

Critical Effect of the N2 Amino Group on Structure, Dynamics, and Elasticity of DNA Polypurine Tracts

Filip Lankaš,^{*‡} Thomas E. Cheatham III,[†] Nad'a Špačková,^{*§} Pavel Hobza,^{*} Jörg Langowski,[‡] and Jiří Šponer^{*}

^{*}J. Heyrovsky Institute of Physical Chemistry, Czech Academy of Sciences, and Center for Complex Molecular Systems and Biomolecules, 182 23 Praha 8, Czech Republic; [†]Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112-5820 USA;

[‡]Division Biophysics of Macromolecules, German Cancer Research Centre, 69120 Heidelberg, Germany; and [§]Institute of Biophysics, Czech Academy of Sciences, 612 65 Brno, Czech Republic

ABSTRACT Unrestrained 5–20-ns explicit-solvent molecular dynamics simulations using the Cornell et al. force field have been carried out for d[GCG(N)₁,GCG]₂ (N, purine base) considering guanine-cytosine (G·C), adenine-thymine (A·T), inosine-5-methyl-cytosine (I·mC), and 2-amino-adenine-thymine (D·T) basepairs. The simulations unambiguously show that the structure and elasticity of N-tracts is primarily determined by the presence of the amino group in the minor groove. Simulated A-, I-, and AI-tracts show almost identical structures, with high propeller twist and minor groove narrowing. G- and D-tracts have small propeller twisting and are partly shifted toward the A-form. The elastic properties also differ between the two groups. The sequence-dependent electrostatic component of base stacking seems to play a minor role. Our conclusions are entirely consistent with available experimental data. Nevertheless, the propeller twist and helical twist in the simulated A-tract appear to be underestimated compared to crystallographic studies. To obtain further insight into the possible force field deficiencies, additional multiple simulations have been made for d(A)₁₀, systematically comparing four major force fields currently used in DNA simulations and utilizing B and A-DNA forms as the starting structure. This comparison shows that the conclusions of the present work are not influenced by the force field choice.

INTRODUCTION

Sequence-dependent local conformational variability of DNA was first evidenced at the atomic resolution more than two decades ago (Wing et al., 1980). The structural basis of local variations, however, remains a matter of discussion. It is widely assumed that DNA local conformational variability and deformability play a considerable role in many biochemical and biological processes (Dickerson, 1999).

One of the most striking illustrations of sequence effects is the macroscopic curvature of DNA (Crothers et al., 1990; Marini et al., 1982; Diekmann, 1986; Olson and Zhurkin, 1996; Wu and Crothers, 1984). Phased runs of four to six consecutive adenines (a structure often termed as A-tract) induce a significant curvature. Despite intense research, the atomic-resolution picture of the A-tract-induced curvature remains elusive, and several models with distinct stereochemistry have been proposed (De Santis et al., 1990; Goodsell and Dickerson, 1994; Koo et al., 1986; Maroun and Olson, 1988; Olson et al., 1993). These include straight and intrinsically curved models of A-tract structures and suggest an important role played by the junctions between the A-tract and non-A-tract regions.

Fiber diffraction studies suggest that poly(dA)·poly(dT) adopts a distinct geometry that is characterized by a large propeller twist (around -26°), very narrow minor groove (~ 3.4 Å), and negative inclination of the basepairs with

respect to the helical axis (Alexeev et al., 1987; Arnott et al., 1983; Coll et al., 1987). The high-propeller twist structure of poly(dA)·poly(dT) differs from the fiber B-DNA form adopted by a common sequence. Interestingly, the same architecture as for poly(dA)·poly(dT) was suggested for poly(dI)·poly(dC) and poly(dA-dI)·poly(dT-dC) (Leslie et al., 1980).

The fiber structural characteristics of poly(dA)·poly(dT) are essentially supported by oligonucleotide crystal structures of tracts with five or six consecutive adenines, showing a propeller twist around -20° and a minor groove width of 3.5–4.0 Å (Crothers and Shakked, 1999; DiGabriele et al., 1989; DiGabriele and Steitz, 1993; Nelson et al., 1987). Crystallographic studies bolster a view that A-tracts are straight with a clearly defined spine of hydration in their narrow minor groove (DiGabriele et al., 1989; DiGabriele and Steitz, 1993; Nelson et al., 1987). Hydroxyl radical footprinting experiments (Burkhoff and Tullius, 1987) show a progressive narrowing of the minor groove in the 5' to 3' direction, an observation which appears to be supported by NMR studies (Hud et al., 1999; Chuprina et al., 1991; Katahira et al., 1990). It has also been suggested, based on NMR experiments in the presence of NH_4^+ , that the minor groove of A-tracts is partly occupied by monovalent cations (Hud et al., 1999). A very recent NMR structure of d(GGCA₆CGG)·d(CCGT₆GCC) determined with residual dipolar coupling suggests a slightly reduced propeller twist of $\sim -16^\circ$ (MacDonald et al., 2001). Interestingly, A-tract-induced curvature and likely also the high-propeller A-tract structure persist only in the low-temperature region and are abolished under pre-melting conditions (Diekmann et al., 1987; Haran and Crothers, 1989; Herrera and Chaires, 1989; Chan et al., 1993; Marini et al., 1984; Jerkovic and Bolton,

Submitted May 8, 2001, and accepted for publication February 21, 2002.

Address reprint requests to Dr. Filip Lankaš, DKFZ, H0500 Im Neuenheimer Feld 280, 69120 Heidelberg, Germany. Tel.: +49-6221-423398; Fax: +49-6221-42523390; E-mail: filip.lankas@jh-inst.cas.cz.

© 2002 by the Biophysical Society

0006-3495/02/05/2592/18 \$2.00

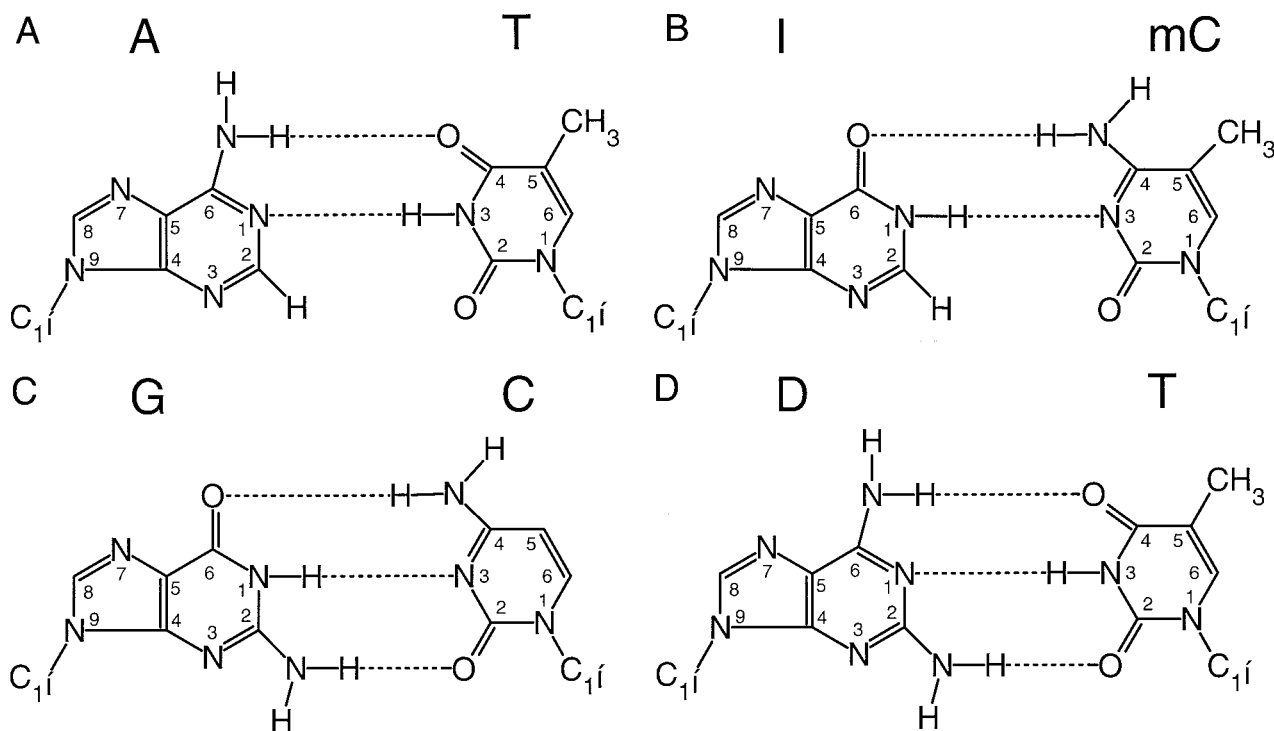


FIGURE 1 (A–D) Schematic representation of the four basepairs studied. The bottom part of each formula points to the minor groove. Note that the I-mC basepair is sterically similar to A-T while the D-T pair is isostructural with G-C. In contrast, the charge distribution and dipole moments of I-mC resemble G-C and those of D-T are similar to A-T. However, our results show that the presence or absence of the minor groove amino group rather than the electrostatic similarity decides the conformation of the tracts.

2000). This indicates that the A-tract structure is energetically close to the common B-DNA geometry.

Recent studies brought attention to unusual structural and dynamical aspects of another purine tract motif, a sequence of consecutive guanines in one strand, namely the G-tract (Dornberger et al., 1999; Ng et al., 2000; Trantirek et al., 2000). G-tracts appear to have mechanical properties contrasting the A-tract (Lankas et al., 2000). G-tracts show an increasing ability to adopt an A-DNA-like structure (Cheatham et al., 1998; Lankas et al., 2000; Ng et al., 1999, 2000; Trantirek et al., 2000) depending on the tract length and environment. G-tracts are characterized by anomalously weak intrinsic base stacking due to unfavorable electrostatic and polarization contributions (Sponer et al., 2000b).

To obtain more information about the structure and dynamics of A-tracts and their influence on the curvature, experiments have been carried out with selective substitutions of certain A-T basepairs by a 6-oxopurine/cytosine (or 5-methyl cytosine) basepair, I-C (metC) (Koo and Crothers, 1987). The I-C (mC) basepair has two H-bonds similar to the A-T basepair. However, its H-bonding is substantially stronger because the electronic distribution and molecular dipole moments of I are very similar to G (Hobza and Sponer, 1999; Sponer et al., 1996a, 2000a). Thus, the I-mC basepair is sterically isostructural with the A-T basepair (see Fig. 1, A and B), while electrostatically this basepair

strongly resembles the G-C one (Hobza and Sponer, 1999; Sponer et al., 1996a, 2000a). The latter fact is often ignored. The intrinsic stacking in an I-tract is weak due to the repulsive intrastrand electrostatics and polarization effects similar to a G-tract (Koo and Crothers, 1987; Shum and Crothers, 1983; Sponer et al., 2000a).

Interestingly, while a common base substitution within the A-tract usually leads to a drastic reduction of the curvature, the A→I substitution affects the gel migration only slightly, and AAIAA and AIAIA sequences only modestly reduce the curvature (Diekmann et al., 1987; Koo and Crothers, 1987; Mollegaard et al., 1997). This indicates that the A-tract structure is not stabilized by major groove-bifurcated H-bonds, as these cannot be formed in the case of AIAIA tracts (Diekmann et al., 1992; Koo and Crothers, 1987). Nevertheless, the AI step is capable of forming a close contact (3.1–3.2 Å) of the two amino groups across the major groove (Luisi et al., 1998; Shatzky-Schwartz et al., 1997; Sponer and Kypr, 1994). The close amino group contact, which was first identified in AT steps (Sponer and Hobza, 1994; Sponer and Kypr, 1994), can be stabilized by asymmetrical pyramidalization of the two contacted amino groups (Hobza and Sponer, 1999; Luisi et al., 1998; Sponer and Hobza, 1994).

The I-tract itself does not lead to a marked curvature (Koo and Crothers, 1987). Nevertheless, solution experi-

ments suggest, in line with fiber diffraction data (see above) that the I-tract is characterized by a very narrow minor groove similar to A-tract (Bailly et al., 1995; Diekmann et al., 1992). Therefore, it appears that the I-tract has inherently the same structure as the A-tract, while its 3' end junction is different from that of the A-tract. Here the very different electrostatic potentials of A-T and I-C basepairs seem to be important (Sponer et al., 2000a). In contrast, an oligonucleotide x-ray study shows a somewhat reduced propeller twist in an I₃mC₃ tract sequence compared to A₃T₃ and AIATmCT (Shatzky-Schwartz et al., 1997). However, more high-resolution data on longer I_n tract sequences would be necessary to achieve an unambiguous picture.

Another substitution useful in studies of A-tracts is an insertion of a 2,6-diaminopurine-uracil (thymine) basepair, D-U(T). The D-T basepair is sterically similar to the G-C basepair (see Fig. 1, C and D), and has three hydrogen bonds, but these are substantially weaker than in the G-C basepair due to the very low polarity of D (Sponer et al., 1996a). Phased D-tracts produce no curvature, the D-tract appears to have a wide minor groove, and a substitution of a single D into an A-tract has a detrimental effect on the curvature (Bailly and Waring, 1998; Koo and Crothers, 1987; Mollegaard et al., 1997). All the experimental data suggest that the amino group in the minor groove has a very significant effect on the structure of oligo and polypurine tracts. At the same time, the experiments indicate that the pyrimidine methyl group in the major groove does not have a substantial effect on the curvature (Koo and Crothers, 1987).

In the present paper we study properties of polypurine tracts (N-tracts) using large-scale molecular dynamics simulations with explicit inclusion of solvent and counterions and with an accurate treatment of electrostatic interactions (Beveridge and McConnell, 2000; Cheatham and Brooks, 1998; Cheatham and Kollman, 2000). This computational technique represents a powerful tool to study DNA structure and dynamics at the atomic resolution level. We have carried out simulations of d[GCG(G)₁₁GCG]₂, d[GCG(I)₁₁GCG]₂, d[GCG(A)₁₁GCG]₂, d[GCG(D)₁₁GCG]₂, and d[GCG(AI)₅AGCG]₂ duplexes utilizing I-mC and D-T basepairs. The five structures are designated as G-tract, I-tract, A-tract, D-tract, and AI-tract throughout this paper. Each trajectory has been extended to a comparably long time of 5 ns, and the A-, I-, and G-tract trajectories were prolonged up to 20 ns. The simulated structures contain an 11-bp N-tract (a full repeat with 10-bp steps) in its center capped by two GCG duplex segments to protect the N-tracts from end effects. The primary aim of our study is to investigate intrinsic properties of N-tracts, and mainly the effect of the molecular interactions of nucleobases on the N-tract architecture. Our simulations are not directly designed to study DNA bending, as we did not investigate either the tract phasing or tract-nontract junctions. However, a proper understanding of the intrinsic structure and dynamics of N-tracts themselves is a key step toward understanding their role in DNA bending. As will be

shown below, simulations of long tracts as utilized in this study point out their intrinsic properties, allow for a substantial increase of the effective time scale of the study, and improve convergence of the simulated trajectories.

One of the main tasks of our study was to investigate the capability of contemporary MD techniques to reproduce the intrinsic properties of the N-tracts as compared to data from x-ray, fiber diffraction, footprinting, and other experiments. Despite the great recent advances, the accuracy of current simulation techniques is still limited by the force field approximations and the nanosecond time scale of the simulations. Because the formation of the distinct architecture of the A-tract substate is likely due to a very subtle balance of different contributions, the outcome of the simulations can be affected by force field approximations, which again justifies our choice of long N-tracts. The long N-tracts allow us to examine their intrinsic properties without perturbation from proximal tract-nontract junctions, which may represent a separate difficult task for the simulation technique. Nevertheless, to address the force field issue directly, we have carried out additional simulations of d(A)₁₀ utilizing four major force fields currently used in DNA simulations (Cornell et al., 1995; Foloppe and MacKerell, Jr., 2000; Cheatham et al., 1999; Langley, 1998; Wang et al., 2000). The results confirm that conclusions of the present study do not strongly depend on the force field choice.

One-nanosecond-length simulations of a long A- and G-tract (10 bp) have been reported before as a part of an effort to investigate the B-A DNA equilibrium (Cheatham et al., 1998). Other groups attempted to characterize properties of short A-, G-, and I-tracts (Feig and Pettitt, 1998; Pastor et al., 1999; Sherer et al., 1999; Sproun et al., 1999; Trantirek et al., 2000; Young and Beveridge, 1998; Strahs and Schlick, 2000; Sponer et al., 2000b). Our study for the first time considers the D-tracts. Further, the extent of our simulation is considerably larger than in the preceding studies.

MATERIALS AND METHODS

Isothermal-isobaric (NPT) ensemble molecular dynamics were performed on each of the five 17-bp duplex oligonucleotides together with water and neutralizing counterions. The AMBER 5.0 package with the Cornell et al. (1995) force field was used. The modified bases were parametrized by means of the RESP methodology (Bayly et al., 1993; Cornell et al., 1993). At first, canonical B-DNA was built using the Nucgen module in the AMBER package; 32 sodium ions were then added to maintain charge neutrality. For this purpose, the Cion procedure of the AMBER Edit module was used, which constructs a crude 1 Å grid and then iteratively places ions at the points of least electrostatic potential. After adding the ions the whole system was immersed in a rectangular box of TIP3P water molecules (Jorgensen et al., 1983). The box dimensions were chosen to achieve a minimum distance of 10 Å from all DNA atoms to the walls, resulting in a box of ~41 Å × 41 Å × 78 Å with ~4200 water molecules. This leads to a counterion concentration of ~0.4 M (based on box dimensions). The system was then equilibrated through a series of energy minimizations and short MD runs, and heated to 300 K. The particle mesh Ewald method was used to treat electrostatic interactions. The total length of each trajectory was 5 ns (20 ns for the A-, I-, and G-tract); snapshots

were taken every 2 ps. The simulations were run on the IBM SP2 at the German Cancer Research Centre.

Due to incomplete energy conservation and small losses of energy (due to application of constant pressure, finite SHAKE tolerances, small truncations in the direct space, interaction and spline-fitting of the reciprocal interactions to a grid), when coupled with velocity scaling to maintain temperature, a small growth in center of mass translation in periodic systems can appear (Harvey et al., 1998, Chiu et al., 2000). To avoid this, it is necessary to periodically remove center of mass translation, and we did remove it every 10 ps of our simulations. However, it is not necessary to remove the rotation (and it is in fact ill-defined because it will add a net-torque to the overall system). Although rotation can be removed through the addition of weak restraints, we did not remove the center of mass rotation. This could, in principle, lead to a close approach of the periodic images. However, although our DNA rotated somewhat in the box during dynamics, as is expected (due to the relatively fast reorientation times of small duplexes in solution), not more than one or two terminal basepairs left the box for several tens of picoseconds. Since the three terminal basepairs were excluded from the analysis and their structure remained intact, we conclude that the fragment rotation did not unduly influence the structure or fluctuations.

The average structures were calculated from each trajectory (with the exclusion of the first nanosecond) using the Carnal module of the AMBER package. The SCHNAaP code (Lu et al., 1997), which implements the CEHS helicoidal analysis scheme (El Hassan and Calladine, 1995), was used to obtain conformational parameters of the averaged structures. The minor groove width was calculated as the minimum distance from the phosphorus atom in the first strand to a P atom in the second strand reduced by 5.8 Å for phosphate van der Waals radii (Lu et al., 1997). Only the central 11-bp part of each oligonucleotide was analyzed; the terminal GCG fragments were omitted from the analyses.

The elastic properties of the oligonucleotides were evaluated by analyzing the correlations of fluctuations of structural variables along the trajectories. The method is described in detail elsewhere (Lankas et al., 2000). Each snapshot was first analyzed by the CURVES conformational analysis program (Lavery and Sklenar, 1988, 1989), which has the unique property of constructing an optimal, curvilinear global helical axis of the DNA fragment. Four deformation variables were defined: the fragment length (taken as the global axis length), the total twist (calculated as the sum of twists of the individual basepair steps), and two bending angles describing the bending of the helical axis into the grooves of the central basepair and to the perpendicular direction, respectively. Ensemble correlations of the deformation variables were calculated, and the known relation between the correlations of quantities with Gaussian distribution and elements of the stiffness matrix (Landau and Lifshitz, 1980) was utilized to obtain the harmonic elastic constants, including anisotropic bending and all the coupling terms. The isotropic bending constant, A_{iso} , was calculated as the harmonic mean of the two anisotropic ones, A_1 and A_2 : $2/A_{\text{iso}} = 1/A_1 + 1/A_2$. Again, only the central 11-bp portion of the oligomers was analyzed.

A separate study was performed to investigate the force field dependence of the results. In this case, 7-ns length simulations of a model of polyA-polyT consisting of 10 consecutive basepairs were run with the Cornell et al. (Cornell et al., 1995), MacKerell all_27 (Foloppe and MacKerell, Jr., 2000), and Langley BMS (Langley, 1998) force fields using CHARMM (Brooks et al., 1983), version c26n1. Canonical models were built using the Nucgen module of AMBER 5.0. Both A- and B-forms were considered as the starting structures. Net-neutralizing ions were initially placed 4 Å perpendicularly off the phosphate oxygen bisector and then relaxed via a gas phase 1000-step adopted-basis set Newton-Raphson minimization (with the DNA held fixed and the radii of the ions artificially increased to 5.0 Å to mimic the hydration shell). This complex was then solvated with TIP3P water (Jorgensen et al., 1983) in a rhombic dodecahedron (angles 60°, 90°, 60°) with side lengths of ~50 Å. In all of the CHARMM calculations described, the particle mesh Ewald method (Darden et al., 1993) was applied throughout with fairly stringent accuracy,

including a $54 \times 54 \times 54$ FFT charge grid, an order of six for the B-spline interpolation, $\kappa = 0.34$, and a 10 Å cutoff (with the pairlist built out to 12 Å and updated heuristically). The van der Waals interactions were smoothly shifted to zero at the cutoff. During all the molecular dynamics simulations with CHARMM, a 2-fs time step, SHAKE (Ryckaert et al., 1977), and constant temperature and pressure were applied with the Langevin Piston (Feller et al., 1995) (pressure = 1.0 atm, piston mass = 500, piston gamma = 20.0) and Nosé-Hoover (mass = 1000) methods (Nosé, 1984; Hoover, 1995). The calculations were run on multiple processors of a cluster of Pentium processors (on the National Institutes of Health LoBoS I and LoBoS II computers, <http://www.lobos.nih.gov>) using a MPI parallelized version of CHARMM.

The first round of equilibration was performed with the Cornell et al. force field and involved the imposition of harmonic best-fit restraints (with a force constant of $25.0 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$) on all the DNA atoms. All other atoms were free to move. Minimization was performed to relax the water and ions and applied 250 steepest descent (sd) steps followed by 750 steps applying the adopted-base Newton Raphson (abnr) algorithms implemented in CHARMM. After this, 5 ps of dynamics were performed ramping the temperature from 50 to 300 K in 1-ps intervals, followed by an additional 20 ps of dynamics at 300 K at constant pressure and temperature. This was followed by 500 steps of abnr minimization and an additional 50 ps of dynamics at 300 K. This final set of coordinates was used as the starting point for each of the force fields and followed by further equilibration that involved rounds of minimization (250 abnr steps) and 5 ps of dynamics at 300 K, with the harmonic best-fit restraint force constant set at 20.0, 15.0, 10.0, and 5.0 kcal/mol/Å² and then one final round without any restraints applied. After this, unrestrained production dynamics were started under the same conditions as described above. Center of mass translation was removed every 5 ps. Helicoidal parameters were calculated for the average structures corresponding to five 2-ns portions of each trajectory (1–3, 2–4, 3–5, 4–6, and 5–7 ns). An average over all the basepairs in the five structures was then calculated. Besides that, the standard deviations of helicoidal parameters were computed using the 1–7-ns portion of each trajectory from each configuration at 1-ps intervals. Dials and Windows were used for the calculation of helicoidal parameters; these are very similar to the values calculated via SCHNAaP. The reported minor groove widths are averages over all frames, at 1-ps intervals, for the five internal closest phosphate-phosphate distances across the minor groove, minus 5.8 Å for the phosphate radii.

In addition, 7-ns-length simulations of the 10-mer polyA-polyT were run in AMBER version 6.0 using the parm99.dat force field starting from Nucgen-generated models of A-DNA and B-DNA. Ions were placed via LEaP using a crude electrostatic potential estimation and the DNA was solvated into truncated octahedral unit cells with sides of ~55 Å. A less stringent equilibration protocol was used that involved 1000 steps of steepest-descent minimization and 100 ps of dynamics (ramping the temperature from 100 to 300 K in the first 10 ps) with the DNA held fixed. This was followed by unrestrained minimization of 1000 steps and then production dynamics at 300 K. In all the simulations, the particle mesh Ewald method was applied with a $54 \times 54 \times 54$ FFT charge grid, order = 4, $\kappa = \sim 0.30768$, and a 9 Å cutoff. The pairlist was updated heuristically and build out to 10 Å. A 2-fs time step was applied and temperature and pressure were maintained with Berendsen coupling (Berendsen et al., 1984), with both coupling times of 1 ps. Center of mass translation was removed every 5 ps (Harvey et al., 1998, Chiu et al., 2000).

RESULTS

RMSD and geometries of the N-tracts

Analysis of root-mean-square deviations

Let us start the discussion by evaluating the global structural characteristics of the simulated N-tracts ($N = A, I, G, D,$

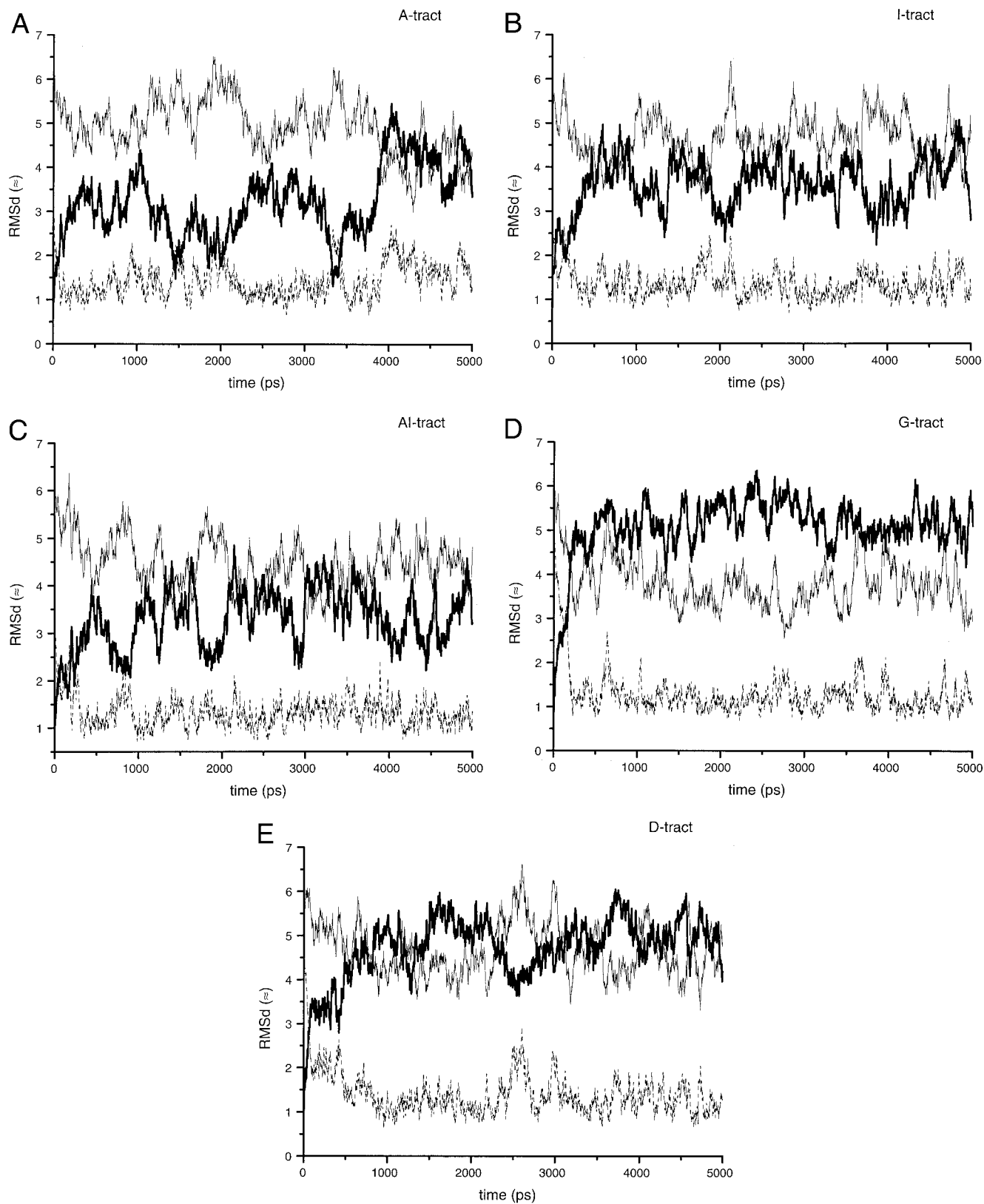


FIGURE 2 (A–E) Root-mean-square deviations (RMSD) of the atomic positions along the molecular dynamics trajectories. Only the central 11-bp portions of the oligomers, i.e., the tracts under study, were analyzed. The GCG ends were excluded from the analysis. RMSD with respect to canonical B-DNA (*thick line*), canonical A-DNA (*thin line*), and with respect to the average structure from the last 4 ns of the simulation (*dashed line*) are shown.

TABLE 1 Root-mean-square deviations

	A-Tract	I-Tract	AI-Tract	G-Tract	D-Tract
B-DNA	3.0	3.4	3.2	5.1	4.8
A-DNA	4.7	4.5	4.2	3.5	4.5

Root-mean-square deviations (in Å) of the central 11-bp segments of the averaged MD structures (1–5 ns) with respect to the canonical B-DNA and A-DNA forms. The canonical B-DNA form was used as the starting structure for the simulations.

AI). Fig. 2 shows the time-development of the root-mean-square-deviations (RMSD) of the atomic positions of the simulated molecules from the canonical B-DNA (*thick line*) and A-DNA (*thin line*). The figure further shows (*dashed line*) the development of the RMSD of the simulated structures with respect to the averaged structure (averaging taken over the 1–5 ns interval). This quantity helps us to estimate the internal stability of the trajectory (Spackova et al., 2000). All RMSD values are calculated considering the inner 11 bp only, disregarding the nontract sequence at either end. Table 1 presents RMSD between the averaged N-tract structures and canonical DNA forms. The RMSD between canonical A and B-form models is ~ 6.3 Å.

It is obvious that the A-tract remains closer to the canonical B-DNA form throughout the entire simulation. During the last nanosecond the RMSD value with respect to the canonical B-form appears to increase. However, a helicoidal analysis of the MD structure averaged over the last nanosecond shows that the A-tract still maintains essential B-DNA features and does not shift toward A-DNA canonical form (data not shown). Furthermore, the RMSD values calculated with respect to the averaged simulated structure (*dashed line*) are in general much lower than the RMSD

with respect to the A- or B-form and suggest that the structures fluctuate about a state distinguishable from both A- and B-DNA.

The I-tract and AI-tract show development of the RMSD along the trajectory very similar to the A-tract. In contrast, the G-tract shows a swift transition toward an A-like structure within the first 0.5 ns and remains in that conformational region throughout the rest of the simulation. Finally, the D-tract shows similar RMSD values with respect to both canonical A- and B-forms, both being relatively high. There also appears to be some substantial structural transition occurring during the first 0.5 ns. The RMSD values with respect to the averaged structures are low for both G- and D-tracts and confirm stability of the trajectories.

Conformations of the N-tracts

Table 2 summarizes the average basepair and basepair step structural parameters of the central 11-bp segments in the averaged MD structures (averaging taken over the 1–5 ps interval). The parameters most important to highlight differences among various N-tracts are shown in bold. The stereoviews of the averaged structures are in Fig. 3, where only the 9-bp central part is shown.

Table 2 clearly shows that the major difference among N-tracts concerns the propeller twist. A-tract, I-tract, and mainly AI-tract are characterized by a large and clearly developed propeller twist in all their steps. G-tract and D-tract structures show essentially no propeller twist. Thus, the magnitude of propeller twist is clearly determined by the presence or absence of the amino group in the minor groove. It should be noted, however, that the propeller twist ob-

TABLE 2 Mean helicoidal parameters of the central 11bp segments of the averaged MD structures

	A-Tract	I-Tract	AI-Tract	G-Tract	D-Tract	Starting
Shear	0.17	−0.10	0.04	−0.01	−0.03	0.00
Stretch	5.54	5.54	5.55	5.58	5.63	5.60
Stagger	−0.19	−0.21	−0.04	−0.20	−0.24	0.04
Buckle	−1.25	−0.04	−0.12	−4.61	−1.79	0.75
Propeller	−12.37	−13.44	−15.75	−2.34	−3.25	4.23
Opening	7.10	8.54	7.66	4.26	3.90	1.25
Shift	−0.10	0.05	−0.06	0.15	−0.04	0.06
Slide	−1.10	−1.04	−0.86	−1.70	−1.71	−0.19
Local rise	3.34	3.38	3.33	3.33	3.40	3.38
Tilt	−1.41	−0.38	−1.28	0.97	−1.01	0.13
Roll	0.70	2.62	2.87	4.29	1.95	−3.66
Local twist	31.49	30.71	31.83	28.66	28.09	35.82
X-disp	−2.00	−2.47	−2.11	−4.52	−4.13	0.23
Y-disp	0.09	0.10	0.07	−0.07	−0.02	−0.09
Global Rise	3.33	3.28	3.22	3.05	3.30	3.38
Inclination	0.61	3.89	5.88	8.50	3.42	−5.93
Tip	2.30	0.15	1.44	−2.00	1.90	−0.10
Global Twist	31.36	30.64	31.85	28.99	28.41	36.00

The length values are in angstroms, angles are in degrees. The average structures were calculated from the last 4 ns of the trajectories. The SCHNAAP program (Lu et al., 1997) adopting the CEHS scheme (El Hassan and Calladine, 1995) was used to calculate the structural parameters. Parameters showing substantial differences are in bold.

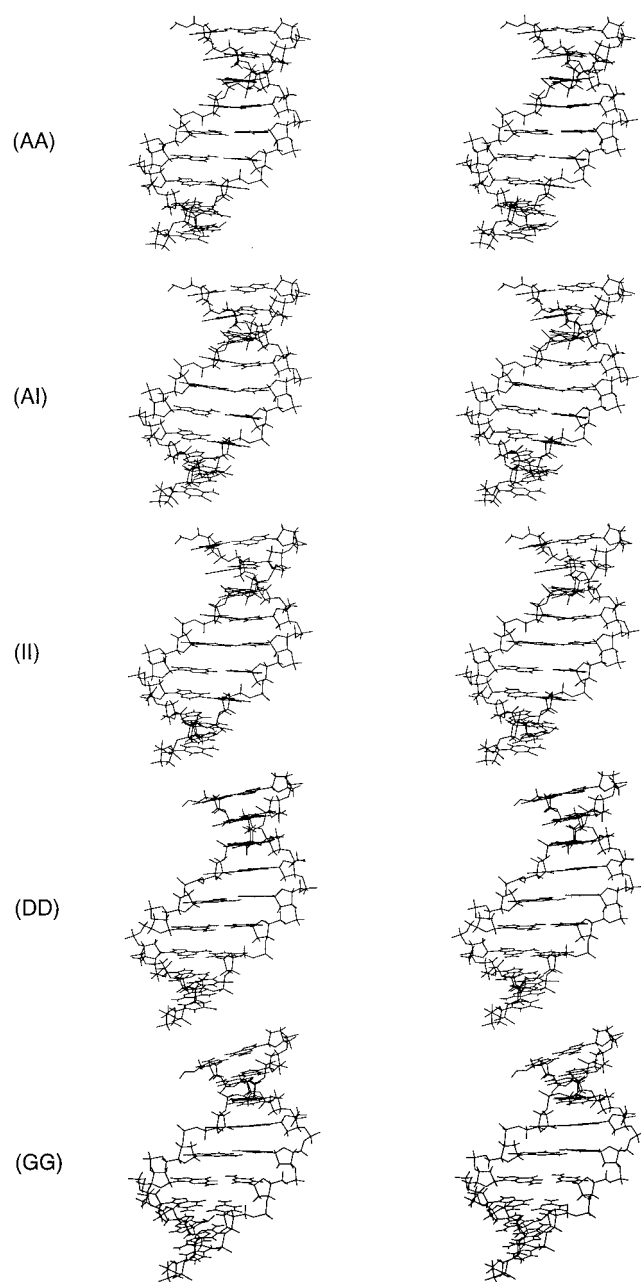


FIGURE 3 Stereoviews of the average structures for the five tracts. Average structures were calculated from the last 4 ns of the simulations. Only the central 9 bp from each tract are shown.

served in the simulated A-tract (around -12°) is lower than that revealed by x-ray and fiber studies. However, the A-tract NMR structure refined with residual dipolar data shows a propeller twist around -16° , that is, closer to the MD values compared to crystallographic data (MacDonald et al., 2001). It is interesting to note, however, that our MD simulations show considerably more uniform values of propeller twist and other local conformational parameters along the A-tract compared to these NMR data. The tendency of simulations to provide a slightly lower than expected pro-

pellor twist is well-established (Cheatham and Kollman, 2000). The major groove cross-strand distances between N6(A) and O4(T) in the molecular dynamics structures are in the range of 4.0–4.4 Å, i.e., considerably longer than in the crystals and NMR structures (MacDonald et al., 2001) where a formation of bifurcated H-bonds has been suggested. Although on average the distances are longer, during the simulations contacts below 3.5 Å are observed locally (at individual basepair steps), which suggests that transiently these bifurcated hydrogen bonds may be forming. However, given their relatively short duration (as evidenced by the absence of this structural element in the average structures), this suggests that these interactions are not a major stabilizing element. Our result is in agreement with other MD studies (Strahs and Schlick, 2000). Note, however, that the role of bifurcated H-bonds in the formation of A-tract structures is not unambiguously accepted (Diekmann et al., 1992). Furthermore, it has been suggested that a formation of bifurcated H-bonds would be accompanied by a slight sp^3 hybridization of the amino group nitrogen atoms, a structural organization that is not possible to capture with the currently applied nonpolarizable force fields, which assume specific hybridization states and do not explicitly model lone pairs (Sponer and Hobza, 1994). Despite the inability of the current force fields to model this inherent distortion of the amino group, when these contacts are observed there is a clear distortion of the amino group out of the base plane. With a force field that has an ability to model the change in hybridization state, more out-of-plane distortion may be evident, which may further stabilize these contacts.

Another interesting parameter is the helical twist. The helical twist appears to be underestimated in all structures, yet all three propeller-twisted tracts have a clearly larger helical twist compared with the D-tract and G-tract. While the absolute value of helical twist is perhaps not satisfactory, the relative differences among the different structures appear meaningful. Again, the tendency of simulations to provide a slightly lower than expected helical twist, especially with the presently used force field, is known and has been recently discussed (Cheatham et al., 1999). Although the overall twist magnitude is lower than expected, the polyA and polyA-like sequences do have a relatively higher twist compared with polyG or a general sequence. This has been previously observed with this force field (Cheatham et al., 1998) and is consistent with experiment (Peck and Wang, 1981; Rhodes and Klug, 1981; Strauss et al., 1981).

Another important local basepair step parameter is slide. Slide is within the range of -0.9 to -1.1 Å for the propeller-twisted tract structures, while it is considerably larger (in absolute value) for the G-tract and D-tract. The increase of slide is typical for a transition of the structure in the direction toward the A-form. This is confirmed by evaluating global parameter X-displacement, which is related to slide.

A-like geometry of a double helix is also characterized by inclination of the basepairs with respect to the global helical axis, and (considering a local conformational frame) by a positive basepair roll (see, for example, Sponer and Kypr (1991); Wahl and Sundaralingam (1997)). Furthermore, A-like geometry shows a reduced rise calculated along the helix axis. Note, however, that vertical distance of basepairs in a local reference frame is unchanged as a result of base stacking forces; see Sponer and Kypr (1993). Table 2 shows that only the G-tract is shifted toward the A-form considerably, though still being far from a classical A-DNA structure. The D-tract, despite its A-like slide, has other parameters more consistent with the B-form.

Variation of the structural parameters along the tracts

Table 2 summarizes the main structural parameters of the complete N-tracts as an average over all the 11 internal basepairs and basepair steps. When investigating the individual values for each step or basepair, due to the sequence homogeneity within a given N-tract, one should expect only subtle local variations in the values. If large variations are seen, this indicates that either the simulations were not sufficiently long to convert from the starting structure to a representative substate or they were not of sufficient length to average over representative substates. (Note that conversion between the substates may take longer than a few nanoseconds; see Feig and Pettitt, 1998). The RMSD data shown in Table 1 suggest that the structures do converge to a representative structure within the 5-ns time scale, therefore any local variations are likely due to transitions among various conformational substates.

The averaged structures of G-, D-, and A-tracts show very minor variations of the helical twist of 1 to 3° along the whole tract. In case of the I-tract structure, we have observed a considerable variation of twist in the second part of the I-tract, where steps 10–12 show global helical twists of 33.9, 26.6, and 33.6°, respectively. This variation is reflected also in the negative shift of the three basepairs (−0.04, −0.10, −0.23 Å) and reduced slide (−0.98, −0.99, −0.60 Å). The local rise of the 11th basepair is high (3.47 Å) and changes occur also in the backbone angles α (−69.2, −65.9, −71.5°; tract average $68.0 \pm 1.5^\circ$) and γ (53.8, 52.3, 51.9°; tract average $54.2 \pm 1.6^\circ$). This likely is connected with a development of some local conformational substate at the beginning of the simulation in that region that is not repaired during the 5-ns simulation (cf. Discussion in Feig and Pettitt (1998) and Spackova et al. (2000)). Global twist in all remaining steps in the I-tract is within the range of 29.1 to 31.8°. The minor structural variation in steps 10–12 of the I-tract is the only significant local structural inhomogeneity observed in an otherwise homogeneous sequence in these simulations, suggesting that the 5-ns time frame for sampling was sufficient.

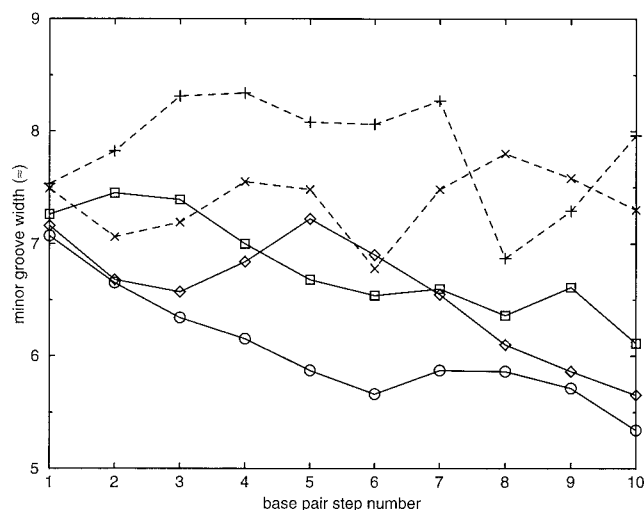


FIGURE 4 Minor groove widths of the five tracts at individual basepair steps, shown in the 5' to 3' direction of the first strand. The widths were calculated as the minimum distance from the P atom in the first strand to a P atom in the second strand, reduced by 5.8 Å to account for the phosphorus van der Waals radii. A-tract, circles; I-tract, squares; AI-tract, diamonds; G-tract, plus sign; D-tract, ×.

The AI-tract exhibits a clear sequence-dependence of the helical twist. The average helical twist of the AI step is 33.3°, while average twist of the IA step is 30.4°, both showing only very minor variations along the tract. This difference is also reflected by the redistribution of basepair roll between the two steps: 0.7° in the AI steps and 5.0° in the IA steps.

Another interesting structural effect concerns the propeller twisting along the tracts. The propeller twisting along the tracts is quite uniform, again showing that the simulations reached the basic convergence. However, the averaged propeller twist in the second half (pairs 10–14) of the propeller-twisted tracts is systematically larger than in the first part (pairs 4–8). The difference is 1.8°, 1.2°, and 1.3° for A-, I-, and AI-tracts, respectively. The largest propeller in the A- and I-tracts occurs for the 13th basepair, while the difference of propeller between the 13th and 5th basepairs is 3.2° and 5.0°, respectively.

We did not analyze the structure of the three terminal GC basepairs at either end because their behavior may be affected by the end-effects. Nevertheless, it is interesting to see that the GC basepairs immediately adjacent to the A-, I-, and AI-tracts have propeller twists of −8° to −10°, i.e., reduced compared to the tract. The same basepairs adjacent to the nonpropellered D- and G-tracts show smaller propellers within the range of −3° to −8°.

Minor-groove width and sugar puckers

Fig. 4 shows the minor groove width for individual basepair steps of all the sequences studied. The width was measured

TABLE 3 Base stacking energies

	Total	Pur-Pur	Pyr-Pyr	Pur-Pyr (3')	Pur-Pyr (5')
G-tract, (GpG)(CpC)	-12.2 (+4.4)	-2.3 (+5.3)	-2.2 (+2.1)	-4.7 (-0.3)	-3.0 (-2.7)
I-tract, (Ipl)(mCpmC)	-12.5 (+4.2)	-1.7 (+5.4)	-2.3 (+3.7)	-5.1 (-2.1)	-3.4 (-2.9)
A-tract, (ApA)(TpT)	-14.4 (+2.3)	-7.0 (+0.2)	-2.9 (+2.7)	-3.2 (-0.2)	-0.9 (-0.4)
D-tract, (DpD)(TpT)	-14.4 (+2.2)	-7.5 (-0.4)	-2.1 (+2.4)	-4.3 (+0.4)	-0.5 (-0.2)
AI-tract, (Apl)(mCpt)	-17.1 (-0.2)	-8.5 (-1.5)	-7.3 (-1.0)	-1.4 (+1.1)	+0.2 (+1.1)
AI-tract, (Ipa)(TpC)	-11.9 (+4.9)	-8.2 (-0.7)	-4.7 (+1.2)	-0.4 (+2.4)	+1.5 (+2.1)

Base stacking energies (kcal/mol) of NpN basepair steps in the averaged simulated structures (first column) and their decomposition into the four individual base-base contributions. Values in parentheses represent the electrostatic component of stacking. The van der Waals component of stacking (not shown) can be obtained by subtracting the electrostatic term from the base stacking energies. The last two columns show the interstrand stacking energies between bases at 3' and 5' ends of the step, respectively

in the 5' to 3' direction of the first strand and was defined as the minimal P-P distance reduced by 5.8 Å to account for phosphorus van der Waals radii. While the values are all between 7 and 7.5 Å at the first basepair step, their distribution along the tract differs substantially with base sequence. The A-tract exhibits almost a linear decrease of the minor groove width down to 5.7 Å at the 6th step, ending at a value of 5.3 Å at the 10th step. The narrowing of the A-tract minor groove has been observed both in experiment (Burkhoff and Tullius, 1987; Hud et al., 1999) and in simulations (Young et al., 1997a). The I- and AI-tracts also show such a narrowing along the tract, though less pronounced and with some variations that might reflect the limited simulation time. However, G- and D-tracts do not undergo any overall groove compression.

Finally, the sugar puckers, averaged over both strands of an N-tract, follow the general tendency outlined above: the pucker of A-, I-, and AI-tracts (on average 118–119°) is consistently higher than that of G- (100°) or D-tracts (112°), confirming the general shift toward the A-form of G- and to a lesser extent also of D-tract. The simulated structures show systematic differences in sugar pucker in purine and pyrimidine. In A-, I-, and AI-tracts the pucker of purine strands is larger by ~10–15° compared with the pyrimidine strands. This is in agreement with literature data (Strahs and Schlick, 2000; MacDonald et al., 2001). This difference is reduced to ~5° in the D-tract, while in the G-tract an opposite difference (puckering angle higher in the pyrimidine strand) of ~23° is seen. Let us point out that the Cornell et al. force field has a slight imbalance of the sugar pucker parameters, as extensively discussed in the literature (Feig and Pettitt, 1998; Cheatham et al., 1999). Proper balance of sugar puckering is one of the major problems of the second-generation force fields relying on constant charges and pairwise additivity and, to our best knowledge, none of the available force fields is truly perfect in this respect. This is one of the reasons why we decided not to discuss the sugar-pucker issue in detail. The simulations should nevertheless be able to reflect relative differences among various sequences, this being the main task of our effort. Strahs and Schlick (2000) reported a large positive roll of ~12° at the 5' end CA and GA steps in their A-tract

simulations. In contrast, in our A-tract simulations, the roll of the 5' end GA step is smaller, ~6°. The 5' end GI step in our I-tract simulation shows a roll value of 5°, while the GA step of the AI-tract shows a basepair roll value of 7°. It should be noted, however, that in both studies the termini of the A-tracts are already rather close to the ends of the simulated molecules and might be subject to end-effects.

Base stacking energies

Table 3 summarizes intrinsic stacking energies of the NpN steps in the simulated polypurine tracts. The stacking energies are calculated using the Cornell et al. (1995) force field utilized in the simulations. The sugar-phosphate units are replaced by hydrogen atoms while charges of these hydrogens are adjusted to keep the bases neutral. The table shows mean values of base stacking energies of the inner three steps in the averaged simulated structures. The stacking energies are calculated assuming a dielectric constant of 1 and thus show the intrinsic stacking of bases in these geometries with no solvent screening (Sponer et al., 2000b). The base stacking energy consists of two parts. The van der Waals component of stacking (not shown) is overlap-dependent and includes the dominating dispersion attraction as well as steric effects. The electrostatic component of stacking (values in parentheses) reflects the mutual interaction of molecular electrostatic potentials of bases (Hobza and Sponer, 1999; Sponer et al., 1997, 2000a). Note that the Cornell et al. force field provides an excellent approximation of high-level *ab initio* calculations of base stacking (Hobza et al., 1997; Sponer et al., 1997), although the force field obviously neglects polarization effects (Sponer et al., 1997, 2000b).

Table 3 clearly shows that the GpG stacking is rather weak and is characterized by a substantial intrastrand electrostatic repulsion (Sponer et al., 1997, 2000b). It has been suggested that this salient stacking might play a role in establishing the distinct features of G-tracts (Sponer et al., 1997, 2000b; Ng et al., 2000; Lankaš et al., 2000). However, the Ipl step has an almost identical balance of individual intrinsic stacking energy contributions as the GpG step,

while the ApA step has entirely different distribution and magnitude of the electrostatic stacking energy terms. Considering the fact that the simulated Ipl steps adopt almost identical geometries as the ApA steps, we can conclude that the geometry of NpN tracts is not primarily determined by the electrostatic component of stacking. The DpD stacking decomposition is very similar to the ApA one. The structure of DpD steps is, however, much closer to the GpG steps, and this confirms the role of the amino group protruding into the minor groove as well as the likely limited effect exerted by the pyrimidine methyl group. Finally, the ApI and IpA steps of the AI-tract show rather significant difference of their individual electrostatic stacking energy terms, resulting also in different total stacking energies. This is accompanied by rather subtle local conformational variations (see above). Thus our results indicate that the presence of efficient solvent screening reduces the role of the electrostatic components of stacking (Florian et al., 1999; Sponer et al., 2000b; Friedman and Honig, 1992; Luo et al., 2001). Note that solvent screening is capable of eliminating electrostatic repulsion even between two consecutive protonated basepairs, at least in some DNA architectures (Gallego et al., 1999; Spackova et al., 1998).

In contrast to the variable electrostatic stacking terms, the van der Waals contributions (not shown in the table) are essentially independent of the sequence. Their total values are highly conserved, ranging from -16.6 to -16.9 kcal/mol for the six different NpN steps. There are moderate variations in the Pyr-Pyr and Pur-Pyr(3') terms, correlated with the presence or absence of the N2-amino group.

In addition to the base-base interactions it would be possible to analyze other contributions to the potential energy as well, e.g., sugar-base interactions. However, the interbase interactions most probably contain the majority of sequence-dependent effects.

Interaction of the tracts with monovalent ions

Several preceding MD studies have revealed that monovalent cations are capable of penetrating into the DNA grooves to establish direct interaction with nucleobases and even to affect the fine aspects of the DNA local architecture (Hamelberg et al., 2000, 2001; Spackova et al., 2000; Stefl and Koca, 2000; Young et al., 1997a, b; McConnell and Beveridge, 2001; Cheatham and Young, 2001). Specifically, the penetration of monovalent cations into the minor groove at ApT B-DNA steps has been studied in detail (Hamelberg et al., 2000, 2001; Stefl and Koca, 2000; Young et al., 1997a, b; McConnell and Beveridge, 2000). Of special importance also is the stabilization of DNA quadruplex molecules by direct interaction of monovalent cations with the guanine quartets (Spackova et al., 1999, 2001; Stefl et al., 2001). Thus, we have systematically monitored the direct interactions of sodium cations with nucleobases in the course of our simulations in the period 3–5 ns (we omitted

the first three nanoseconds of simulations to provide a little more time for the ions to equilibrate before analyzing their positions). To our surprise, we found almost no interaction of sodium cations with the minor groove edges of bases in our simulated N-tracts. We have found only a single such event of a cation coordination lasting ~ 0.1 ns (with a cation interacting directly with O₂ atoms of cytosines during the G-tract simulation). Instead, we have found more frequent penetration of cations into the major groove. For the A-tract we found (in the period of 3–5 ns) six major contacts lasting 0.1–0.6 ns (~ 0.25 ns on average) distributed over the whole tract and involving six distinct cations. For the D-tract, we have evidenced four major contacts (0.1–0.35 ns) involving three distinct cations, and for the I-tract, four major contacts (0.1–0.2 ns) involving three different cations. More major-groove direct cation-base contacts have been seen for the G-tract (13 contacts on a time scale of 0.1–0.55 ns involving 7 cations) and AI-tract (12 contacts lasting again 0.1–0.55 ns and involving 7 different cations). The direct major-groove cation-base contacts preferably involve the N7 position of purine bases and less frequently the carbonyl oxygen atoms, when available. Despite the fact that direct cation-base contacts in the major groove are more frequent, we do not believe they critically influence the structure of the simulated N-tracts, because for the major portion of time the bases remain uncontacted with the cations. Furthermore, while the cation-base contacts are most abundant in case of the AI- and G-tracts, these molecules adopt substantially different architectures.

To further investigate the interaction of the ions, the A-tract, G-tract, and I-tract simulations were extended to 20 ns. Analysis of the ion interaction over the time range of 5–20 ns confirmed our previous observations. Specifically, very little direct interaction of sodium ions was observed directly to bases in the minor groove (defined as interaction of 3.5 Å or less to N3, O₂, and N2-H22 atoms) for each of the sequences. The only consistent set of interactions in the minor groove observed are with the phosphate O1P atoms (which point into the minor groove) where direct ion interactions are observed throughout the sequence with occupancies in the 5–15% range (with lifetimes on the order of 50 ps or less and maximum lifetimes of 300–400 ps). Consistent with the analysis over shorter time frames, most of the ion interaction (beyond direct interaction with the phosphates) is transient interaction with bases (primarily N7 atoms) in the major groove. Occupancies in the range of 5–10% are common throughout the sequence with measured lifetimes of typically < 200 ps with frequent exchange of ions.

The absence of cations in the minor groove of the A-tract might appear surprising, especially given the observations of direct ion interaction seen by others with the same force field and equivalent simulation protocol (Hamelberg et al., 2000; Young et al., 1997a; McConnell and Beveridge, 2001; Cheatham and Young, 2001). However, ion localiza-

tion is likely influenced by issues including the starting geometry (which in our case involved placing the ions initially at favorable electrostatic potential regions), the cation concentration (around 350 mM), the ion parameters, and the time-length of the simulation. To investigate issues related to the time scale and initial ion placement we extended the simulations to 20 ns. In the longer simulations we continued to observe very low direct occupancy of Na^+ ions in the minor groove, along with significant exchange of the ions, suggesting that the initial ion placement and short (3–5 ns) time scale was not an issue. This could be a factor in the earlier ApT studies, where ion localization at the ApT steps was seen very early in the simulations (Hamelberg et al., 2000; Young et al., 1997a); however, a higher occupancy at these steps is expected due to its low negative electrostatic potential (Lavery and Pullman, 1985). A more likely cause for low ion occupancy in the current set of simulations is the low cation concentration. Although the cation concentration in the simulations is on the order of 350 mM, this only represents enough salt to directly neutralize the DNA and more realistically represents the case of “no added salt.” Without additional salt, direct interaction with the minor groove bases may be disfavored. This is consistent with recent work investigating similar A-tract sequences (McConnell and Beveridge, 2001), where in significantly shorter (3 ns) simulations low (<2.5%) occupancies were observed at ApA steps in the minor groove under similar, net-neutralizing, salt conditions. It is also consistent with observations made by Strahs and Schlick for their A6 tract simulations (Strahs and Schlick, 2000). At higher added salt concentrations, on the order of ~ 200 mM added Na^+/Cl^- , significantly more ion occupancy in A-tract regions is observed ($\sim 25\%$ or more) (Cheatham and Young, 2001). Based on our observations, and under a low salt limit, we can nevertheless conclude that the structural and elastic properties of the N-tracts, as revealed by the present simulations, are not affected by direct cation-base contacts in the minor groove. Despite less direct association in the minor groove than expected, it should be nevertheless noted that the lack of Na^+ association with the bases in the minor groove is not in disagreement with NMR studies that show a considerable presence of NH_4^+ there. As noted by Denisov and Halle (2000), NH_4^+ appears to bind to the A-tracts by an order of magnitude more strongly than Na^+ . Note also that the current experimental evidence concerning direct association with nucleobases in the minor groove of B-DNA is somewhat conflicting, and it may well be that the degree of cation association is highly dependent on the type of cation, concentration, temperature, and other factors (Denisov and Halle, 2000; McConnell and Beveridge, 2000; Stellwagen et al., 2001).

Elastic properties of the tracts

There are important biological processes in which not only the structure, but also the mechanical deformability of DNA

plays a crucial role. These include specific DNA-binding to proteins or processes related to long-range interactions and higher-order organization of DNA in the nucleus. Many such biological functions occur within the linear elasticity regime. In that case the elastic energy can be expressed as a quadratic function of appropriate deformation variables.

Various methods have been devised to theoretically calculate the elastic constants. One of the most promising is the “method of fluctuations”: if an ensemble of deformation variables with Gaussian distribution is available, then the correlations of these variables are directly related to the harmonic force constants (Landau and Lifshitz, 1980). Olson et al. (1998) used this approach to calculate empirical deformabilities of dinucleotide steps based on an ensemble of x-ray protein-DNA complexes. They obtained force constants within an unknown multiplicative factor, the effective temperature of the ensemble.

It is, however, also desirable to know elastic properties of much larger portions of DNA than basepair steps. In a recent attempt, Bruant et al. (1999) used ensembles of snapshots from 1-ns molecular dynamics trajectories to calculate the isotropic bending, twisting, and stretching rigidities of two 15-bp oligomers. While their bending and twisting constants were in reasonable agreement with experiment, the stretch modulus was much higher than expected from the macroscopic experiments.

In our previous work (Lankaš et al., 2000) we performed 5-ns molecular dynamics runs on four 17-bp DNA oligomers with different sequences and calculated the elastic constants including anisotropic bending and all the coupling terms. In particular, the value of twist-stretch coupling was established, and the twist-bend coupling constant, whose existence was theoretically predicted some time ago (Marko and Siggia, 1994; O’Hern et al., 1998), was calculated for the first time. Although the values of the force constants were in very good overall agreement with experimental values for random sequences, a pronounced sequence-dependence of elastic properties was observed, especially regarding the stretch modulus.

Large-scale molecular dynamics simulations thus represent a unique source of data to study DNA elasticity, including its sequence dependence. Here, we use the method from our previous work to compare the deformabilities of the five tracts under study.

The elastic constants are shown in Table 4. The values of the stretch modulus, Y_f , strongly confirm the similarity of A-, I-, and AI-tracts versus G- and D-tracts. The former are much stiffer than the latter, with AI and D as extreme examples. Just the opposite, yet not so markedly pronounced, holds for the twist persistence length C : the former group is in general softer with respect to twisting than the latter one. No clear tendency has been observed in the case of the isotropic bending persistence length A_{iso} ; here the AI-tract alone stands out as the most flexible. It thus seems that the bending stiffness is not a simple consequence of the

TABLE 4 Elastic constants of the tracts

	Y_f (pN)	C (nm)	A_{iso} (nm)	D (nm)	A_{lt}^* (nm)
A-Tract	1564	60	92	15	52
I-Tract	1362	97	91	10	30
AI-Tract	2094	107	72	19	39
G-Tract	784	126	92	18	45
D-Tract	730	160	87	10	52

Y_f , stretch modulus; C, twist persistence length; A_{iso} , isotropic bending persistence length; D, twist-stretch coupling; A_{lt} , coupling between twist and bending into the grooves of the central basepair. The elastic constants were evaluated from the last 4 ns of the trajectories. See Lankas et al. (2000) for methodological details.

*Value for the 6-bp fragment. See text.

presence or absence of the C2 amino group. The same applies to the twist-stretch coupling D .

The coupling between twist and bending into the major groove of the tract's central basepair, A_{lt} , is the most important coupling term for fragments shorter than one helical turn. On the dinucleotide level it corresponds to the anticorrelation between twist and roll, a phenomenon observed in the crystallographic data (Gorin et al., 1995; Olson et al., 1998). This coupling diminishes as the fragment length approaches one helical turn; therefore we report in the table the values for 6-bp fragments (averaged over all possible 6-mers of the tract). All the other coupling terms are small and are not reported in the Table.

As far as the anisotropic bending is concerned, it is markedly pronounced for small numbers of basepairs in a fragment; our calculations show that about two times less energy is needed to bend a trimer into the grooves of its central basepair than in the perpendicular direction. This is in line with the view that changes in roll should be much easier to perform than changes in tilt (Olson et al., 1998; Zhurkin et al., 1979). This anisotropy vanishes at the half of the helical turn, and is quite small at the full length of our tracts, in accordance with our previous results (Lankas et al., 2000).

Besides the global elastic properties mentioned above, one might be interested in the DNA deformability on the basepair or basepair step level. The corresponding force constants may be inferred from the molecular dynamics trajectories as well (Lankas et al., manuscript in preparation).

Structure of $d(A)_{10}$: Comparison of four force fields

The lower than experimentally observed propeller twist for the A-tract sequence studied (see above) led us to extend our study and to investigate the force field dependence and possible bias in the observed values. To characterize this, simulations of a 10-mer polyA-polyT were performed with four of the most reliable and currently available empirical, all-atom nucleic acid force fields. These include the Cornell

et al. (1995), Cheatham et al. (parm98/99) (Cheatham et al., 1999; Wang et al., 2000), the MacKerell all_27 nucleic acid force field (Foloppe and MacKerell, Jr., 2000), and Langley's BMS force field (Langley, 1998). All simulations were carried out for 7 ns. Furthermore, to verify the effect of the starting structure, the simulations were carried out taking both A- and B-DNA structures as starting points. Thus, we have carried out altogether ~60 ns of simulations of the 10-mer polyA-polyT. With each of the force fields, a conversion from A-DNA to B-DNA is seen (data not shown) with convergence to a common set of structures for each applied force field. Thus, in the following, we present only data obtained for the structures initiated in the B-DNA form geometry, and these are summarized in Table 5. Comparing the B-DNA simulation results across the force fields suggests that all of these force fields tend to underestimate the average propeller twist compared to crystallographic data. The best behavior, arguably, comes from the BMS force field, which shows propeller twisting even slightly larger than suggested recently by NMR spectroscopy (MacDonald et al., 2001). It should be noted that in the simulated 10-mer polyA-polyT structures a certain variation is seen in the values at each basepair, suggesting that there is still somehow insufficient sampling in the current set of 7-ns simulations. As has been noted by us and others (Feig and Pettitt, 1998; Spackova et al., 2001; Sponer et al., 2000b), nanosecond-length simulations may not be sufficient to fully converge the structures. There are relatively large variations (standard deviations, see Table 5) observed during the simulation for the propeller twist, compared to helical twist, suggesting a wide range of motion of the basepairs. Note, however, that the structures of the N-tracts in our 17-bp duplex simulations studied herein and discussed previously are considerably more uniform. This may be related to the presence of terminal GCG segments shielding the core N-tract from end-effects. Nevertheless, convergence issues aside, all the force fields show similar trends.

It is difficult to rationalize at this point whether the lower than expected average propeller twist in the simulated A-tracts is due to lack of convergence or insufficient sampling, a potential bias in the experimental structures, or a systematic bias of the force field. The latter is distinctly possible due to the neglect of hydrogen bond directionality, polarization effects, or other force field approximations such as neglect of anisotropy of the short-range repulsion (Spuner et al., 1996b). However, despite the observation of lower propeller twist values, it is reassuring that all three different force field philosophies tend to give comparable results and show similar trends. A comparison of the Cornell et al. results to the other force fields suggests that the root of the "lower than expected" propeller twist in the Cornell et al. force field is not simply due to the known deficiencies of this force field. Although the Cornell et al. force underestimates the average sugar pucker phases and helical repeat, the MacKerell all_27 force field, which does a much better

TABLE 5 Selected structural parameters of $d(A)_{10}$ as obtained with four major force fields

	Propeller Twist (°)	Twist (°)	Roll (°)	Minor Groove Width (Å)
Cornell et al.	-14.1 -13.6 to -14.8 (13.0)	31.8 31.4 to 32.5 (5.06)	0.39 0.3 to 0.6 (8.06)	6.29 (1.07)
MacKerell	-13.0 -12.1 to -13.7 (12.0)	34.7 33.4 to 35.2 (6.98)	-0.76 -2.76 to -0.09 (7.58)	7.69 (0.93)
Langley BMS	-18.3 -16.6 to -19.8 (14.2)	34.9 34.6 to 35.4 (5.96)	1.2 0.4 to 1.25 (9.84)	7.69 (1.36)
Parm99	-14.4 -13.6 to -15.4 (12.7)	32.8 32.6 to 33.1 (5.33)	0.53 0.41 to 0.62 (8.03)	6.90 (1.32)

The propeller twist (°), overall helical twist (°), and average roll (°) for all the force fields were calculated separately from 2-ns straight coordinate averaged structures (at 1–3-, 2–4-, 3–5-, 4–6-, and 5–7-ns intervals where the snapshots were recorded at 0.5-ps intervals). Shown are the mean values over all basepairs in all structures, range of values across the five separate structures and below this, the standard deviations (in parentheses). The standard deviations (calculated for each basepair step from the trajectory data taken at 1-ps intervals over the entire 1–7-ns time range and averaged over the basepairs) do not give an estimate of the relative error, but give an idea of the dynamic fluctuations of the values. The values of the propeller twist and helical twist were calculated with Dials and Windows and are comparable to the values calculated by the SCHNAAP procedure. Except where noted, the minor groove widths (and fluctuations) were calculated over the interval of 1–7 ns by averaging the P_5-P_{19} , P_6-P_{18} , P_7-P_{17} , P_8-P_{16} , and P_9-P_{15} distances.

job in this regard, shows larger than expected minor groove widths and lower than expected propeller twist. As the parm98/99 results suggest a larger magnitude for the propeller twist, clearly the sugar pucker phases and helical repeat (which are improved in this force field) are involved. The increase in propeller twisting compared to the Cornell et al. force field, however, is not substantial. Despite the deficiencies and absence of specific hydrogen bond directionality and explicit polarization, it is reassuring that the BMS force field, which uses the Cornell et al. charge model and all_22 internal (bond/angle) parameters and an iteratively improved dihedral potential, does better. This suggests that with further effort to balance and tweak these simple empirical potentials it will be possible to generate an even more reliable (and general) nucleic acid force field.

Nevertheless, the most important conclusion regarding the purpose of the present study is that all the applied force fields show qualitatively similar results. Thus, the main results with the Cornell et al. force field described herein, regarding the intrinsic flexibility and variations in helicoidal parameters and minor groove widths across the polypurine sequences studied, seem to be general.

DISCUSSION AND CONCLUSIONS

We have performed 5–20-ns unrestrained simulations of $d[GCG(N)_{11}GCG]_2$ duplex structures. The simulations have been carried out with explicit inclusion of solvent and counterions and with the particle mesh Ewald method for an accurate treatment of long-range electrostatic forces. This technique is currently the most robust computational tool available for studies of structure and dynamics of oligonucleotides. We have used the Cornell et al. force field, which

is known to provide a realistic description of base stacking interactions (Hobza et al., 1997).

The simulation reproduces many essential features of the A-tracts. Mainly, it shows that the A-tract structure is substantially propeller-twisted, has a narrow minor groove, and the width of the minor groove systematically decreases along the tract in the 3' to 5' direction. The narrowing of the groove appears to be saturated after approximately six consecutive adenines. The A-tract is found to be rigid with respect to stretching and relatively flexible with respect to twisting, while its bending rigidity does not seem to differ much from that of the other tracts.

However, the simulation also shows rather substantial differences compared to the experimental findings. The propeller twist is (in absolute value) underestimated, being -12° in the simulations, while oligonucleotide crystals and fibers show a value around -20° and -26° , respectively. The calculated helical twist of the A-tract is underestimated by $\sim 3-4^\circ$. Also, the narrowing of the minor groove in the simulated structure is not as significant as in the experiments. These observations are in agreement with literature data (Cheatham and Kollman, 2000). These data indicate that while qualitatively correct, the force field is not capable of quantitatively reproducing some fine structural aspects of the A-tract. The most likely reasons might be the imbalance of the parameters for the sugar-phosphate backbone, imbalance of solute-solvent interactions, possible nonplanarity of the N2 amino group, or the absence of polarizabilities in the force field. Nevertheless, the simulation shows key features of the A-tracts and, therefore, it is justified to carry out simulations with different sequences. The relative differences among various sequences should be affected by the force field inaccuracies to a substantially smaller degree, as

all structures should be biased in the same direction, leaving sequence differences fully expressed. Furthermore, it has been explicitly demonstrated that the Cornell et al. force field provides a very good and well-balanced description of the base-base interactions (Hobza et al., 1997).

It should be noted that our simulations have been carried out with the original version of the Cornell et al. force field (Cornell et al., 1995). Recently, new variants of the force field, parm98 (Cheatham et al., 1999) and parm99 (Wang et al., 2000) have been released, with modified sugar pucker parameters. This influences the B-A equilibria provided by the force fields, helical twist parameters, and some other features of the simulated structures. Furthermore, there exist two other force fields designed for DNA simulations, BMS (Langley, 1998) and CHARMM (Foloppe and MacKerell, Jr., 2000). This has inspired us to carry out simulations of a long adenine d(A)₁₀ tract using all four parametrizations. These additional simulations show that, at least considering the presently available generation of force fields, our results obtained with the original Cornell et al. force field are general. These results are also in line with our preceding paper (Spackova et al., 2000), where we have found almost no difference in simulations of an adenine-zipper duplex with parm94 and parm98. This view is also in agreement with results obtained by others (Cheatham and Kollman, 2000). It is to be noted that all three AMBER force field variants have identical parametrization of DNA bases, which is the most important aspect of the force field for the purpose of the present study. Thus the choice of the force field (considering presently available parametrizations) should have no substantial influence on the outcome and conclusions of the present paper.

Simulations of d[GCG(I)₁₁GCG]₂ and d[GCG(AI)₅AGCG]₂ duplexes show unambiguously that both I- and AI-tracts adopt a high-propeller structure with a narrow minor groove width very similar to that of A-tract. This finding is in full agreement with all available experimental data (Bailly et al., 1995; Diekmann et al., 1992; Leslie et al., 1980). The only experimental result that might be in disagreement with our data is a crystal structure of a short medium-resolution I₃mC₃ tract (Shatzky-Schwartz et al., 1997). An analogous crystal structure with AIATmCT sequence shows a high propeller twist (Shatzky-Schwartz et al., 1997), entirely in agreement with our present data. The mechanical deformability of the I- and AI-tract is characterized by a high stretch modulus and rather low twist persistence length, again entirely in accordance with the properties of the A-tract.

The A-T and I-mC basepairs have the same van der Waals contour; however, they differ substantially in their molecular electrostatic potentials. As a result, AA and II steps differ considerably in the balance of intrastrand and interstrand contributions to stacking and in the interplay between van der Waals and electrostatic contributions (Sponer et al., 1997, 2000a). The base stacking in the II step

is weak and is characterized by an intrastrand electrostatic repulsion. The electrostatic component of stacking in an I-tract is considerably closer to stacking in a G-tract than to stacking in an A-tract. The alternating AI structure, where rather nonpolar AT basepairs are interdispersed between very polar IC basepairs, has electrostatic interactions obviously closer to the A-tract sequence. In our simulations, the structure of the I-tract remains similar to the A-tract, and the same is true for the AI-tract. Thus, the simulations clearly show that the steric effects associated with the presence or absence of the minor-groove amino group within the tract is the critical factor governing the formation of A-tract-like structure, while the electrostatic differences of the stacking remain unexpressed. It does not rule out, however, that the substantial difference in molecular electrostatic potentials of A-T and I-mC basepairs is not important at the tract-nontract junction, especially if the tract region is enveloped by polar G-C basepairs. The similarity of A- and AI-tracts has been anticipated because the strong polarity of the I-mC basepair is not expressed through stacking terms when these pairs are interdispersed between rather nonpolar A-T basepairs.

The conclusion concerning the critical role of the minor-groove amino group is fully supported by the simulations of G- and D-tracts (note that the electrostatic component of stacking in the D-tract is weak, resembling the A-tract and very distinct from the I- and G-tracts). This is in full agreement with experimental data, including the higher propensity of the G-tract to undergo a conformational shift toward A-form like structure. The elastic properties also sharply contrast those of the A-, I-, or AI-tract: both the simulated G- and D-tract have low stretch modulus and high twist rigidity. As far as the bending rigidity and twist-stretch coupling is concerned, no clear tendency has been observed and, perhaps more surprisingly, the values of the bending rigidity do not differ much among the tracts. This, of course, does not rule out that other sequences may exist with unusually low or high values of the bending rigidity.

To further investigate a possible dependence of our results on the time scale of the simulations, we prolonged three of the trajectories (for A-, G-, and I-tract) up to 20 ns each. The analysis of these extended simulations fully confirms our findings. In fact, the average conformational parameters differ in most cases by not more than several percent from those reported in Table 2. The elastic properties also keep pointing out the difference between the A- and I-tract, which are hard to stretch but relatively easy to twist, and the G-tract, where the opposite is true, with the bending rigidity similar for all the tracts. However, it has to be pointed out that while the general tendencies are reproduced, the values of the elastic constants somehow differ compared with the shorter simulations. The stretch modulus is 1793, 1400, and 657 pN, the twist persistence length 87, 72, and 112 nm, and the isotropic bending persistence length 88, 81, and 82 nm (for the A-, I-, and G-tract,

respectively). The biggest difference regards the twisting persistence length of the A-tract, where the value of 87 nm rather than 60 nm came out from the long simulation (see Table 4). This suggests that the anomalously low A-tract twisting persistence length of 60 nm was an artifact arising from insufficient sampling in the 5-ns simulation, and indicates that substantially longer trajectories than 5 ns may be necessary to capture the fluctuation-dependent properties of DNA oligomers, even in the case of homopolymeric sequences.

Being aware of all the limitations imposed by the force field approximations and time scale of our simulations, we intentionally restrict ourselves in further analyzing tiny details of the simulated structures, as we do not want to push the method beyond the limits of its applicability.

Finally, let us compare our results with related preceding computational studies. A particularly interesting study has been reported by Sherer and co-workers (1999). They have carried out 1-ns simulations of three self-complementary duplexes with short (three-base) tracts interdispersed by a purine-pyrimidine step: $d(\text{CGCAAATTTGCG})_2$, $d(\text{CGCAIATmCTGCG})_2$, and $d(\text{CGCIImCmCmCGCG})_2$. They have found large propeller twist in their A- and AI-tracts, slightly larger than in our simulations. However, they reported a moderate reduction of the propeller twist in their I-tract. This is in agreement with the crystal structure of Shatzky-Schwartz of the same sequence (Shatzky-Schwartz et al., 1997) but, at first sight, not in agreement with our simulations and majority of other experimental data. As they have used the same force field, the difference between the two studies may be due to the size of the tract and the junctions in the dodecamers. Furthermore, 1-ns simulations are not sufficiently converged, especially when simulating nonhomogeneous sequences (Feig and Pettitt, 1998; Spomer et al., 2000b). This can be illustrated by substantial differences (asymmetry) between the two halves of the self-complementary structures reported by Sherer et al. (1999; cf. upper and lower part of Table 1 in Sherer et al.). Considering all these aspects, the modest reduction of propeller reported by Sherer et al. for III₃ is not so dramatic and we believe they are not in disagreement with our data. This is, however, not the case of the 2.5-ns simulation of $d(\text{C-TITIIIGGGC})_2$ duplex studied by Pastor et al. (1999), showing virtually no propeller twist within the I-tract. This could be caused by the different sequence, mainly a G-tract mounted on the I-tract with the same polarity. This leads to an arrangement of six consecutive basepair steps with substantially unfavorable electrostatic stacking (note again the striking similarity of electronic structures of G and I). However, another factor could be the force field. Pastor et al. utilized a 1995 version of the CHARMM force field (MacKerell et al., 1995) which is known to be a little biased toward the A-form helices (Feig and Pettitt, 1998). Such a shift would abolish the A-tract-like structure and could be especially favored by the G-tract part of the duplex simu-

lated by Pastor et al. Furthermore, the CHARMM force field differs from the Cornell et al. force field in the description of base stacking energetics; the latter force field is substantially closer to our reference ab initio calculations (Hobza et al., 1997).

Several groups reported simulations of long (Cheatham et al., 1998) and short (Feig and Pettitt, 1998; Trantirek et al., 2000; Spomer et al., 2000b) G-tracts. The studies in general suggest that this sequence has a certain propensity to be partly shifted toward the A-form helix, which is also supported by the x-ray study (Ng et al., 2000). This is in line with our simulations.

Extensive molecular dynamics studies of phased four- and five-base A-tracts have been reported, with a similar simulation length as utilized in our study, and with the same force field (Sproun et al., 1999; Young and Beveridge, 1998). Although the authors do not provide details about the propeller twist distribution in the A_n segments, their overall description is consonant with our simulations. As discussed above in detail, our data for the A-tract appear to be consistent with simulations by Strahs and Schlick (2000).

We can conclude that our series of 5-ns/20-ns simulations of DNA duplexes containing 11-bp homopurine stretches, together with a careful evaluation of experimental and preceding simulation data, clearly show that the I- and AI-tracts intrinsically prefer a high-propeller conformational substate very similar to the native A-tract. However, G-tract and D-tract adopt a low propeller architecture with certain A-DNA features, considerably more expressed for the G-tract. The elastic properties also differ between the two types of tracts. The absence of the minor-groove amino group in purine position 2 (i.e., the van der Waals contour of the pair) is the major factor allowing a formation of the A-tract-like structure, due to steric and dispersion effects, including changes of hydration caused by the protruding amino group (Lan and McLaughlin, 2001). The electrostatic component of stacking seems to be unimportant in that respect. We do not suggest, however, that different stacking electrostatic properties of various basepairs are not critically important in some other aspect DNA structure and dynamics (Bailly et al., 1999; Gallego et al., 1994). We believe, despite some limitations discussed above, that our results confirm that modern large-scale simulations of nucleic acids represent a powerful tool to study qualitative and sometimes quantitative aspects of the effect of base composition on DNA structure and dynamics, and represent a useful complement to available experimental techniques.

The simulations of the five N-tracts were carried out using the high-performance computers at the German Cancer Research Center (DKFZ) in Heidelberg. The extended A-tract, I-tract, and G-tract simulations, along with the polyA-polyT 10-mer simulations with different force fields, were performed at the National Institutes of Health using the LoBoS I and II cluster of Pentiums in Bernie Brooks's laboratory at the Center for High Performance Computing at the University of Utah, and continued during

the acceptance phase of the Terrascale System at the Pittsburgh Supercomputing Center, for which we gratefully acknowledge access to resources. We also thank Alex MacKerell and David Langley for early access to their force fields.

This project (LN00A032, Center for Complex Molecular Systems and Biomolecules) was supported by the Ministry of Education of the Czech Republic.

REFERENCES

- Alexeev, D. G., A. A. Lipanov, and I. Y. Skuratovskii. 1987. Poly (dA)₅poly(dT) is a B-type double helix with a distinctively narrow minor groove. *Nature*. 325:821–823.
- Arnott, S., R. Chandrasekaran, I. H. Hall, and L. C. Puigjaner. 1983. Heteronomous DNA. *Nucleic Acids Res.* 11:4141–4155.
- Bailly, C., S. Echeperre, F. Gago, and M. J. Waring. 1999. Recognition elements that determine affinity and sequence-specific binding to DNA of 2QN, a biosynthetic bis-quinoline analogue of echinomycin. *Anti-Cancer Drug Design*. 14:291–303.
- Bailly, C., N. E. Mollegaard, P. E. Nielsen, and M. J. Waring. 1995. The influence of the 2-amino group of guanine on DNA conformation-uranyl and DNase-I probing of inosine diaminopurine substituted DNA. *EMBO J.* 14:2121–2131.
- Bailly, C., and M. J. Waring. 1998. The use of diaminopurine to investigate structural properties of nucleic acids and molecular recognition between ligands and DNA. *Nucleic Acids Res.* 26:4309–4314.
- Bayly, C. I., P. Cieplak, W. D. Cornell, and P. A. Kollman. 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* 97:10269–10280.
- Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.
- Beveridge, D. L., and K. J. McConnell. 2000. Nucleic acids: theory and computer simulation, Y2K. *Curr. Opin. Struct. Biol.* 10:182–196.
- Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* 4:187–217.
- Bruant, N., D. Flatters, R. Lavery, and D. Genest. 1999. From atomic to mesoscopic descriptions of the internal dynamics of DNA. *Biophys. J.* 77:2366–2376.
- Burkhoff, A. M., and T. D. Tullius. 1987. The unusual conformation adopted by the adenine tracts in kinetoplast DNA. *Cell*. 48:935–943.
- Chan, S. S., K. J. Breslauer, R. H. Austin, and M. E. Hogan. 1993. Thermodynamics and premelting conformational changes of phased d(A)₅ tracts. *Biochemistry*. 32:11776–11784.
- Cheatham, T. E. III, and B. R. Brooks. 1998. Recent advances in molecular dynamics simulation towards the realistic representation of biomolecules in solution. *Theor. Chem. Acc.* 99:279–288.
- Cheatham, T. E. III, P. Cieplak, and P. A. Kollman. 1999. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.* 16:845–862.
- Cheatham, T. E. III, and P. A. Kollman. 2000. Molecular dynamics simulation of nucleic acids. *Annu. Rev. Phys. Chem.* 51:435–471.
- Cheatham, T. E. III, J. Srinivasan, D. A. Case, and P. A. Kollman. 1998. Molecular dynamics and continuum solvent studies of the stability of polyG-polyC and polyA-polyT DNA duplexes in solution. *J. Biomol. Struct. Dyn.* 16:265–280.
- Cheatham, T. E. III, and M. A. Young. 2000. Molecular dynamics simulations of nucleic acids: successes, limitations, and promise. *Biopolymers* 56:232–256.
- Chiu, S.-W., M. Clark, S. Subramaniam, and E. Jakobsson. 2000. Collective motion artifacts arising in long-duration molecular dynamics simulations. *J. Comp. Chem.* 21:121–131.
- Chuprina, V. P., A. A. Lipanov, O. Y. Fedoroff, S.-G. Kim, A. Kintanar, and B. R. Reid. 1991. Sequence effects on local DNA topology. *Proc. Natl. Acad. Sci. U.S.A.* 88:9087–9091.
- Coll, M., C. A. Frederick, A. H.-J. Wang, and A. Rich. 1987. A bifurcated hydrogen-bonded conformation in the d(A.T) base-pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc. Natl. Acad. Sci. U.S.A.* 84:8385–8389.
- Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117:5179–5197.
- Cornell, W. D., P. Cieplak, C. I. Bayly, and P. A. Kollman. 1993. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J. Am. Chem. Soc.* 115:9620–9631.
- Crothers, D. M., T. E. Haran, and J. G. Nadeau. 1990. Intrinsically bent DNA. *J. Biol. Chem.* 265:7093–7096.
- Crothers, D. M., and Z. Shakked. 1999. DNA bending by adenine-thymine tracts. In *Oxford Handbook of Nucleic Acid Structure*. S. Neidle, editor. Oxford University Press.
- Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an N²log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.
- Denisov, V. P., and B. Halle. 2000. Sequence-specific binding of counterions to B-DNA. *Proc. Natl. Acad. Sci. U.S.A.* 97:629–633.
- De Santis, P., A. Palleschi, M. Savino, and A. Scipioni. 1990. Validity of the nearest-neighbor approximation in the evaluation of the electrophoretic manifestations of DNA curvature. *Biochemistry*. 29:9269–9273.
- Dickerson, R. E. 1999. Helix structure and molecular recognition by B-DNA. In *Oxford Handbook of Nucleic Acid Structure*. S. Neidle, editor. Oxford University Press.
- Diekmann, S. 1986. Sequence specificity of cured DNA. *FEBS Lett.* 195:53–56.
- Diekmann, S., J. M. Mazzealli, L. W. McLaughlin, E. von Kitzing, and A. A. Travers. 1992. DNA curvature does not require bifurcated hydrogen bonds or pyrimidine methyl groups. *J. Mol. Biol.* 225:729–738.
- Diekmann, S., E. von Kitzing, L. McLaughlin, J. Ott, and F. Eckstein. 1987. The influence of exocyclic substituents of purine bases on DNA curvature. *Proc. Natl. Acad. Sci. U.S.A.* 84:8257–8261.
- DiGabriele, A. D., M. R. Sanderson, and T. A. Steitz. 1989. Crystal-lattice packing is important in determining the bend of a DNA dodecamer containing an adenine tract. *Proc. Natl. Acad. Sci. U.S.A.* 86:1816–1820.
- DiGabriele, A. D., and T. A. Steitz. 1993. A DNA dodecamer containing an adenine tract crystallizes in a unique lattice and exhibits a new bend. *J. Mol. Biol.* 231:1024–1039.
- Dornberger, U., M. Leijon, and H. Fritzsche. 1999. High basepair opening rates in tracts of GC basepairs. *J. Biol. Chem.* 274:6957–6962.
- El Hassan, M. A., and C. R. Calladine. 1995. The assessment of the geometry of dinucleotide steps in double-helical DNA: a new local calculation scheme. *J. Mol. Biol.* 251:648–664.
- Feig, M., and B. M. Pettitt. 1998. Structural equilibrium of DNA represented with different force fields. *Biophys. J.* 75:134–149.
- Feller, S. E., Y. H. Zhang, R. W. Pastor, and B. R. Brooks. 1995. Constant-pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* 103:4613–4621.
- Florian, J., J. Spomer, and A. Warshel. 1999. Thermodynamic parameters for stacking and hydrogen bonding of nucleic acid bases in aqueous solution: ab initio/Langevin dipoles study. *J. Phys. Chem. B.* 103:884–892.
- Foloppe, N., and A. D. MacKerell, Jr. 2000. All-atom empirical force field for nucleic acids. I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comp. Chem.* 21:86–104.
- Friedman, R. A., and B. Honig. 1992. The electrostatic contribution to DNA base-stacking interactions. *Biopolymers*. 32:145–159.

- Gallego, J., E. B. Golden, D. E. Stanley, and B. R. Reid. 1999. The folding of centromeric DNA strands into intercalated structures: a physicochemical and computational study. *J. Mol. Biol.* 285:1039–1052.
- Gallego, J., F. J. Luque, M. Orozco, C. Burgos, J. Alvarez-Builla, M. M. Rodrigo, and F. Gago. 1994. DNA sequence-specific reading by ecinomycin: role of hydrogen-bonding and stacking interactions. *J. Med. Chem.* 37:1602–1609.
- Goodsell, D. S., and R. E. Dickerson. 1994. Bending and curvature calculations in B-DNA. *Nucleic Acids Res.* 22:5497–5503.
- Gorin, A. A., V. B. Zhurkin, and W. K. Olson. 1995. B-DNA twisting correlates with base-pair morphology. *J. Mol. Biol.* 247:34–48.
- Hamelberg, D., L. McFail-Isom, L. D. Williams, and W. D. Wilson. 2000. Flexible structure of DNA: ion dependence of minor-groove structure and dynamics. *J. Am. Chem. Soc.* 122:10513–10520.
- Hamelberg, D., L. D. Williams, and W. D. Wilson. 2001. Influence of the dynamic positions of cations on the structure of the DNA minor groove: sequence-dependent effects. *J. Am. Chem. Soc.* 123:7745–7755.
- Haran, T. E., and D. M. Crothers. 1989. Cooperativity in A-tract structure and bending properties of composite T_nA_n blocks. *Biochemistry.* 28:2763–2767.
- Harvey, S. C., R.-K. Z. Tan, and T. E. Cheatham III. 1998. The flying ice cube: Velocity rescaling in molecular dynamics leads to violation of energy equipartition. *J. Comp. Chem.* 19:726–740.
- Herrera, J. E., and J. B. Chaires. 1989. A premelting conformational transition in poly(dA)-poly(dT) coupled to daunomycin binding. *Biochemistry.* 28:1993–2000.
- Hobza, P., M. Kabelac, J. Sponer, P. Mejzlik, and J. Vondrasek. 1997. Performance of empirical potentials (AMBER, CFF95, CFVV, CHARMM, OPLS, POLTEV), semiempirical quantum chemical methods (AM1, MNDO/M, PM3), and ab initio Hartree-Fock method for interaction of DNA bases: comparison with nonempirical beyond Hartree-Fock results. *J. Comput. Chem.* 18:1136–1150.
- Hobza, P., and J. Sponer. 1999. Structure, energetics, and dynamics of the nucleic acid basepairs: nonempirical ab initio calculations. *Chem. Rev.* 99:3247–3276.
- Hoover, W. G. 1995. Canonical dynamics: equilibrium phase distributions. *Phys. Rev. A.* 31:1695–1697.
- Hud, N. V., V. Sklenar, and J. Feigon. 1999. Localization of ammonium ions in the minor groove of DNA duplexes in solution and the origin of DNA A-tract bending. *J. Mol. Biol.* 286:651–660.
- Jerkovic, B., and P. H. Bolton. 2000. The curvature of dA tracts is temperature dependent. *Biochemistry.* 39:12121–12127.
- Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. M. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
- Katahira, M., H. Sugeta, and Y. Kyogoku. 1990. A new model for the bending of DNAs containing the oligo(dA) tracts based on NMR observations. *Nucleic Acids Res.* 18:613–618.
- Koo, H.-S., and D. M. Crothers. 1987. Chemical determinants of DNA bending at adenine-thymine tracts. *Biochemistry.* 26:3745–3748.
- Koo, H.-S., H.-M. Wu, and D. M. Crothers. 1986. DNA bending at adenine-thymine tracts. *Nature.* 320:501–506.
- Lan, T., and L. W. McLaughlin. 2001. Minor groove functional groups are critical for the B-form conformation of duplex DNA. *Biochemistry.* 40:968–976.
- Landau, L. D., and E. M. Lifshitz. 1980. *Statistical Physics, Part 1.* Butterworth-Heinemann, Oxford, U.K.
- Langley, D. R. 1998. Molecular dynamic simulations of environment and sequence-dependent DNA conformations: the development of the BMS nucleic acid force field and comparison with experimental results. *J. Biomol. Struct. Dyn.* 16:487–509.
- Lankaš, F., J. Sponer, P. Hobza, and J. Langowski. 2000. Sequence-dependent elastic properties of DNA. *J. Mol. Biol.* 299:695–709.
- Lavery, R., and B. Pullman. 1985. The dependence of the surface electrostatic potential of B-DNA on environmental factors. *J. Biomol. Struct. Dyn.* 2:1021–1032.
- Lavery, R., and H. Sklenar. 1988. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* 6:63–91.
- Lavery, R., and H. Sklenar. 1989. Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.* 6:655–667.
- Leslie, A. G. W., S. Arnott, R. Chandrasekaran, and R. L. Ratliff. 1980. Polymorphism of DNA double helices. *J. Mol. Biol.* 143:49–72.
- Lu, X.-J., M. A. El Hassan, and C. A. Hunter. 1997. structure and conformation of helical nucleic acids: analysis program (SCHNAaP). *J. Mol. Biol.* 273:668–680.
- Luisi, B., M. Orozco, J. Sponer, F. J. Luque, and Z. Shakked. 1998. On the potential role of the amino nitrogen atom as a hydrogen acceptor in macromolecules. *J. Mol. Biol.* 279:1123–1136.
- Luo, R., H. S. R. Gilson, M. J. Potter, and M. K. Gilson. 2001. The physical basis of nucleic acid base stacking. *Biophys. J.* 80:140–148.
- MacDonald, D., K. Herbert, X. Zhang, T. Polgruto, and P. Lu. 2001. Solution structure of an A-tract DNA bend. *J. Mol. Biol.* 306:1081–1098.
- MacKerell, A. D., Jr., J. Wiorkiewicz-Kuczera, and M. Karplus. 1995. An all-atom empirical energy function for the simulation of nucleic acids. *J. Am. Chem. Soc.* 117:11946–11975.
- Marini, J. C., P. N. Effron, T. C. Goodman, C. K. Singleton, R. D. Wells, R. M. Wartell, and P. T. Englund. 1984. Physical characterization of a kinetoplast fragment with unusual properties. *J. Biol. Chem.* 259:8974–8979.
- Marini, J. C., S. D. Levene, D. M. Crothers, and P. T. Englund. 1982. Bent helical structure in kinetoplast DNA. *Proc. Natl. Acad. Sci. U.S.A.* 79:7664–7668.
- Marko, J. F., and E. D. Siggia. 1994. Bending and twisting elasticity of DNA. *Macromolecules.* 27:981–998.
- Maroun, R. C., and W. K. Olson. 1988. Base sequence effects in double-helical DNA. 3. Average properties of curved DNA. *Biopolymers.* 27:585–603.
- McConnell, K. J., and D. L. Beveridge. 2000. DNA structure: what's in charge? *J. Mol. Biol.* 304:803–820.
- McConnell, K. J., and D. L. Beveridge. 2001. Molecular dynamics simulations of B'-DNA: sequence effects on A-tract-induced bending and flexibility. *J. Mol. Biol.* 314:23–40.
- Mollegaard, N. E., C. Bailly, M. J. Waring, and P. E. Nielsen. 1997. Effects of diaminopurine and inosine substitutions on A-tract induced DNA curvature. Importance of the 3'-A-tract junction. *Nucleic Acids Res.* 25:3497–3502.
- Nelson, H. C. M., J. T. Finch, B. F. Luisi, and A. Klug. 1987. The structure of an oligo(dA)-oligo(dT) tract and its biological implications. *Nature.* 330:221–226.
- Ng, H.-L., M. L. Kopka, and R. E. Dickerson. 1999. The deformability of G-tracts: the crystal structure of a complete turn of an A/B-DNA intermediate. Proceedings of the Tenth Conversation. The University, Albany, New York, June 15–19.
- Ng, H.-L., M. L. Kopka, and R. E. Dickerson. 2000. The structure of a stable intermediate in the A↔B DNA helix transition. *Proc. Natl. Acad. Sci. U.S.A.* 97:2035–2039.
- Nosé, S. 1984. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* 52:255–268.
- O'Hern, C. S., R. D. Kamien, T. C. Lubensky, and P. Nelson. 1998. Elasticity theory of a twisted stack of plates. *Eur. Phys. J. B.* 1:95–102.
- Olson, W. K., A. A. Gorin, X.-J. Lu, L. M. Hock, and V. B. Zhurkin. 1998. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. U.S.A.* 95:11163–11168.
- Olson, W. K., N. L. Marky, R. L. Jernigan, and V. B. Zhurkin. 1993. Influence of fluctuations on DNA curvature: a comparison of flexible and static wedge models of intrinsically bent DNA. *J. Mol. Biol.* 232:530–551.
- Olson, W. K., and V. B. Zhurkin. 1996. Twenty years of DNA bending. In *Biological Structure and Dynamics. Proceedings of the Ninth Conversation.* R. H. Sarma and M.H. Sarma, editors. Adenine Press, Albany, N.Y.

- Pastor, N., A. D. MacKerell, Jr., and H. Weinstein. 1999. TIT for TAT: the properties of inosine and adenosine in TATA box DNA. *J. Biomol. Struct. Dyn.* 16:787–810.
- Peck, L. J., and J.-C. Wang. 1981. Sequence dependence of the helical repeat of DNA in solution. *Nature.* 292:375–378.
- Rhodes, D., and A. Klug. 1981. Sequence-dependent helical periodicity of DNA. *Nature.* 292:378–380.
- Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Chem. Phys.* 23: 327–341.
- Shatzky-Schwartz, M., N. D. Arbuckle, M. Eisenstein, D. Rabinovich, A. Bareket-Samish, T. E. Haran, B. F. Luisi, and Z. Shakked. 1997. X-ray and solution studies of DNA oligomers and implications for the structural basis of A-tract-dependent curvature. *J. Mol. Biol.* 267:595–623.
- Sherer, E. C., S. A. Harris, R. Soliva, M. Orozco, and C. A. Laughton. 1999. Molecular dynamics studies of DNA A-tract structure and flexibility. *J. Am. Chem. Soc.* 121:5981–5991.
- Shum, B. W., and D. M. Crothers. 1983. High-resolution NMR studies of A-containing and G-containing oligonucleotides. *Biopolymers.* 22: 919–933.
- Spackova, N., I. Berger, M. Egli, and J. Sponer. 1998. Molecular dynamics of hemiprotonated intercalated four-stranded i-DNA: stable trajectories on a nanosecond scale. *J. Am. Chem. Soc.* 120:6147–6151.
- Spackova, N., I. Berger, and J. Sponer. 1999. Nanosecond molecular dynamics simulations of parallel and antiparallel guanine quadruplex DNA molecules. *J. Am. Chem. Soc.* 121:5519–5534.
- Spackova, N., I. Berger, and J. Sponer. 2000. Nanosecond molecular dynamics of zipper-like DNA duplex structures containing sheared G.A mismatch pairs. *J. Am. Chem. Soc.* 122:7564–7572.
- Spackova, N., I. Berger, and J. Sponer. 2001. Structural dynamics and cation interactions of DNA quadruplex molecules containing mixed guanine/cytosine quartets revealed by large-scale molecular dynamic simulations. *J. Am. Chem. Soc.* 123:3295–3307.
- Sponer, J., I. Berger, N. Spackova, J. Leszczynski, and P. Hobza. 2000a. Aromatic base stacking in DNA: from ab initio calculations to molecular dynamics simulations. *J. Biomol. Struct. Dyn. Conversation.* 11: 383–407.
- Sponer, J., J. Florian, H.-L. Ng, and N. Spackova. 2000b. Local conformational variations observed in B-DNA crystals do not improve base stacking. Computational analysis of base stacking in d(CATGGGC-CCATG)₂ B[\rightarrow]A intermediate crystal structure. *Nucleic Acids Res.* 28:4893–4902.
- Sponer, J., H. A. Gabb, J. Leszczynski, and P. Hobza. 1997. Base-base and deoxyribose-base stacking interactions in B-DNA and Z-DNA. A quantum-chemical study. *Biophys. J.* 73:76–87.
- Sponer, J., and P. Hobza. 1994. Bifurcated hydrogen bonds in DNA crystal structures. An ab initio quantum chemical study. *J. Am. Chem. Soc.* 116:709–714.
- Sponer, J., and J. Kypr. 1991. Different interstrand and intrastrand contributions to stacking account for roll variations at the alternating purine-pyrimidine sequences in A-DNA and A-RNA. *J. Mol. Biol.* 221: 761–764.
- Sponer, J., and J. Kypr. 1993. Theoretical analysis of the base stacking in DNA. Choice of the force field and comparison with oligonucleotide crystal structures. *J. Biomol. Struct. Dyn.* 11:277–292.
- Sponer, J., and J. Kypr. 1994. Close mutual contacts of the amino groups in DNA. *Int. J. Biol. Macromol.* 16:3–6.
- Sponer, J., J. Leszczynski, and P. Hobza. 1996a. Structures and energies of hydrogen-bonded DNA basepairs. a nonempirical study with inclusion of electron correlation. *J. Phys. Chem.* 100:1965–1974.
- Sponer, J., J. Leszczynski, and P. Hobza. 1996b. Base stacking in cytosine dimer. A comparison of correlated ab initio calculations with three empirical potential models and density functional theory calculations. *J. Comput. Chem.* 17:841–850.
- Sprou, D., M. A. Young, and D. L. Beveridge. 1999. Molecular dynamics studies of axis bending in d(G₅-(GA₄T₄C)₂-C₅) and d(G₅-(GT₄A₄C)₂-C₅): effects of sequence polarity on DNA curvature. *J. Mol. Biol.* 285:1623–1632.
- Stefl, R., and J. Koca. 2000. Unrestrained molecular dynamics simulations of [d(AT)(5)](2) duplex in aqueous solution: hydration and binding of sodium ions in the minor groove. *J. Am. Chem. Soc.* 122:5025–5033.
- Stefl, R., N. Spackova, I. Berger, J. Koca, and J. Sponer. 2001. Molecular dynamics of DNA quadruplex molecules containing inosine, 6-thioguanine and 6-thiopurine. *Biophys. J.* 80:455–468.
- Stellwagen, N. C., S. Magnusdottir, C. Gelfi, and P. G. Righetti. 2001. Preferential counterion binding to A-tract DNA oligomers. *J. Mol. Biol.* 305:1025–1033.
- Strahs, D., and T. Schlick. 2000. A-tract bending: insights into experimental structures by computational models. *J. Mol. Biol.* 301:643–663.
- Strauss, F., C. Gaillard, and A. Prunell. 1981. Helical periodicity of DNA, poly(dA).poly(dT) and poly(dA-dT).poly(dA-dT) in solution. *Eur. J. Biochem.* 118:215–222.
- Trantirek, L., R. Stefl, M. Vorlickova, J. Koca, V. Sklenar, and J. Kypr. 2000. An A-type double helix of DNA having B-type puckering of the deoxyribose rings. *J. Mol. Biol.* 297:907–922.
- Wahl, M. C., and M. Sundaralingam. 1997. Crystal structures of A-DNA duplexes. *Biopolymers.* 44:45–63.
- Wang, J., P. Cieplak, and P. A. Kollman. 2000. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comp. Chem.* 21:1049–1074.
- Wing, R. M., H. P. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R. E. Dickerson. 1980. Crystal-structure analysis of a complete turn of B-DNA. *Nature.* 287:755–758.
- Wu, H.-M., and D. M. Crothers. 1984. The locus of sequence-directed and protein-induced DNA bending. *Nature.* 308:509–513.
- Young, M. A., and D. L. Beveridge. 1998. Molecular dynamics simulations of an oligonucleotide duplex with adenine tracts phased by a full helix turn. *J. Mol. Biol.* 281:675–687.
- Young, M. A., B. Jayram, and D. L. Beveridge. 1997a. Intrusion of counterions into the spine of hydration in the minor groove of B-DNA. *J. Am. Chem. Soc.* 119:59–69.
- Young, M. A., G. Ravishanker, and D. L. Beveridge. 1997b. A 5-nanosecond molecular dynamics trajectory for B-DNA: analysis of structure, motions, and solvation. *Biophys. J.* 73:2313–2336.
- Zhurkin, V. B., Y. P. Lysov, and V. I. Ivanov. 1979. Anisotropic flexibility of DNA and the nucleosomal structure. *Nucleic Acids Res.* 27: 1081–1096.