleic acid structure is considered, involving 10 bases. First, a coordinate frame is attached to each base, using a standard procedure (black frames). Three groups of bases are then chosen: the initial (bases 1–3 and 10, red), middle (bases 4 and 9, green) and final (bases 6 and 7, blue) group. Bases 5 and 8 do not belong to any group. For each group, the average frame is computed. The global bending magnitude is defined as the angle between the average frame vector $\mathbf{i}_3$ of the initial group and the average frame vector $\mathbf{f}_3$ of the final group. The bending direction is evaluated with respect to the average frame of the middle group. A detailed description of the method is given in Section S1 of the Supplementary Data.

Table 1. Overview of the simulations

Sequence	Code	t_{sim} (ns)	T (K)	Ions
$GGCA_4T_4GCC$	AT_short	150	300	150 mM KCl
	AT_short_neut	50	300	K^+
	AT_short_lowT	50	283	K^+
	AT_short_Na	50	300	150 mM NaCl
$CCGT_4A_4CGG$	TA_short	150	300	150 mM KCl
	TA_short_neut	50	300	K^+
	TA_short_lowT	50	283	K^+
	TA_short_Na	50	300	150 mM NaCl
$G_2CA_4T_4GCA_4T_4GC_2$	AT_long	150	300	150 mM KCl
$C_2GT_4A_4CGT_4A_4CG_2$	TA_long	150	300	150 mM KCl

Trajectory length is denoted by t_{sim}, T stands for simulation temperature. Simulations were performed either with net-neutralizing K^+ or Na^+ ions and 150 mM of the corresponding added salt (KCl or NaCl), or with net-neutralizing K^+ ions only.

Table 2. Global bending magnitude and direction

Sequence	Code	Bending magnitude (°)	Direction (°)
GGCAAAATTTTGCC	AT_short	9.9 (10.3, 10.1)	178 (190, 162)
	AT_short_neut	8.4 (8.0, 8.7)	174 (164, 182)
	AT_short_lowT	10.1 (10.6, 9.9)	182 (177, 188)
	AT_short_Na	8.7 (8.2, 9.3)	173 (180, 162)
GCAAAATTTTGC	AT_nmr	32.0 (24.5, 35.5)	180 (167, 194)
CCGTTTTAAAACGG	TA_short	1.9 (0.7, 4.0)	202 (129, 211)
	TA_short_neut	2.6	63
	TA_short_lowT	0.9 (0.9, 1.2)	27 (337, 67)
	TA_short_Na	1.4 (2.8, 4.5)	222 (324, 180)
CGTTTTAAAACG	TA_nmr	23.5 (20.2, 27.2)	176 (164, 182)
GGCA$_4$T$_4$GCA$_4$T$_4$GCC	AT_long	17.4 (18.6, 16.6)	357 (351, 4)
CCGT$_4$A$_4$CGT$_4$A$_4$CGG	TA_long	4.4 (4.3, 5.6)	30 (36, 16)

The first values are for the entire filtered MD trajectories, or for the ensemble of NMR models, values in parentheses are for the two halves of each trajectory (MD data) or the maximum and minimum value for the NMR models. Base frames at underlined locations, in the reference and the complementary strands, are used to measure bending as described in Section 2, Figure 1, and Section S1 of the Supplementary Data. Directions around 0° or 360° correspond to bending into the major groove in the oligomer centre, directions around 180° imply bending towards the minor groove. For TA_short_neut, only the first 18 ns have been analysed since the rest of the trajectory is contaminated by a flip of the backbone torsion angle γ.

NMR models as measured by *3DNA* (refined values). To assess convergence of the MD results, the filtering and analysis were repeated on the two halves of each trajectory separately.

To further validate the results on global bending, we also used another, independent method. For each of the average structures, we took the central 10-bp fragment and built a 15-mer by chaining the individual fragment copies one after another in space, using *3DNA*. For the junction steps between monomers, we took the conformational parameters averaged over the two steps flanking the 10-bp fragment in the average structure. The shape of the 15-mer is close to a superhelix in all cases. Chaining identical DNA fragments was already used by Stefl *et al.* (30). Here, we refine the procedure as follows. We fit an optimal superhelix to the DNA centerline with the program *Helfit* (48) which, in contrast with previously published algorithms, optimizes all the helix parameters simultaneously. For each 15-mer we have 14 junction points between successive monomers, which we project onto the superhelix axis. The averaged distance between the projected points defines a fragment of the optimal superhelix. The angle between the tangent vectors at its two ends provides another estimate of the global bending magnitude of the monomer unit.

Finally, we measured the bending magnitude by the bending angle of the global, curvilinear axis assigned to the average structures by the programs *Curves* (6) and *Curves+* (7). Terminal base pairs of AT_short and TA_short structures were discarded to keep the oligomer length consistent with the other bending measurements.

RESULTS AND DISCUSSION

Global bending

Table 2 compares bending magnitudes and directions for the simulated and NMR structures. Underscore marks the bases whose frames, together with those of the paired

bases in the complementary strand, were used to measure bending. They correspond from left to right to the initial, middle and final group of frames (see Section 2, Figure 1, and Section S1 of the Supplementary Data). Thus, each group includes four base-fixed frames to average. The middle group represents the oligomer centre. It follows from the definition of the base-fixed frames, and from the frame averaging procedure, that the x-axis of the middle average frame ($\mathbf{m}_1$ in Figure 1) points towards the major groove.

The bending direction close to 0° (or 360°) then implies bending into the major groove in the oligomer centre, whereas the direction ~180° indicates bending into the minor groove.

All the simulated AT oligomers are bent into the minor groove in the oligomer centre, in line with experiments on A-tracts in solution. The magnitude of bending is roughly 10° The values of bending magnitude and direction computed from the two halves of each trajectory agree with the values for the whole trajectory, which indicates good convergence of the results. The values are only weakly dependent on the simulated temperature and salt conditions, with a slight reduction of bending in the neutral solution (no added salt) and in NaCl compared with KCl solution. The TA oligomers are essentially straight, as indicated by very small bending magnitude and poorly defined bending direction.

Experiments on A-tracts show a premelting transition which changes bent structures at lower temperatures to nearly straight structures at elevated temperatures, with the transition curve centred around 30–40°C (14). Since our standard simulation temperature (300 K, or 27°C) is at the edge of this interval, we performed additional simulations at 283 K (or 10°C) to make sure that the 300 K simulations are not influenced by premelting. Moreover, 283 K is close to the temperatures at which the NMR and electrophoretic mobility experiments of Stefl *et al.* (30) were done. We observe essentially no change in bending when comparing the 283 K and 300 K simulations

(Table 2), suggesting that our results are not influenced by the premelting transition.

The results for simulated AT or TA tandem-tract structures (AT_long and TA_long) are entirely consistent with the results for the single-tract structures. In fact, the first AT tract in the AT tandem is bent by 9.1°, the second by 8.9°, both into the minor groove in their respective centres, so that the tandem as a whole is bent by roughly twice that amount but into the major groove in its centre. The TA single tracts are nearly straight, and so is the TA tandem. Thus, we observe no cooperativity in bending when passing from a single tract to a tandem.

The simulation results can be compared with the electrophoretic mobility data of Stefl *et al.* (30). In agreement with the findings of Koo and Crothers (20), we assume that the bending magnitude is proportional to the square root of the quantity R_L-1, where $R_L = L_{app}/L_{real}$ and L_{app}, L_{real} are, respectively, the apparent length of the bent sequence and its real, contour length. We take $N = 15$ in Figure 5 of Stefl *et al.* (i.e. we consider their measured sequence of 15-tract monomers). We assume a 20° bend for the A_6 tract, in agreement with cyclization experiments (21), topological measurements (22) and NMR data (28). From this value and from the quadratic relationship between R_L-1 and the bending magnitude, we obtain a 13° bend for an AT tract and 5° for a TA tract, both in NH_4Cl and KCl, and 10° for AT and 3° for TA in NaCl. The bending magnitudes of our simulated structures (Table 2) are consistent with these electrophoresis data.

In contrast, the NMR data of Stefl *et al.* (30) differ from the simulation and the electrophoretic mobility results. Data in Table 2 show that the NMR structures of both the AT and the TA tracts are significantly bent into the minor groove, although the bending is weaker for the TA tract. The range of values for the individual NMR models confirms this finding.

In order to check that our measured bending magnitudes are not unduly influenced by the bending measurement method used, we also employed the helix fitting procedure described in Section 2, which is independent of the frame averaging method. The results are shown in Table 3. We see that the simulated structures yield slim, elongated superhelices with monomer units bent by 12.5° for AT tract and 3.3° for TA tract. The bending angle is consistent with the one computed for single tracts using the frame averaging method, as reported in Table 2. Moreover, the bending angle is very close to the one we deduced from the electrophoretic mobility experiments.

Table 3. Parameters of superhelices fitted to 15-mers of the average structure fragments and bending magnitudes of a monomer fragment inferred from the fitted helices

Code	Superhelix radius (Å)	Pitch (Å)	Bending angle (°)
AT_short	22.7	333.9	12.5
TA_short	7.1	396.5	3.3
AT_nmr	55.8	117.6	30.8
TA_nmr	53.7	212.4	25.3

Finally, we measured the bending magnitude of the average structures using the global curvilinear axis computed by *Curves* and *Curves+*. The results are shown in Table 4, together with the bending magnitudes by frame averaging and helix fitting reproduced from Tables 2 and 3, respectively. Although the *Curves* and *Curves+* values somehow differ from one another and from the values obtained by frame averaging and helix fitting, they confirm the overall trend, namely moderately bent AT tract and nearly straight TA tract from simulations, and a significant bend for AT tract and to a lesser extent also for TA tract from NMR.

Thus, the various bending measurement methods yield consistent results both for the simulated and NMR data, and confirm the contrast between the simulated structures and electrophoretic mobility data on one side and the NMR structures on the other side.

Local structure

The global bending magnitude and direction result from the local structure of the DNA oligomers. Besides that, the local structure of A-tracts is also biologically important in itself (15,49).

The profiles of average roll and tilt are shown in Figure 2. The simulated values, despite the fact that they were obtained in a range of simulation conditions, are very close to each other. The only visible anomaly appears in the last two steps at the 3′-end in the TA_short_lowT structure (blue line in the right part of the figure). It is probably related to the permanently broken 3′-end pair in that structure that we allowed to pass through the filtering (see Section S2 of the Supplementary Data). The NMR and the simulated values are quite similar, but with two notable exceptions: roll in the central AT step of the AT tract (–12° from NMR and –2° from MD) and roll in the two innermost TT/AA steps in the TA tract (roughly –4° from NMR and close to 0 from MD). A detailed examination of the available X-ray and other NMR structures (29,50–54) found no conclusive evidence supporting such an extreme negative roll in the AT step within AT tracts,

Table 4. Summary of bending magnitudes computed using different methods

Code	Bending magnitude (°)			
	Frame averaging	Helix fitting	Curves (UU)	Curves+
AT_short	9.9	12.5	7.1	8.8
TA_short	1.9	3.3	4.2	1.0
AT_nmr	32.0	30.8	33.4	26.1
TA_nmr	23.5	25.3	16.0	14.4

The values from frame averaging and helix fitting discussed above (see Tables 2 and 3) are supplemented with the bending angles of the global curvilinear axis provided by *Curves* (the angle of the axis vectors, or UU) and *Curves+*. The values are to be compared with those inferred from electrophoretic mobility experiments, i.e. 13° for an AT sequence and 5° for a TA sequence (see text). The bending of the simulated oligomers (AT_short and TA_short) is consistent with the electrophoresis data. In contrast, bending of the NMR structures is much more pronounced, both for AT and for TA sequences.
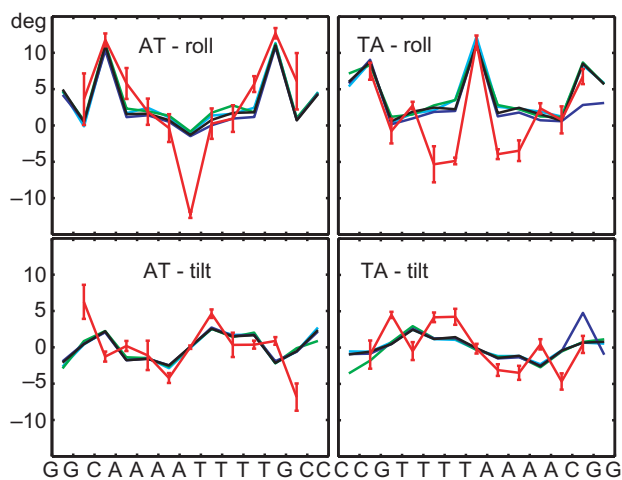
Figure 2. Profiles of average roll and tilt in the simulated and NMR structures. The lines connect vertices marking values for the individual steps, placed in the middle of each indicated step. Values for AT_small (black), AT_short_lowT (blue), AT_short_Na (green), AT_short_neut (cyan) simulations and the NMR data AT_nmr (red) are shown. The colour coding for TA tracts is analogous. Error bars represent 1 SD for the ensemble of NMR models.

Figure 3. Profiles of average minor groove width. Colour coding and error bars are as in Figure 2.

Figure 4. Profiles of average propeller twist. Colour coding and error bars are as in Figure 2.

nor the moderately negative roll in the AA/TT steps within TA tracts.

Figures analogous to Figure 2 for twist and slide are in the Supplementary Data. There the agreement is quite good for AT tract but qualitative differences are seen for TA tract. The complete sets of intra- and inter-base pair parameters for the AT_short and TA_short simulated structures are also provided in the Supplementary Data.

Other characteristic features of A-tracts include a progressive narrowing of the minor groove in the 5′ to 3′ direction of the adenine strand (55), and high negative propeller twist (14,17). The minor groove narrowing in MD structures (Figure 3) is saturated at the third step from the 5′-end of the adenine strand. Similar saturation has been observed in X-ray (56) and NMR (28) structures of the A_6 tract. The simulated propeller twist profiles (shown in Figure 4) exhibit analogous saturation at the third base pair. Although A-tracts are commonly defined as having four or more A·T pairs in a row, early experiments by Koo *et al.* (23) revealed slight electrophoretic anomaly also for A_3 tracts. Thus, A-tracts as short as A_3 may show unique local structural features. This has also been suggested in a recent study on the role of DNA shape in protein–DNA recognition (49).

The profiles of conformational parameters for the simulated AT and TA tract tandems (AT_long and TA_long) almost precisely repeat the profiles for the simulated monomers twice in a row. We demonstrate this by the minor groove width profiles shown in the Supplementary Data.

We also checked for the presence of bifurcated hydrogen bonds in the major groove of the AA steps, involving the N6 atom of adenine and the O4 atom of thymine in the next base pair along the 5′ to 3′ direction of the adenine strand. They were observed in some A-tract structures but not in others, and are not a prerequisite for
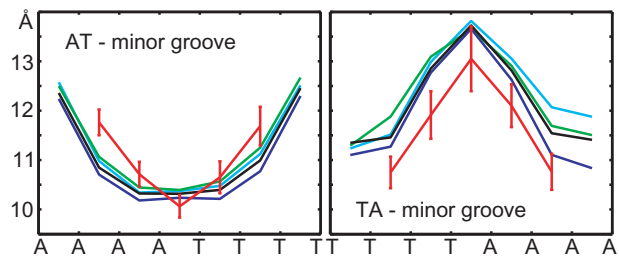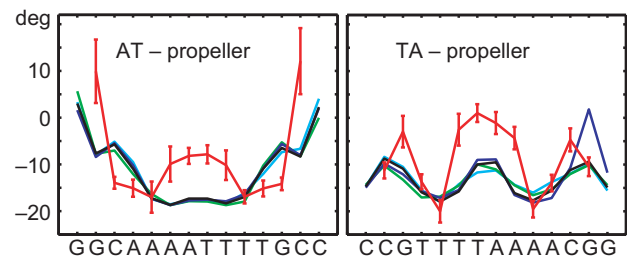
A-tract induced bending (14,17). A careful examination of the N6–O4 distances in the MD and NMR structures studied here provides no evidence for the presence of such bonds.

We finally examined the N6–N6 distances in the AT step, extensively studied by Sponer and Hobza (57) using crystal structures and quantum chemical calculations. The frequent close N6–N6 major groove contacts in AT steps were explained by activated partial sp^3 pyramidalization of the adenine amino groups (57,58). Although the force field does not account for this effect, our simulated N6–N6 distances are roughly midway between the optimal (3.26 Å) and repulsive values (3.52 Å) found by Sponer and Hobza. This indicates the overall tendency of the AT step to bring the two N6 amino groups into close contact. The NMR distances in the structures examined here (4.2 ± 0.1 Å) are too large for any contact to take place and do not agree with X-ray data.

In our analysis, we filter out snapshots with broken intra-base pair H-bonds and non-canonical backbone states (γ-flips). In this way, we limit our sampling to the B-DNA structural family. The H-bond breaks we observe are different from the base pair breathing events detected by imino proton exchange, which take place at much longer time scales. For the A_4T_4 tract, the experimental base pair lifetimes are roughly 100 ms, whereas the open state lifetimes are of the order of 10 ns (59). The fast opening events in our simulations may be related to a subtle underestimation of base pair stability by the force field, for instance due to the lack of an explicit polarization term. Nevertheless, the H-bond breaks we observe are by and large limited to terminal pairs. Indeed, the maximum population of a broken H-bond outside the terminal pairs is only 0.8%. A detailed check of the simulated

data shows that, except for terminal pairs and steps (and next to terminal ones in two cases), the influence of H-bond breaks on the averaged intra- and inter- base pair parameters is minor. This is in line with a recent simulation study (60) that, however, revealed a dramatic influence of broken H-bonds on the standard deviations of the parameters.

The situation is different in the case of γ-flips. In the crystal structures of naked B-DNA, the flips are present only in exceptional circumstances (61–63). Thus, simulations of naked B-DNA should exhibit a very low population of γ-flips. Longer simulations using the earlier `parm94` or `parm99` force fields resulted in accumulation of irreversible γ-flips that had a detrimental effect on the helical structure. Irreversible flips have been eliminated in `parmbsc0` (42), but random, infrequent reversible flips do occur. In fact, it is not possible to block the flips completely. Any realistic force field should make the γ/trans state easily accessible, since it is common in protein–DNA complexes and in RNA. However, the flips seem to be overpopulated at the time scale of our simulations. Indeed, the longest flip we observe (in TA_long) lasts for >70 ns, roughly half of the total simulation time. Once formed, a flip significantly affects local conformation: we observe, for instance, a decrease of twist and roll by 4° and of slide by 0.5 Å. Moreover, the effect is not limited to the flipped step itself but is non-local (see Supplementary Data for additional information). As a solution, we propose to filter out any shapshsot with a γ-flip, which yields a conformation in the limit of negligible population of γ-flips. Since no flip >80 ns has been observed even in the 1.2 μs B-DNA simulation (64), it is well possible that the population of γ-flips indeed becomes negligible at very long simulation times. However, for simulations of the order of 100 ns, we rather suggest to filter out the flips, since the occasional long-living flips may bias the values of structural parameters.

CONCLUSIONS

In this article, we propose a method to measure global bending magnitude and direction in nucleic acids structures. The method relies on a properly defined average of the base-fixed coordinate frames within suitably chosen groups of DNA or RNA bases. The bending magnitude and direction is then computed using the average frames of the groups.

The proposed method is applied to AT and TA tracts of double-helical DNA, which we study using our own atomistic MD simulations as well as published NMR and electrophoretic mobility data. We show that the simulated structures exhibit characteristic local conformational features and their global bending magnitudes agree with the electrophoresis data, whereas the NMR structures show much more pronounced bends. We expose similarities and differences among the results obtained from MD, NMR and electrophoresis, hoping that such a comparison may stimulate further research in the field.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Wing,R., Drew,H., Takano,T., Broka,C., Tanaka,S., Itakura,K. and Dickerson,R.E. (1980) Crystal structure analysis of a complete turn of B-DNA. *Nature*, **287**, 755–758.
2. Dickerson,R.E., Bansal,M., Calladine,C.R., Diekmann,S., Hunter,W.N., Kennard,O., Lavery,R., Nelson,H.C.M., Olson,W.K., Saenger,W. *et al.* (1989) Definitions and nomenclature of nucleic acid structure parameters. *J. Mol. Biol.*, **205**, 787–791.
3. Dickerson,R.E., Kopka,M.L. and Pjura,P. (1983) A random-walk model for helix bending in B-DNA. *Proc. Natl Acad. Sci. USA*, **80**, 7099–7103.
4. Lu,X.-J. and Olson,W.K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, **31**, 5108–5121.
5. Strahs,D. and Schlick,T. (2000) A-tract bending: insights into experimental structures by computational models. *J. Mol. Biol.*, **301**, 643–663.
6. Lavery,R. and Sklenar,H. (1989) Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.*, **6**, 655–667.
7. Lavery,R., Moakher,M., Maddocks,J.H., Petkeviciute,D. and Zakrzewska,K. (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.*, **37**, 5917–5929.
8. Lankas,F., Sponer,J., Hobza,P. and Langowski,J. (2000) Sequence-dependent elastic properties of DNA. *J. Mol. Biol.*, **299**, 695–709.
9. Mazur,A.K. (2006) Evaluation of elastic properties of atomistic DNA models. *Biophys. J.*, **91**, 4507–4518.
10. Goodsell,D.S. and Dickerson,R.E. (1994) Bending and curvature calculations in B-DNA. *Nucleic Acids Res.*, **22**, 5497–5503.
11. Lankas,F., Lavery,R. and Maddocks,J.H. (2006) Kinking occurs during molecular dynamics simulations of small DNA minicircles. *Structure*, **14**, 1527–1534.
12. Curuksu,J., Zakrzewska,K. and Zacharias,M. (2008) Magnitude and direction of DNA bending induced by screw-axis orientation: influence of sequence, mismatches and abasic sites. *Nucleic Acids Res.*, **36**, 2268–2283.
13. Hagerman,P.J. (1990) Sequence-directed curvature of DNA. *Annu. Rev. Biochem.*, **59**, 755–781.
14. Crothers,D.M. and Shakked,Z. (1999) DNA bending by adenine-thymine tracts. *Oxford Handbook of Nucleic Acid Structure*. Oxford University Press, Oxford, pp. 455–470.
15. Haran,T.E. and Mohanty,U. (2009) The unique structure of A-tracts and intrinsic DNA bending. *Q. Rev. Biophys.*, **42**, 41–81.
16. Barbic,A., Zimmer,D.P. and Crothers,D.M. (2003) Structural origins of adenine-tract bending. *Proc. Natl Acad. Sci. USA*, **100**, 2369–2373.
17. Neidle,S. (2008) *Principles of Nucleic Acid Structure*. 1st edn. Elsevier, London.
18. Hagerman,P.J. (1986) Sequence-directed curvature of DNA. *Nature*, **321**, 449–450.
19. Zinkel,S.S. and Crothers,D.M. (1987) DNA bend direction by phase sensitive detection. *Nature*, **328**, 178–181.

20. Koo,H.-S. and Crothers,D.M. (1988) Calibration of DNA curvature and unified description of sequence-directed bending. *Proc. Natl Acad. Sci. USA*, **85**, 1763–1767.

21. Koo,H.-S., Drak,J., Rice,J.A. and Crothers,D.M. (1990) Determination of the extent of DNA bending by an adenine-thymine tract. *Biochemistry*, **29**, 4227–4234.

22. Tchernaenko,V., Radlinska,M., Drabik,C., Bujnicki,J., Halvorson,H.R. and Lutter,L.C. (2003) Topological measurement of an A-tract bend angle: comparison of the bent and straightened states. *J. Mol. Biol.*, **326**, 737–749.

23. Koo,H.-S., Wu,H.-M. and Crothers,D.M. (1986) DNA bending at adenine thymine tracts. *Nature*, **320**, 501–506.

24. Diekmann,S. (1987) Temperature and salt dependence of the gel migration anomaly of curved DNA fragments. *Nucleic Acids Res.*, **15**, 247–265.

25. Chan,S.S., Breslauer,K.J., Austin,R.H. and Hogan,M.E. (1993) Thermodynamics and premelting conformational changes of phased $(dA)_5$ tracts. *Biochemistry*, **32**, 11776–11784.

26. DiGabriele,A.D., Sanderson,M.R. and Steitz,T.A. (1989) Crystal lattice packing is important in determining the bend of a DNA dodecamer containing an adenine tract. *Proc. Natl Acad. Sci. USA*, **86**, 1816–1820.

27. Dickerson,R.E., Goodsell,D.S. and Neidle,S. (1994) "…the tyranny of the lattice". *Proc. Natl Acad. Sci. USA*, **91**, 3579–3583.

28. MacDonald,D., Herbert,K., Zhang,X., Polgruto,T. and Lu,P. (2001) Solution structure of an A-tract DNA bend. *J. Mol. Biol.*, **306**, 1081–1098.

29. McAteer,K., Aceves-Gaona,A., Michalczyk,R., Buchko,G.W., Isern,N.G., Silks,L.A., Miller,J.H. and Kennedy,M.A. (2004) Compensating bends in a 16-base-pair DNA oligomer containing a T_3A_3 segment: a NMR study of global DNA curvature. *Biopolymers*, **75**, 497–511.

30. Stefl,R., Wu,H., Ravindranathan,S., Sklenar,V. and Feigon,J. (2004) DNA A-tract bending in three dimensions: solving the dA_4T_4 vs dT_4A_4 conundrum. *Proc. Natl Acad. Sci. USA*, **101**, 1177–1182.

31. Sanghani,S.R., Zakrzewska,K., Harvey,S.C. and Lavery,R. (1996) Molecular modelling of $(A_4T_4NN)_n$ and $(T_4A_4NN)_n$: sequence elements responsible for curvature. *Nucleic Acids Res.*, **24**, 1632–1637.

32. Young,M.A. and Beveridge,D.L. (1998) Molecular dynamics simulations of an oligonucleotide duplex with adenine tracts phased by a full helix turn. *J. Mol. Biol.*, **281**, 675–687.

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

34. McConnell,K.J. and Beveridge,D.L. (2001) Molecular dynamics simulations of B'-DNA: sequence effects on A-tract-induced bending and flexibility. *J. Mol. Biol.*, **314**, 23–40.

35. Beveridge,D.L., Dixit,S.B., Barreiro,G. and Thayer,K.M. (2004) Molecular dynamics simulations of DNA curvature and flexibility: helix phasing and premelting. *Biopolymers*, **73**, 380–403.

36. Dixit,S.B., Pitici,F. and Beveridge,D.L. (2004) 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. *Biopolymers*, **75**, 468–479.

37. Lankas,F., Cheatham,T.E. III, Spackova,N., Hobza,P., Langowski,J. and Sponer,J. (2002) Critical effect of the N2 amino group on structure, dynamics and elasticity of DNA polypurine tracts. *Biophys. J.*, **82**, 2592–2609.

38. Sherer,E.C., Harris,S.A., Soliva,R., Orozco,M. and Laughton,C. (1999) Molecular dynamics studies of DNA A-tract structure and flexibility. *J. Am. Chem. Soc.*, **121**, 5981–5991.

39. Mocci,F. and Saba,G. (2003) Molecular dynamics simulations of A·T-rich oligomers: sequence-specific binding of Na^+ in the minor groove of B-DNA. *Biopolymers*, **68**, 471–475.

40. 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. and Kollman,P.A. (1995) A second generation force field for the simulation of proteins, nucleic acids and organic molecules. *J. Am. Chem. Soc.*, **117**, 5179–5197.

41. Wang,J., Cieplak,P. and Kollman,P.A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.*, **21**, 1049–1074.

42. Perez,A., Marchan,I., Svozil,D., Sponer,J., Cheatham,T.E. III, Laughton,C.E. and Orozco,M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of α/γ conformers. *Biophys. J.*, **92**, 3817–3829.

43. Curuksu,J., Zacharias,M., Lavery,R. and Zakrzewska,K. (2009) Local and global effects of strong DNA bending induced during molecular dynamics simulations. *Nucleic Acids Res.*, **37**, 3766–3773.

44. Dang,L.X. (1995) Mechanism and thermodynamics of ion selectivity in aqueous solutions of 18-crown-6 ether: a molecular dynamics study. *J. Am. Chem. Soc.*, **117**, 6954–6960.

45. Lavery,R., Zakrzewska,K., Beveridge,D., Bishop,T.C., Case,D.A., Cheatham,T.E. III, Dixit,S., Jayaram,B., Lankas,F., Laughton,C. *et al.* (2009) A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucleic Acids Res.*, **38**, 299–313.

46. Olson,W.K., Bansal,M., Burley,S.K., Dickerson,R.E., Gerstein,M., Harvey,S.C., Heinemann,U., Lu,X.-J., Neidle,S., Shakked,Z. *et al.* (2001) A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.*, **313**, 229–237.

47. Moakher,M. (2002) Means and averaging in the group of rotations. *SIAM J. Matrix Anal. Appl.*, **24**, 1–16.

48. Enkhbayar,P., Damdinsuren,S., Osaki,M. and Matsushima,N. (2008) HELFIT: helix fitting by a total least squares method. *Comput. Biol. Chem.*, **32**, 307–310.

49. Rohs,R., West,S.M., Sosinsky,A., Liu,P., Mann,R.S. and Honig,B. (2009) The role of DNA shape in protein-DNA recognition. *Nature*, **461**, 1248–1253.

50. Mack,D.R., Chiu,T.K. and Dickerson,R.E. (2001) Intrinsic bending and deformability at the T-A step of CCTTTAAAGG: a comparative analysis of T-A and A-T steps within A-tracts. *J. Mol. Biol.*, **312**, 1037–1049.

51. Edwards,K.J., Brown,D.G., Spink,N., Skelly,J.V. and Neidle,S. (1992) Molecular structure of the B-DNA dodecamer d(CGCAA ATTTGCG)$_2$: an examination of propeller twist and minor-groove water structure at 2.2 Å resolution. *J. Mol. Biol.*, **226**, 1161–1173.

52. Woods,K.K., Maehigashi,T., Howerton,S.B., Sines,C.C., Tannenbaum,S. and Williams,L.D. (2004) High-resolution structure of an extended A-tract: [dCGCAAATTTGCG]$_2$. *J. Am. Chem. Soc.*, **126**, 15330–15331.

53. Wu,Z., Dalaglio,F., Tjandra,N., Zhurkin,V.B. and Bax,A. (2003) Overall structure and sugar dynamics of a DNA dodecamer from homo- and heteronuclear dipolar couplings and ^{31}P chemical shift anisotropy. *J. Biomol. NMR*, **26**, 297–315.

54. Morozov,A.V., Fortney,K., Gaykalova,D.A., Studitsky,V.M., Widom,J. and Siggia,E.D. (2009) Using DNA mechanics to predict *in vitro* nucleosome positions and formation energies. *Nucleic Acids Res.*, **37**, 4707–4722.

55. Burkhoff,A.M. and Tullius,T.D. (1987) The unusual conformation adopted by the adenine tracts in kinetoplast DNA. *Cell*, **48**, 935–943.

56. Nelson,H.C.M., Finch,J.T., Luisi,B.F. and Klug,A. (1987) The structure of and oligo(dA)·oligo(dT) tract and its biological implications. *Nature*, **330**, 221–226.

57. Sponer,J. and Hobza,P. (1994) Bifurcated hydrogen bonds in DNA crystal structures. An ab initio quantum chemical study. *J. Am. Chem. Soc.*, **116**, 709–714.

58. Yildirim,I., Stern,H.A., Sponer,J., Spackova,N. and Turner,D.H. (2009) Effects of restrained sampling space and nonplanar amino groups on free-energy predictions for RNA with imino and sheared tandem GA base pairs flanked by GC, CG, iGiC or iCiG base pairs. *J. Chem. Theor. Comput.*, **5**, 2088–2100.

59. Warmländer,S., Sponer,J.E., Sponer,J. and Leijon,M. (2002) The influence of the thymine C5 methyl group on spontaneous base pair breathing in DNA. *J. Biol. Chem.*, **32**, 28491–28497.

60. Lankas,F., Gonzalez,O., Heffler,L.M., Stoll,G., Moakher,M. and Maddocks,J.H. (2009) On the parameterization of rigid base and

basepair models of DNA from molecular dynamics simulations. *Phys. Chem. Chem. Phys.*, **11**, 10565–10588.

61. Varnai,P., Djuranovic,D., Lavery,R. and Hartmann,B. (2002) α/γ transitions in the B-DNA backbone. *Nucleic Acids Res.*, **30**, 5398–5406.

62. Djuranovic,D. and Hartmann,B. (2004) DNA fine structure and dynamics in crystals and in solution: the impact of BI/BII backbone conformations. *Biopolymers*, **73**, 356–368.

63. Svozil,D., Kalina,J., Omelka,M. and Schneider,B. (2008) DNA conformations and their sequence preferences. *Nucleic Acids Res.*, **36**, 3690–3706.

64. Perez,A., Luque,F.J. and Orozco,M. (2007) Dynamics of B-DNA on the microsecond time scale. *J. Am. Chem. Soc.*, **129**, 14739–14745.