doi:10.1006/jmbi.2001.4926 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2001) 314, 23–40

JMB

Molecular Dynamics Simulations of *B′*-DNA: Sequence Effects on A-tract-induced Bending and Flexibility

Kevin J. McConnell and D.L. Beveridge*

Chemistry Department and Molecular Biophysics Program Wesleyan University Middletown, CT 06459, USA

Molecular dynamics (MD) simulations including water and counterions are reported on five examples of A-tract DNA oligonucleotide dodecamer duplexes for which crystal structures are available, the homopolymeric duplex sequences poly(dA) and poly(dG), and two related sequences that serve as controls. MD was performed using the AMBER suite of programs for 3 ns on each sequence. These results, combined with previously reported MDs on 25-mer and 30-mer oligonucleotides on sequences with phased A-tracts carried out under a similar simulation protocol, are used to examine salient issues in the structural chemistry of ApA steps and A-tract induced axis bending. MD modeling sucessfully describes the distinctive *B′* structure of A-tracts in solution as essentially straight (wedge angles of <1°), more rigid than generic *B*-form DNA, with slight base-pair *inclination*, high *propeller* twist and a minor groove narrowing 5′ to 3′. The MD structures in solution agree closely with corresponding crystal structures, supporting the idea that crystal structures provide a good model for A-tract DNA structure in solution. From the collective MD results, bending and flexibility are calculated by step. Pyrimidine-purine steps are predicted to be most intrinsically bent and also most bendable, i.e. susceptible to bending. Pyrimidine-pyrimidine (~ purine-purine) and purine-pyrimidine steps show less intrinsic deformation and deformability. The MD calculated flexibility correlates well with the protein-induced bendability derived independently from the protein DNA crystal structures. The MD results indicate that bending and flexibility of base-pair steps in DNA are highly correlated, i.e. steps that exhibit the most intrinsic deformation from *B*-form DNA turn are also the most dynamically deformable. The MD description of A-tract-induced axis bending shows most consistency with the non A-tract, general-sequence model, in which the sequence curvature originates primarily in base-pair *roll* towards the major groove in non-A-tract regions of the sequence, particularly pyrimidine-purine steps. The direction of curvature is towards the minor groove viewed from opposite the A-tracts, but the A-tracts *per se* exhibit only minor deformation. The MD results are found to be consistent with the directionality of bending inferred for DNA sequences from gel retardation and cyclization experiments.

*Corresponding author

Keywords: DNA; molecular dynamics; *B′* form; A-tracts; sequence effects

Introduction

The structure of the DNA sequence poly(dA) as well as stretches of A:T base-pairs in a DNA sequence assumes a variant of the structure of *B*-DNA known as the *B′* form, which exhibits base-pair *inclination* (all designation of DNA structural parameters follow the Cambridge Convention[1] and are given in italics), larger *propeller*, and a narrower minor groove compared with *B*-form structures. However, fundamental aspects of *B′*-DNA in solution, including the structure of ApA base-pair steps and A-tracts in solution, sequence effects on bending and flexibility, the origins of A-tract stability, and mechanisms and models of A-tract-induced axis bending are not established

Abbreviations used: MD, molecular dynamics; PME, particle mesh Ewald; MPD, methyl-2,4-pentanediol; NOESY, nuclear Overhauser enhancement spectroscopy.

E-mail address of the corresponding author: dbeveridge@wesleyan.edu

0022-2836/01/010023–18 $35.00/0

unequivocally. Recently, with the development of second-generation nucleic acid force-fields[2] and improved treatment of long-range electrostatic interactions using the particle mesh Ewald (PME) method,[3] molecular dynamics (MD) simulations of DNA have shown markedly improved performance,[4] and provide an independent assessment of the structure of A-tracts and A-tract-induced axis bending. We report MD simulations in solution on five examples of A-tract-containing DNA oligonucleotides for which crystal structures are available, the homopolymeric sequences poly(dA) and poly(dG), and two related DNA sequences that serve as controls. All simulations include solvent water and mobile counterions explicitly. MD trajectories were performed using the AMBER suite of programs[5] for at least 3 ns each, 27 ns of MD *in toto*. The dynamical structures from MD modeling are compared with experimentally determined structures to establish further the intrinsic level of accuracy to be expected from MD modeling on DNA. The MD results on these nine sequences combined with four others reported under the same simulation protocol[6,7] provide a basis for further investigation of the structural chemistry of ApA steps and A-tracts in solution, the dynamical structure of the so-called *B′* form of DNA characteristic of A-tracts at ambient and lower temperatures, and the nature of A-tract-induced DNA bending and bendability in solution.

Background

Structural chemistry of *B′*-DNA

The unsettled state of the structure of A-tracts in solution arises, in part, from the difficulty of determining molecular structures in solution directly, and a concern that structures of DNA oligonucleotides determined from crystallography may not be fully indicative of solution structure, due to packing effects and/or possible artifacts introduced by co-crystallization agents such as methyl-2,4-pentanediol (MPD) in the crystal.[8,9] Obtaining accurate, high-resolution structures of DNA in solution by 2D nuclear Overhauser enhancement spectroscopy (NOESY) NMR is well known to be an underdetermined problem,[10] a subject that is addressed in concurrent papers[4] and in new improved structures obtained using dipolar couplings.[11,12] Other spectroscopic methods are of interest but typically address local rather than global aspects of DNA structure.[13,14]

DNA sequence effects can be discussed in terms of the structure of the ten unique dinucleotide steps. Using the conventional abbreviations of Y and R for pyrimidines and purines, respectively, dinucleotide steps may be grouped into the subclasses RpY, RpR (equivalent to YpY), and YpR. Of these steps, ApA (=TpT) of the RpR/YpY class has taken on a particular significance because of its prominent role in the structure of A-tracts in solution and in A-tract-induced axis bending. The A-tract sequence/structure motif is encountered in diverse phenomena of biological interest, including nucleosome structure, in which DNA curvature facilitates formation of the solenoidal DNA-histone complex,[15] as well as in structures of bent DNA complexed with regulatory proteins.[16] DNA sequences with tracts of four to eight A:T base-pairs (A-tracts) in phase with a full turn of *B*-form double helix[17] exhibit anomalously slow gel migration as a consequence of A-tract-induced structural changes, particularly axis curvature.

A more extensive background on the structural chemistry of ApA steps and experimental data on A-tract induced curvature in DNA oligonucleotides is provided in a recent reviews.[18–22] A steric clash argument for essentially straight, relatively rigid A-tracts emanates from the higher *propeller* twist of A:T base-pairs (''a stack of carpenter's sawhorses is more difficult to push over than a stack of flat wooden planks''[23]) and particularly favorable base-pair stacking at ApA steps.[23,24] ApA steps in crystals are, of course not perfectly straight, but show small, non-zero values of *roll* and *tilt*, −0.89 (±2.87) and −0.11 (±2.50), respectively, from current calculations for over 54 oligonucleotide crystal structures.[25] This results in a gentle writhe described by Dickerson and co-workers based on NEWHELIX as a property of all *B*-form DNA.[23] However, this is not necessarily to be taken as the origin of macroscopic curvature.

The possibility of a ''bifurcated'' hydrogen bond across the major groove between the A N6H atom on one strand of a DNA duplex and the T O4 on the other has been proposed on the basis of crystal structure evidence.[26,27] This interaction, if significant, would be facilitated by the relatively large ApA *propeller*, could augment Watson-Crick A:T base-pairing, and potentially rigidify the structure. However, studies of sequences replacing AAA with the inosine (I) mutant AIA, which has no opportunity to form the bifurcated H-bond, show little change in anomalous gel retardation, suggesting that the contribution of bifurcated H-bonds to conformational stability may be slight, if any. However, sequences with III replacing AIA show reduced (i.e. near normal) gel retardation,[28] as well as significantly lower *propeller* twist. The foreshortened A N6H to T O4 distance associated with the bifurcated H-bond across the major groove could arise as a *de facto* consequence of characteristic high *propeller* in A-tracts, rather than as the driving force for it; hence, one must be continually aware of a potential cause and effect problem. Examination of AAA steps in 17 different oligonucleotide crystal structures reveals that only 13% of the examples show structures with interatomic N6H to T O4 distances of 2.8 Å or less.[29] However, NMR data distinctly support an interaction at A N6,[12,30] and resonance Raman spectroscopy indicates an interaction at T O4.[14] While the net contribution from each instance would be expected to be small at best, the effect could, of course, be cumulative. Ghosh & Bansal,[29] based on

analysis of the crystallographic data base, have advanced a case for a minor groove cross-strand hydrogen bond, A C2 to T O2, but this involves a relatively non-polar C-H donor.

DNA structure is known be sensitive to solvent effects, and effects of hydration and counterion atmosphere may play a significant role in A-tract stability. The minor groove of *B*-form DNA features a characteristic ''spine of hydration'', which could be stabilizing the structure[31] (or *vice versa*). Recently, evidence from MD,[32] crystallography[33] and NMR spectroscopy [34,35] has been submitted in support of the idea of fractional occupation of counterions in A + T-rich sequences, and ion binding is thus another possible source of A-tract stability.[36] The case for this was advanced in a series of recent NMR studies by Hud & Feigon.[34,35] However, our most recent analyses of MD simulations indicate that fractional occupancy of Na counterions in the grooves of DNA is ca 10 % at best[37] (note our measure emphasizes the floor of the groove), and NMR indicates even less.[38] Recent crystal structures in the presence of heavier counterions show increased fractional occupancies,[39] as much as 50 % in the case of Tl^+.[40] There is leading evidence as well as a proposed model for ions in the grooves of DNA as bending loci.[37,39,41,42] Stellwagen *et al*.[43] have recently presented electrophoretic evidence for preferential counterion binding to A-tracts with implications for the phenomenon of A-tract-induced axis bending.

Theories of A-tract-induced axis bending of DNA can be categorized as wedge models,[44–46] junction models,[17] and non-A-tract (general sequence) models.[47–49] Non-zero values of *roll* and *tilt* give rise to the so-called ApA wedge angle,[44,46] the arithmetic average of *roll* and *tilt* angles. Positive *roll* at a step signals bending towards the major groove and negative *roll* results in bending towards the minor groove. The wedge model of DNA bending was postulated by Trifinov and co-workers.[46,50,51] In a wedge model for A-tract-induced curvature, the origin of curvature is primarily from bending within A-tracts. The gradual curvature in A-tracts noted above is not a wedge model *per se*, unless it is demonstrated to be the dominant feature in A-tract-induced axis bending.

Crothers and co-workers,[17,52,53] on the basis of gel retardation data, postulated a junction model, which features essentially straight A-tracts (values of *roll* and *tilt* ApA steps near zero), with axis bending arising from a change of base-pair *inclination* at junctions between A-tracts and flanking sequences of mixed composition and possibility of a *tilt* bend localized at the 5′ and 3′ junctions of A-tracts. The possibility of 3′ junction *roll* is said to be eliminated by the experimental data,[17] but a *roll* bend solely at the 5′ junction or distributed among non-A-tract base-pair steps is not precluded.[54]

The non-A-tract model[47–49] is characterized by essentially straight A-tracts, with sequence-dependent axis bending originating from positive *roll* in non-A-tract regions. An issue in dodecamer structures (see Strahs & Schlick;[55] MacDonald *et al*.[12]) with respect to inferences about A-tract-induced axis bending is that the non-A-tract region of a phased A-tract is not fully represented in a 12-mer model. To develop a general perspective on the structural implications of phased A-tracts, results on longer sequences will likely need to be considered.

Bolshoy *et al*.[44] have shown *via* parameter fitting that a dinucleotide model with an ApA wedge can account quantitatively for a large amount of gel retardation data. Goodsell & Dickerson[56] have analyzed various bending models and conclude that the non-A-tract model derived from nucleosome positioning data shows most consistency with the observed results. Liu *et al*.[25] revisited the dinucleotide model of DNA bending and demonstrated that the Bolshoy *et al*. results supporting the wedge model are not necessarily unique, and that a model with essentially straight ApA steps fits the data nearly as well and has improved predictive value.[57] The Liu *et al*. study advanced the case for non-A-tract bending as the primary origin of curvature in phased A-tract sequences, but indicated that a secondary effect may arise from uncompensated small bends at ApA steps.

A-tract-induced DNA curvature in each of the three basic theories of DNA bending, i.e. wedge, junction, and non-A-tract, arises from the idea that A-tracts have a unique structure related to, yet well differentiated from, normal *B*-form DNA, and generally referred to as *B*′-DNA. The effect of this (or any other) unique structure within a sequence is amplified when phased with respect to a full turn of helix, leading to concerted bending. Each of the three theories of DNA bending accounts for the direction of curvature deduced for phased A-tract DNA sequences from gel retardation experiments,[17,52,53] i.e. towards the minor groove with respect to A-tracts and toward the major groove with respect to non-A-tracts (see Figure 7 of Dickerson *et al*.[23]). However, none of the various proposed models has been proved or disproved unequivocally on the basis of experimental data to date.

The gel retardation for phased A-tract sequences is reduced with the addition of cosolvents such as ethanol or, in particular, MPD,[9] typically a component of the mother liquor from which single crystals of DNA oligonucleotides are grown. The effect of MPD has been shown to be localized to A-tracts.[9,58,59] Adding MPD appears to convert *B*′-form A-tracts into *B*-form (or something closer to *B*-form) structures, in which the unique features of A-tracts that are phased to produce concerted bending are diminished and the sequence thus runs more rapidly on the gel. From the point of view of a curved A-tract, wedge model of DNA bending,[46] MPD would ''straighten'' the wedge angle at ApA steps or the curvature of A-tracts in order to bring this about.[9,58] However, the six oligonucleotide crystal structures of DNA oligonu-

cleotides with A-tracts reported to date show no evidence of appreciable wedge-like deformations at ApA steps, and show ''essentially straight A-tracts'' with no inclination.[28,60,61] Olson *et al*.[62] proposed the ''flexible wedge'' model in which the deformation potential at ApA steps exhibited a minimum at a regular straight *B*-form structure but was anisotropic in shape with a preference toward deformation into the minor groove. This model would predict straight A-tracts at low temperature, accounting for the crystal structure results, whereas at high temperature, bent forms of ApA steps would be thermally populated and the dynamical (or ensemble average) average structure of A-tracts would show ApA wedges and curved A-tracts. Other proponents of the wedge model have argued that the crystal structure results are influenced by environmental and/or packing effects, and are not a good indication of the solution structure of DNA oligonucleotides. From the point of view of the non-A-tract bending model, the effect of MPD and similarly acting agents would be to increase the flexibility of essentially straight but more rigid *B*′-form A-tracts to something more *B*-like.[23] Thus, the MPD effect can be explained by either curved or straight A-tracts. However, explaining the MPD effect based on a curved A-tract model requires, in addition, that the crystal structures of A-tracts are atypical due to the effects of MPD, a speculation never confirmed directly. The non-A-tract bending model is consistent with crystallography results as they stand, and thus has at least the ''principle of parsimony'' in its favor.

Theoretical studies and simulations on DNA

Potential energy function calculations and energy minimization formed the basis for an influential theory of sequence effects on DNA structure with bending implications set forth in early work by Zhurkin *et al*.,[63,64] which distinguished the structural features of dinucleotide YR, RR (=YY) and RY steps, and has generally stood the test of time.[65] Sanghani *et al*.[66] described potential energy minimizations on the phased A-tract sequence poly(A_4T_4NN) on which regular axis curvature was imposed, and examined how the structures responded. The results produced features such as the 5′ to 3′ narrowing of the minor groove and implied an equilibrium geometry in reasonable accord with experiment, 26° of bending per A-tract compared with 11°-22° (often quoted as 18°) from cyclization data,[18] and clearly distinguished the helical phased motif of A_4T_4CG (experimentally curved) from T_4A_4NN (normal). Analysis of the results showed that A-tract-induced curvature in A_4T_4CG originates in a negative *roll* at the ApT step and with a positive *roll* at CpG. As noted, this does not coincide exactly with any of the models proposed above, but does follow Zhurkin's proposed pattern of sequence effects.

MD calculations based on all-atom potentials, while computationally intensive with solvent included explicitly, have the advantage that the origin of sequence curvature emerges as a result rather than as a postulate. MD on DNA has been characterized and validated for prototype cases in solution,[67–71] and in the crystalline state.[69] MD simulations have been used recently to investigate DNA sequence effects and bending,[1,2,55,72] and the relative stability of *A*-form and *B*-form structures in solutions of diverse composition.[73,74] The capabilities, limitations and applications to date of MD on nucleic acids have been reviewed recently.[4] MD simulations on DNA can now be performed including water and ionic strengths comparable to experiment, and can provide an independent perspective on the current controversy about the structure of A-tracts in solution and the nature of A-tract-induced axis bending. However, it must be carefully noted that MD models of DNA remain subject to assumptions in the underlying force-field.

Previous MD simulations on A-tract oligonucleotides in solution include early work from this laboratory that was based on the GROMOS force-field,[10] but required restraints to maintain intact Watson-Crick base-pairs. The results, however, supported the idea of straight, relatively rigid ApA steps, exhibiting little deformation from canonical *B*-form structural parameters. MD simulations have been used to compare the structure and dynamics of DNA dodecamers,[72] containing the sequences AAA and AIA. The results indicate that the I-substitution does more to the dynamical structure of the oligonucleotides than might be expected from just eliminating a bifurcated hydrogen bond across the major groove, and point to the importance of DNA flexibility as much as static structure in determining macroscopic behavior. Earlier MD on A-tract oligonucleotides[75] indicate as well that structures with possible bifurcated hydrogen bonding comprise only a small fraction of the MD trajectory, while other features of A-tracts, such as minor groove narrowing, remain intact.

Pastor *et al*.[76] have addressed this issue using the CHARMM 22 force-field in simulations of the TATA box sequence d(CTATAAAAGGGC) and a similar sequence substituting the A with I. They find as well that I-substitution introduces more flexibility, and conclude independently that the bifurcated H-bond in the major groove does not contribute to the stability of the A-tract over the I-tract. Interestingly, Pastor *et al*. describe very different hydration patterns in the minor groove of the two cases, despite the identical chemical nature of the two sequences, and suggest this as an account of the difference in the binding for TATA box-binding protein (TBP) of some three orders of magnitude. Flatters *et al*. report TATA box MD studies[77] and explore the effect of making a double mutation.[78] They found that the wild-type sequence is more like the A-like protein-bound

conformation, while the mutant more closely resembles to the *B*-form, and conclude that MD modeling is able to predict sequence-dependent effects consistent with experiment. In the case of the double mutant, the DNA structure is not as close to the adapted conformation as observed in the crystal structure of the bound form. However, contrary results on the *roll* of YpR steps compared to the MD reported by Pastor *et al*.[79] were noted,[78] possibly a consequence of the differences between the AMBER and CHARMM force-fields.

MD studies of DNA oligonucleotide 25-mer and 30-mer duplexes featuring phased A-tracts have been reported recently from this laboratory,[6,7] in which a series of nanosecond level MDs were carried out for various concentrations of saline solution. A 30 bp duplex composed of three 10 bp repeats of the *Bam*HI recognition sequence was simulated as a control. The MD results, for a concentration of 60 mM KCl, 10 mM $MgCl_2$ added salt plus minimal neutralizing cations, show concerted axis bending to the extent of 15.4° per A-tract, which compares favorably with the bending per turn of 18° (or 11°-21°) inferred from cyclization experiments.[18] The MD model also exhibits a progressive 5′ to 3′ narrowing of the minor groove region of A-tracts, a feature inferred from DNA footprinting.[80,81] MD was subsequently performed on the DNA duplexes $d(G_5\text{-}\{GA_4T_4C\}_2\text{-}C_5)$ and $d(G_5\text{-}\{GT_4A_4C\}_2\text{-}C_5)$ to 3.0 ns and 2.5 ns, respectively,[7] at concentrations comparable to a ligase buffer. Analysis of the results shows that the $d(G_5\text{-}\{GA_4T_4C\}_2\text{-}C_5)$ simulation exhibits enhanced curvature, whereas $d(G_5\text{-}\{GT_4A_4C\}_2\text{-}C_5)$ shows less, consistent with experiment.[82] The locus of bending dominant in the curvature in the MD structure was found in the non-A-tract region, at the central YR C15-G16 step, with an average *roll* angle of 12.8(±6.40)°. The MD A-tracts were essentially straight and did not contribute appreciably to the overall bending. The dynamic structure of $d(G_5\text{-}\{GA_4T_4C\}_2\text{-}C_5)$ exhibited minor groove deformation comprised of expansion at the 5′ end and progressive narrowing towards the 3′ end, consistent with the hydroxyl radical footprinting results on A_4T_4 and T_4A_4 sequences reported by Burkhoff & Tullius.[81] These simulations are included in the analysis described here in order to have a full representation of phased A-tracts and non-A-tract regions in longer sequences included.

Recently, Strahs & Schlick[55] reported 2 ns of MD on the dodecamers d(CGCGAAAAAACG) and d(CGCAAAAAAGCG) duplexes in water with minimal Na ions using the Cornell *et al*. force-field implemented in CHARMM. The MD structure has an overall curvature of 12-14°. The calculated curvature is characterized by a positive *roll* bend of 12° at the 5′junction region, but gradual curvature within A-tracts was noted. The MD structures of these dodecamers were thus interpreted as having features of wedge, junction, and non-A-tract bending models, and the dominant effect in the curvature of the dodecamer sequence seems to be *roll* bending at or near the 5′ junction of A-tract and flanking sequences.

The present study provides additional MD simulations in solution on nine A-tract-containing DNA oligonucleotides, the homopolymeric sequences poly(dA), and poly(dG), and two related sequences, including the CAG 30-mer, which exhibits hyper straight behavior on gels, but still readily undergoes fast cyclization associated with high bending and or bendability. All MD were performed under essentially the same simulation protocol used by Young *et al*.[6] and Sprous *et al*.[7] as described above. These simulations, combined with those of Young *et al*. and Sprous *et al*. previously reported comprise an extensive MD data base on the dynamical structure of ApA steps and A-tracts in various contexts. Subject to the accuracy of the MD model, this provides a basis for further consideration of A-tract issues independent of any particular fluctuations that arise in a particular sequence. We consider the following questions. (a) What is the dynamical structure (*B*′ form) predicted by MD modeling of A-tracts? (b) Are the properties of an ApA step transferable to A-tracts, i.e. what is the extent of context effects? (c) What does MD predict to be the relative bending and bendability of DNA base-pair steps, and how reliable are the predictions? (d) To what extent are bending and bendability correlated in MD models of DNA in solution? (e) What model of DNA bending does the MD model support: the wedge, flexible wedge, junction, or non-A-tract general-sequence model and what, if any, variations on the basic theme are suggested?

Results

A series of nine new isothermal, isobaric ensemble MD simulations were performed, treating the duplex oligonucleotides listed in Table 1. The discussion in the following section is based on these results, combined with those reported by Young *et al*.[6] and Sprous *et al*.[7] on phased A-track 25-mers and 30-mers. The average structures calculated for each of the nine new MD simulations are shown in Figure 1. In all cases, the simulations maintain reasonable *B*-form structures. The root-mean-square deviation (rmsd) of the central ten base-pair steps, heavy atoms only, of the MD structures with respect to corresponding crystal structures and other controls are collected in Table 2. For the six A-tract oligonucleotides, the average MD structures range from 2.3 Å to 2.7 Å rmsd from canonical *B*-form starting structure and 2.0 to 2.5 Å rmsd from the corresponding crystal structures. For each sequence, the average MD structures are found to be closer to the respective crystal structures than to the canonical *B*-form starting structures. The MD on poly(G) and CAG, for which there is no corresponding single crystal structure, show rmsd (central ten base-pairs) of 4.3 Å and 3.6 Å from canonical *B*, respectively, compared with MD of

Table 1. DNA sequences for which MD simulations were performed, denoting abbreviations and references to corresponding structure determinations

Name	Sequence	NDB ID
A4	d(CGCGAAAACGCG)	BDL032, BDL021
A5T	d(CGCAAAAATGCG)	BDL015
A6-1	d(CGCAAAAAAGCG)	BDL006
A6-2	d(CGCGAAAAAACG)	BDL047
A3T3	d(CGCAAATTTGCG)	BDL038
T3A3	d(CGCTTTAAAGCG)	
Poly(A)	$d(A)_{30}$	
Poly(G)	$d(G)_{30}$	
CAG	d(GCAGCAGCTCATGGCCATGCAGCAGCAGCT)	

The names of the systems used here and the sequences of one strand, along with the NDB reference of the crystal structures studied denoting name and sequence of one strand. Crystal structures BDL032 and BDL021 have nicks in the backbone at the center of the complementary strand between T18 and T19.

poly(A), which lies 2.3 Å from canonical *B*. Interestingly, the MD on poly(G) is the furthest from canonical *B* (4.3 Å rmsd), as well as being the sequence with the most propensity to convert to the *A* form. The CAG sequence, with a high C + G content, also shows a large deviation from canonical *B*, 3.6 Å rmsd.

The MD trajectories can be characterized in more detail in terms of sugar puckers ϕ, average helicoidal values of *x-displacement*, *rise*, *twist*, and minor groove *width* (Table 3). *x-displacement* and base-pair *inclination* are diagnostic of the difference between canonical *A* and *B* forms,[83] for which *x-displacement* values are −5.4 Å and −0.7 Å, respectively. *x-displacement* values for MD A-tracts fall between −1.7 Å and −1.9 Å, close to canonical *B* values. Two simulations show an average *x-displacement* further in the direction of canonical *A*-DNA, CAG (−2.5 Å) and poly(G), (−3.5 Å). The poly(G) simulation is closer to the *A* form of DNA with respect to *x-displacement* than any of the others, while the MD value for CAG lies between the *A* and *B* forms of DNA. This trend is observed in several of the other key structural properties, including *rise*, *twist* and minor groove *width*. The average *rise* of the MD structures lies closer to *B*-form DNA, ∼3.4 Å, with poly(G) having an only slightly lower value, 3.3 Å. With respect to minor groove *width*, all of the A-tracks exhibit an average *width* close to that of canonical *B*-form DNA, a consequence of expansion at the 5′ end relative to B72 balanced against narrowing at the 3′ end of the A-tract. The feature of minor groove narrowing 5′ to 3′ is well established in the MD (see below). The MD models of poly(G) and the CAG sequence show minor groove *widths* wider than canonical *B* by ∼1.0 Å, but closer to canonical *B*-form than *A*-form DNA. The average *twist* values from MDs on all sequences show reduction with respect to the canonical *B* value of 36°, with most structures under-twisted by some ∼4-5°. The MD of poly(G) shows an even larger under-twisting, ∼7°.

To elucidate conformational changes occurring 5′ to 3′ within A-tracts in the MD structures, results on relevant helicoidal values are shown in Figure 2. Here, MD average values are plotted for the first, second, third … ApA steps as they occur in all sequences studied. Corresponding average values from crystal structures are included for reference, noting that the MD results refer to a solution and

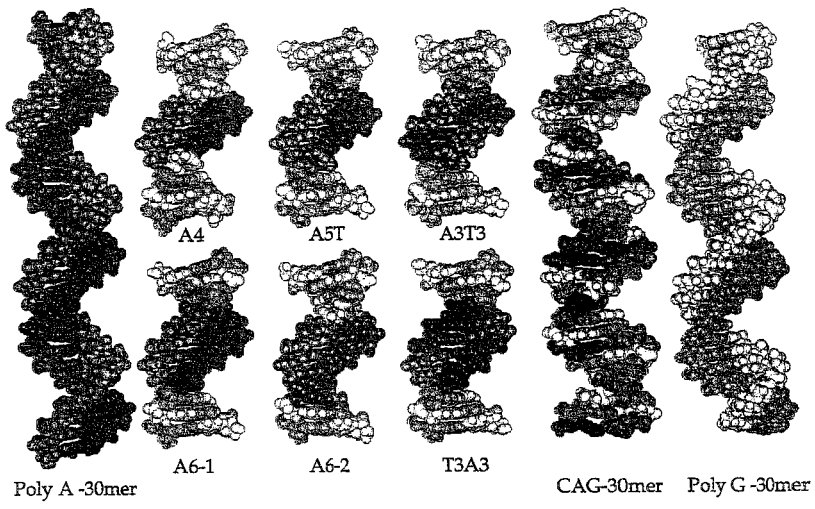

Figure 1. Average structure for 3ns trajectories MD simulations on nine oligonucleotide sequences. For the definition of abbreviations, sequences and other relevant details of the structures, see Table 1.

Table 2. Root-mean-square deviations from MD simulations to crystal structures

Reference:	A4	A5T	A61	A6-2	A3T3	T3A3	G30	CAG	A30
B72	2.68	2.44	2.67	2.27	2.38	2.81	4.31	3.58	2.28
Xtal-1	2.47	2.17	2.46	2.03	1.97				
Xtal-2	2.36	2.34							
Xtal-1-B72	1.36	1.15	1.15	1.13	1.07				
Xtal-2-B72	1.38	1.30							

Heavy-atom rmsd of the central ten residues of each of the MD average structures to canonical B72 and, where appropriate, to crystal structures.

not a crystalline state model and are not expected to be strictly in accord. However, MD structures show trends similar to those found in the crystal structures, although the changes in the crystal structures are generally more pronounced. The MD of A-tracts reproduces the characteristic large negative *propeller*, −14.1° *versus* −10.4° overall. The calculated value from the poly(A) MD is −14.0°, compared with average *propeller* from *B*-form crystal structures of −17.3°. MD reproduces the minor groove width narrowing in A-tracts compared with canonical *B*-DNA, though not quite at the level found in crystal structures. Two other properties for which the MD simulations show progressive changes are *inclination* and *slide*, cf. Figure 3. The trend in *inclination* is actually greater in the MD compared to crystal structures and is progressively more negative 5′ to 3′, i.e. towards A-form DNA values, as the number of ApA steps increases. The ApA steps also show a slight 5′ to 3′ decrease in base-pair *slide*.

The local deformations of dinucleotide steps from MD on A-tract oligonucleotides analyzed using *roll/tilt* ''bending dials''[60] are shown in Figure 3. (For additional perspectives on the use of bending dials and other methods for graphical depiction of DNA bending, see Dickerson [56].) The contour boundaries on the dials defined a region enclosing 98% of the MD data points. The MD results for A-tracts all fall near the ''bulls eye,'' of the bending dials. The calculated bending by step for A -tracts, following the curves procedure, is near 0 in wedge angle. The A-tracts in MD are essentially straight and less flexible, i.e. exhibiting lesser thermal fluctuations than other steps. The black circles indicate specific values for each step reported for the corresponding crystal structures as listed in Table 1. All of the crystal structure values fall within the contour boundary of MD values, indicating close accord between observed and calculated deformations. The results from MD in solution are completely consistent with the crystal structure data, except for the sequence A4, in which nicks on one strand in the center of the crystallographic A-tract slightly contaminate the comparison.

Results from MD on local bending at base-pair steps as a function of sequence as measured by the angles between successive segment vectors are shown in Figure 4, and permit us to examine further the extent of axis bending in A-tracts. Local bending per step determined by this measure does not carry directional information but shows points at which deformations occur. Axis deformations are most prevalent in the flanking sequences at or near the junction. Bending dials (Figure 3) indicate that this occurs primarily via a *roll* bend into the major groove in non-A-tract regions. The MD bending profile for the sequences A6-1 and A6-2 are nearly identical at the dinucleotide step level, even though the A-tract is shifted by one base-pair and the flanking sequences are altered. The bend from the A6-1 sequence appears to occur at the CA step, while the bend in A6-2, which does not have

Table 3. MD calculated helicoidal and morphological parameters for MD average structures

Sequence	x-displacement (Å)	*Rise* (Å)	*Twist* (deg.)	Minor groove width (Å)	ϕ_1	ϕ_2
A4	−1.91 (0.64)	3.35 (0.40)	31.12 (5.49)	6.35 (1.29)	131.04	126.84
A5T	−1.83 (0.61)	3.33 (0.37)	31.83 (4.72)	5.95 (1.22)	130.99	128.83
A6-1	−1.94 (0.66)	3.38 (0.37)	31.78 (4.98)	6.07 (1.27)	129.18	122.22
A6-2	−1.78 (0.51)	3.36 (0.35)	32.13 (5.16)	5.75 (1.07)	130.96	125.18
A3T3	−1.74 (0.63)	3.31 (0.33)	31.63 (5.21)	6.16 (1.28)	128.42	129.52
T3A3	−1.84 (0.57)	3.41 (0.36)	31.90 (4.85)	5.92 (1.13)	127.95	123.65
A30	−1.97 (0.42)	3.38 (0.34)	32.26 (4.36)	5.18 (1.15)	126.71	118.50
G30	−3.58 (0.95)	3.29 (0.43)	28.89 (5.58)	7.10 (1.46)	107.76	112.24
CAG	−2.48 (0.59)	3.28 (0.41)	30.69 (4.94)	6.94 (1.37)	120.91	120.37
B72	−0.70	3.38	35.81	5.89	−168.40	−168.40
A72	−5.41	2.56	30.92	10.00	13.10	13.08
H-DNA	−1.08	3.23	35.56	5.47	7.29	164.82

Average helicoidal values calculated over all base-pairs and base-pair steps with the terminal base-pair and base-pair step removed. Values in parentheses are standard deviation. ϕ_1 and ϕ_2 refer to sugar pseudorestoration angles on the two sugar phosphate backbone strands of the DNA.

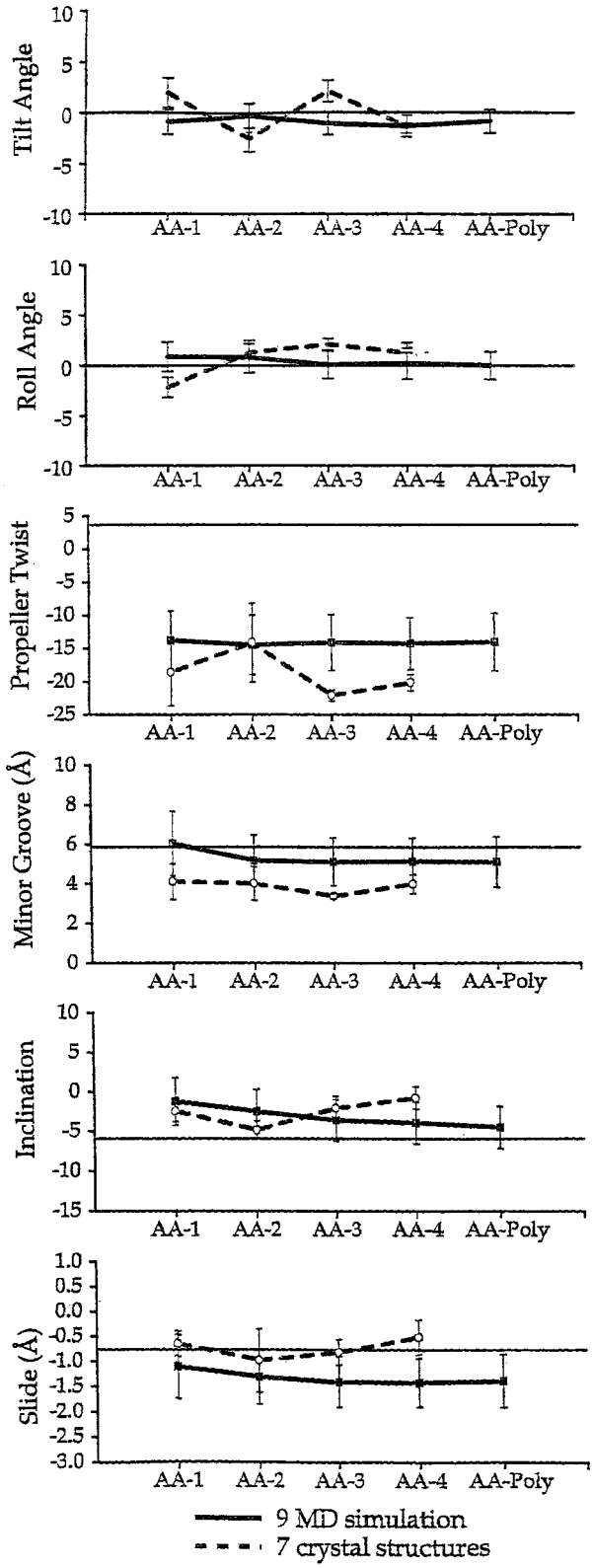

Figure 2. Calculated properties of ApA steps in MD structures of A-tract oligonucleotides with respect to position 5′ to 3′ in an A-tract. For each data point, all instances of the step from the A-tract sequences listed in Table 1 are included.

a CA step, occurs at the CG step. The poly(A) and poly(G) simulations show a constant deformation over all steps, with poly(G) slightly greater than poly(A), but with helix phasing result in an essentially straight DNA structure overall. The MD on the CAG sequence shows a large number of base step bends and kinks in the structure (Figure 4(i)). In this sequence, the steps with the largest bending are the CA·TG, TA and CG steps.

Discussion

Dynamical structure of A-tracts

The MD structures for A-tracts in solution show a high degree of correspondence with crystal structures (Table 4), and this agreement supports the idea that the essence of crystal structure results transfer successfully to the solution state. The dynamical model of A-tracts derived from the MD described here exhibits the properties high *propeller*, a narrowing of the minor groove 5′ to 3′, and essentially straight helix axes, and our MD predicts this to be the characteristic *B′* form of A-tracts in solution. The structure of A-tracts as described by either crystallography or MD varies 5′ to 3′ within a sequence, and thus it is clear that some degree of context effects influence the ApA structure. An indication of the extent of this can be seen by comparing the data as plotted in Figure 3. The variation of helicoidal parameters 5′ to 3′ for an ApA step is slight but systematic, with the largest changes noted in the helicoidal parameters *slide* and *inclination* as well as minor groove *width*. The MD results on poly(A) compared with those on short oligonucleotide sequences shows AA and AAA sequences to be differentiated, but $(A)_n$ for $n > 4$ are essentially the same as poly(A). This result is consistent with the results on the cooperativity in the A-tract unit inferred from analysis of gel data[17] and NMR base-pair opening studies.[84]

Origins of A-tract stability

The prevalence of bifurcated H-bonds in the MD model was monitored by calculating the distance between participating N and O atoms across the major and minor grooves over all structures in the ensemble. The distribution of N6-O4 distances across the major groove calculated from the MD is shown in histogram form in Figure 5. The calculated distribution is centered at 4.1 Å, close to that of canonical *B* and outside of even a generous definition of hydrogen bond length, 3.5 Å. Only 12 % of the MD structures have calculated N6-O4 interaction distances below 3.5 Å. Thus we find (see also Sherer *et al.*[72]) that the MD structures readily exhibit the general features of high *propeller*, narrow minor groove *width* associated with *B′*-DNA without the necessity of an auxiliary bifurcated hydrogen bond across the major groove. Nevertheless, Raman spectra[14] and NMR N15[30,85] experiments support this interaction, and questions

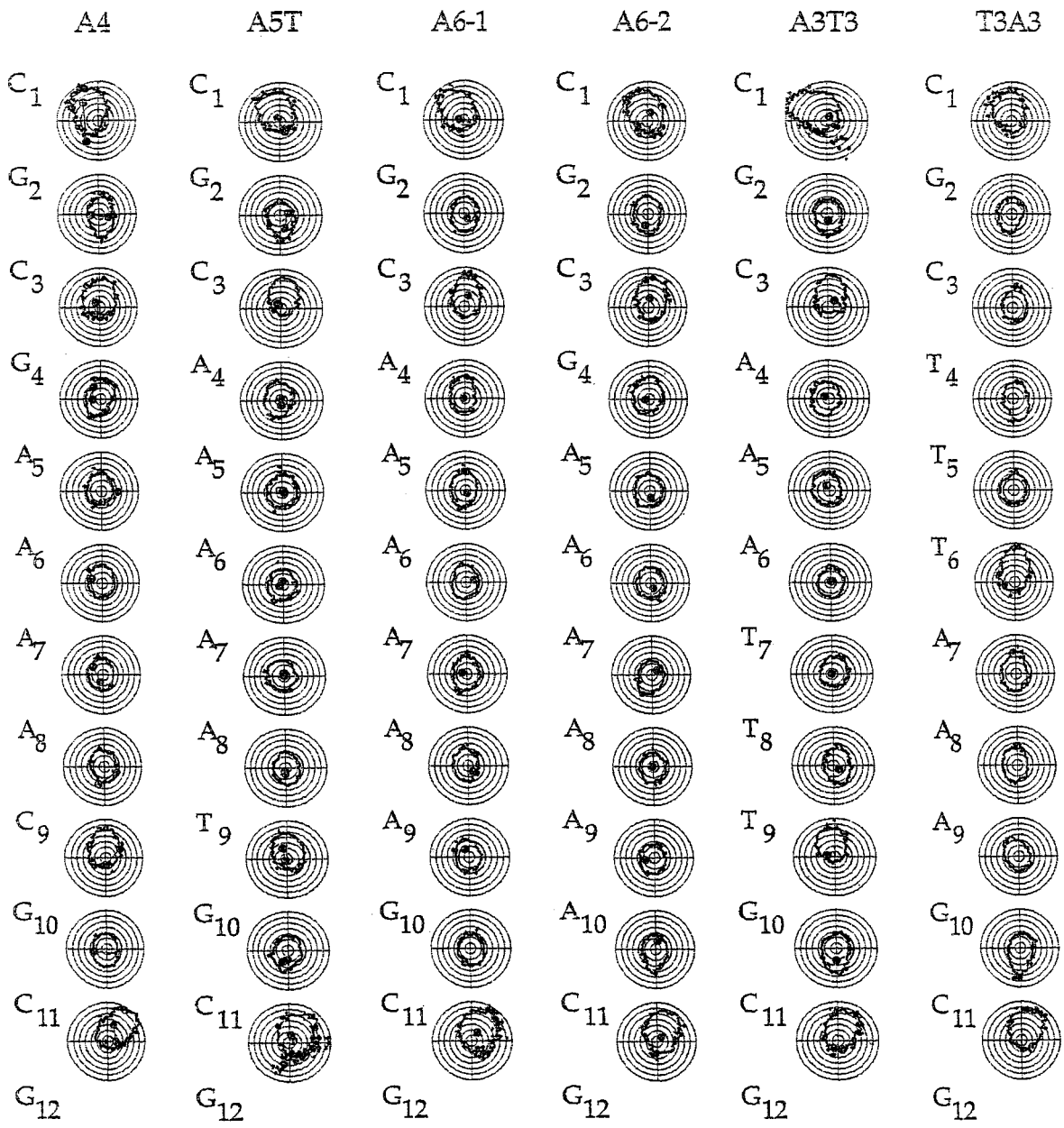

Figure 3. The *roll/tilt* bending dials for MD structures of A-tract oligonucleotides. For a full definition and description of bending dials for DNA duplexes, see Young *et al.*[60] On each dial, 0° (vertical) denotes *roll* toward the major groove, 90° and 270° denote *tilt* towards the sugar phosphate backbone, and 180° denotes *roll* toward the minor groove.

remain about whether this effect is described accurately in the MD and whether it contributes significantly to A-tract stability. Our results may be sensitive to the slight but systematic under-twisting of the helix in the MD model . With respect to the hypothesis of a minor groove bifurcated hydrogen C2H…O2,[29] the calculated distribution of A C2 and the T O2 separations across the major groove is shown in Figure 6. Here the C2-O2 interaction distances are <3.5 Å in 72 % of the structures and skewed toward shorter values. While the MD results are consistent with the hypothesis, this interaction would be expected to be relatively weak.

The suggestion that binding of mobile counterions may contribute the rigidity of the ApA step and to A-tract structure and stability has been advanced.[36] Analysis of the fraction of Na^+ binding from all the MDs surveyed in this study has been carried out, and is shown in Figure 7 broken down with respect to the ten unique dinucleotide steps. All calculated fractional occupancies of counterions in the grooves are <10 %. The only steps in the minor groove to show significant direct Na^+ binding are ApT (6.5 %) and ApA/TpT (2.5 %). In

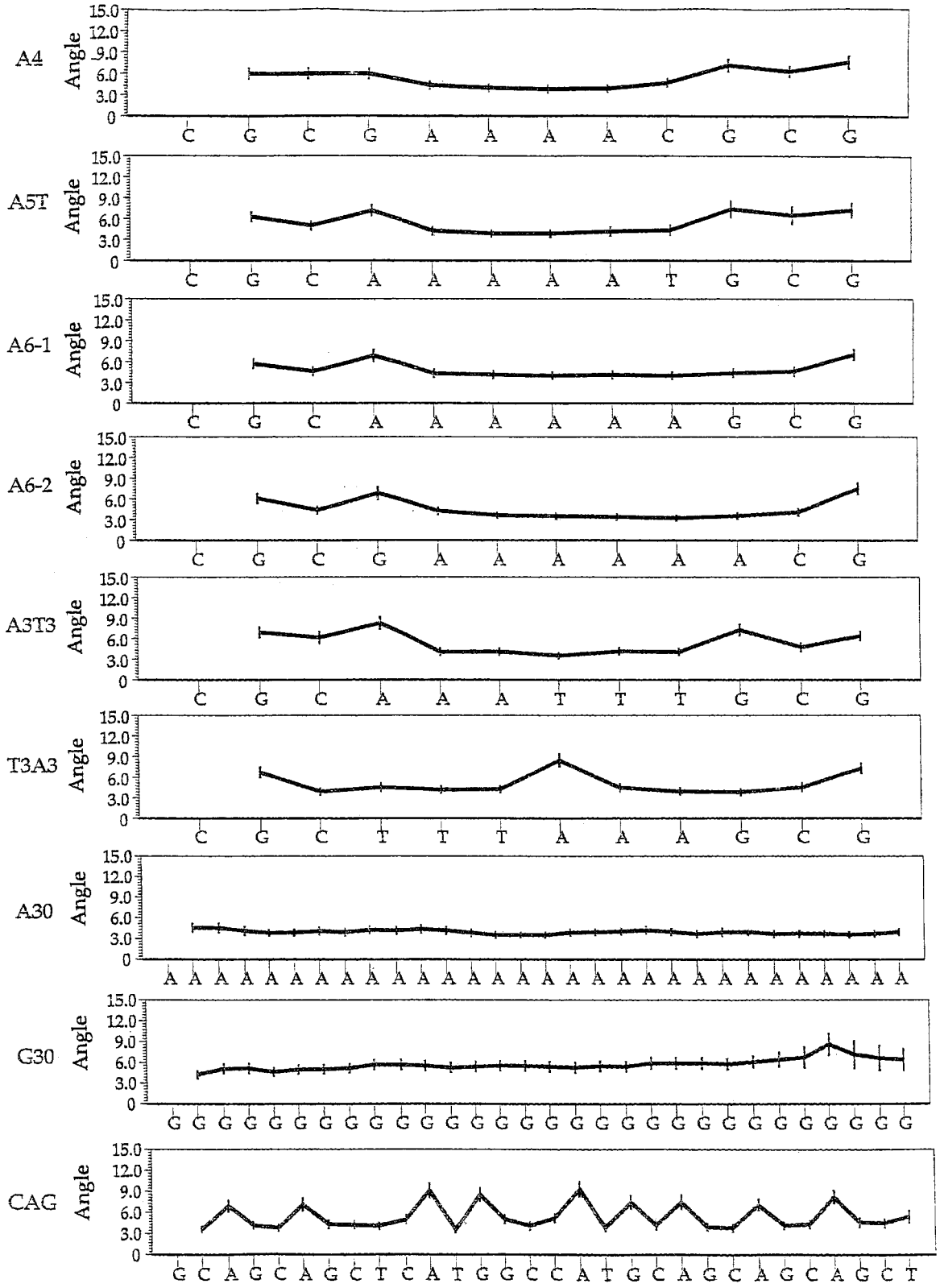

Figure 4. Calculated local bending of base-pair steps in MD structures of A-tract oligonucleotides as a function of sequence. Bending is defined as described for CURVES 5.0.[97]

the major groove, all steps show a significant increase in fraction of ions bound, the largest being and the ApT (8%) and the smallest CG (2%). Since, once again, the characteristics of *B*′-DNA of A-tracts are well reproduced in the MD, Na^+ binding is not indicated to be an appreciable factor in

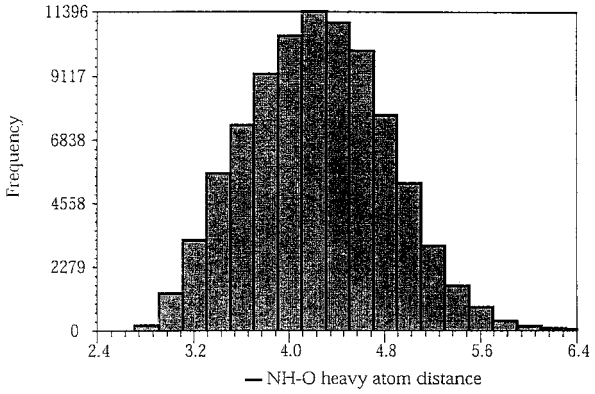

Figure 5. A histogram of N6-O4 distances across the major groove of MD structures of A-tract oligonucleotides.

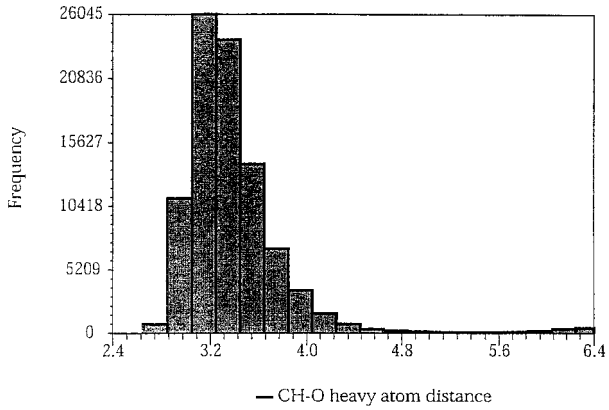

Figure 6. A histogram of C2-O2 distances across the minor groove of MD structures of A-tract oligonucleotides.

stabilizing the conformation of the ApA steps. Similar conclusions have been drawn by Shakked and co-workers from a new dodecamer structure, d(ACCGACGTCGGT).[86] However, in related work[37] we do find in MD evidence for a correlation between ion proximity at CpG bending loci in the major groove of DNA sequences, and crystal structures of oligonucleotides with heavier ions than Na^+ show higher fractional occupancies. (In making a correspondence of these results with those of Williams and co-workers,[87] it is essential to note that different definitions of the minor groove may be involved: our analysis emphasizes the floor of the groove, whereas their definition includes the lips of the groove as well.)

DNA bending and flexibility

Since all instances of the ten unique base-pair steps in DNA are represented in the sequences studied here by MD, we extracted from the collective of MD trajectories the helicoidal properties *tilt, roll* and *twist,* wedge angle and local bending angle by step. Note the number of examples of each vary somewhat, and this must be considered only a provisional view of sequence effects in DNA from MD. The results are shown in Figure 8, together with corresponding results from crystal structures of *B*-form DNA oligonucleotides for reference. The qualification that the calculated and observed sets of data do not refer to the same state and are not strictly comparable applies here as well; nevertheless, it is of interest to know how they do compare. MD and crystal structure average *tilt* values are in good overall agreement, with only small differences at the CT and CC base steps. Three dinucleotide steps TG, TA and GC have average *roll* values larger in the MD model compared with the crystal structure values, but it is important to note that these steps also have widest distribution in the crystallographic data sets. As noted above, MD *twist* values are systematically lower than those of crystal sequences. However, differences in the crystal and MD results may reflect the differences in solution and crystal environments (our unpublished results). The MD calculated helicoidal parameters of the ten unique base-pair steps show a clear trend: YR steps have a small positive *tilt*, a large positive *roll* and a large average bending angle. The RR/YY steps fall into two groups, with the CT·AG and TT steps exhibiting a smaller negative *tilt*, essentially zero *roll*, and a smaller bending angle (close to the 3° threshold value that we use to define ''essentially straight.'' We note the specific values of angles quoted are from CURVES analysis, and can be expected to vary slightly if other methods are used. The TC·GA and CC·GG

Table 4. Root-mean-square deviation (Å) (heavy atoms) of MD calculated and observed A-tracts

	A4	A5T	A6-1	A6-2	A3T3	T3A3
Ave MD-B72	1.42	1.37	1.62	1.48	1.17	1.8
Ave MD-Xtal-1	1.52	1.02	1.46	1.04	0.94	
Ave MD-Xtal-2	1.44	1.31				
Ave MD-HDNA	1.59	1.58	1.79	1.69	2.14	1.61
Xtal-1-B72	1.17	0.94	0.94	0.91	0.76	
Xtal-2-B72	1.18	1.03				
HDNA-Xtal-1	1.15	1.19	1.44	1.30	1.63	
HDNA-Xtal-2	1.41	1.28				
HDNA-B72	1.50	1.35	1.25	1.22	1.53	1.45

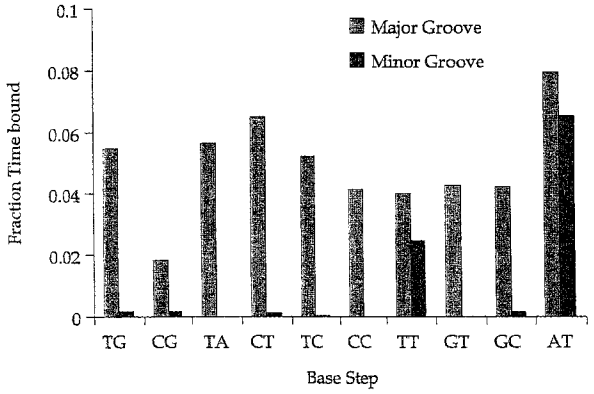

Figure 7. MD calculated fractional occupancy of Na ions as a function of base-pair step.

steps show zero *tilt*, small positive *roll* and a moderate bending angle. The three RY steps have a near zero *tilt* and *roll* value and a significantly smaller base step bending angle than YR steps. This provides an independent conformation of Zhurkin's theory[64] predicting the trend YR > RR = YY > RY, although we find the local bending in YR steps to be significantly larger, whereas the RR and RY steps are smaller and less differentiated.

MD calculated flexibilities for each step were calculated from the areas (thermal fluctuations) of *roll/tilt* data points on the bending dials of Figure 3, using a grid-based numerical method. The results, reduced to averages for each of the ten unique dinucleotide steps, are plotted as a function of sequence in Figure 9. The lowest flexibility was seen at the AT step, followed closely by a group consisting of the other RY steps GC and GT and the AA (=TT) step. For ApA steps, the thermal fluctuations appear to be distributed symmetrically with respect to both *roll* and *tilt* directions, consistent with the potentials used by Maroun & Olson,[49] but do not support the flexible wedge model[62] of DNA bending. The most flexible category of step is YR, with TA > CG > CA.[88] It is interesting to note that Stellwagen *et al*. found that a 20-mer phased CACA sequence migrated faster than a phased A-tract.[43]

In a recent study by Olson *et al*.[90] based on protein DNA crystal structures, a bendability index by base-pair step was derived. A plot of their results *versus* our MD calculated flexibility by step is shown in Figure 10. The correlation coefficient is $r = 0.74$. The MD results also provide us with the opportunity to investigate if local bending and flexibility are correlated. The local bending by step from Figure 8 is plotted *versus* the flexibility by step as listed in Figure 9, and the results are shown in Figure 11. The correlation proved to be strong, with a correlation coefficient $r = 0.89$. We noted above the implication of both bending and flexibility in the interpretation of gel retardation and cyclization data, but just how these two properties contribute to observed measures of DNA curvature is not established unequivocally. Our result, that bending and flexibility are correlated, introduces an additional useful perspective into the interpretation of experimental DNA curvature data.

CAG

Our MD model shows the CAG sequence exhibits a number of local bends, (Figure 4(i)) yet shows no overall curvature (Figure 1), consistent with the observation of no gel retardation. The analysis shows the CA step to also be one of the most flexible, consistent with the interpretation of the cyclization data in terms of the unusual flexibility of the molecule.

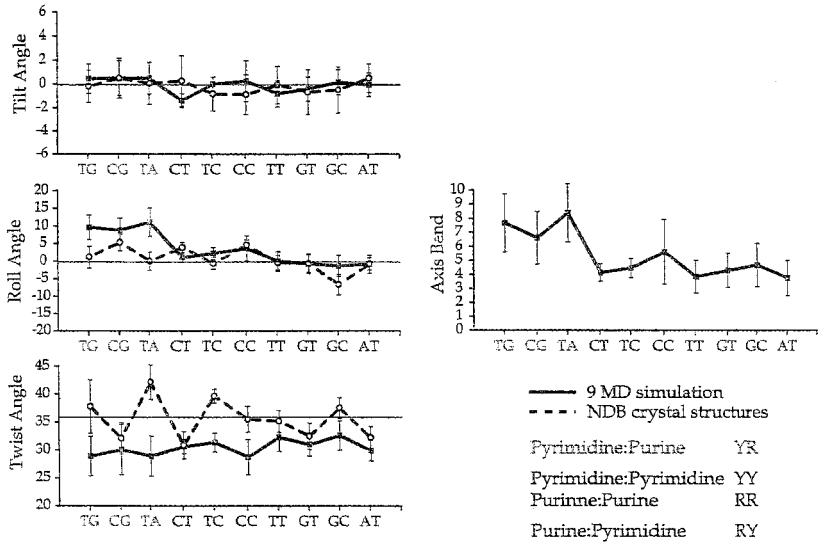

Figure 8. MD calculated *roll*, *tilt*, *twist* and bending angle as a function of base-pair step.

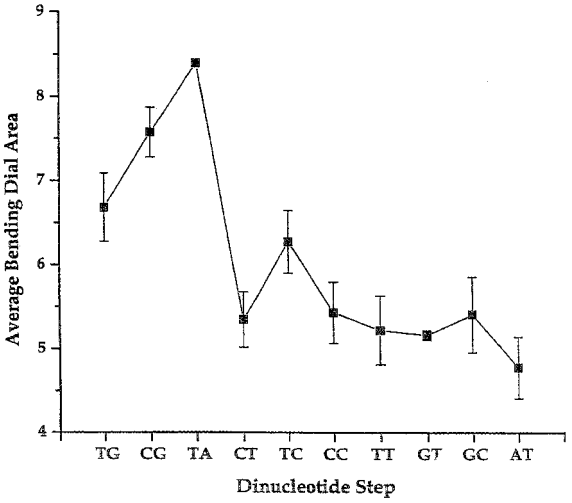

Figure 9. Calculated bendability of base-pair steps in MD structures of DNA oligonucleotides as a function of sequence.

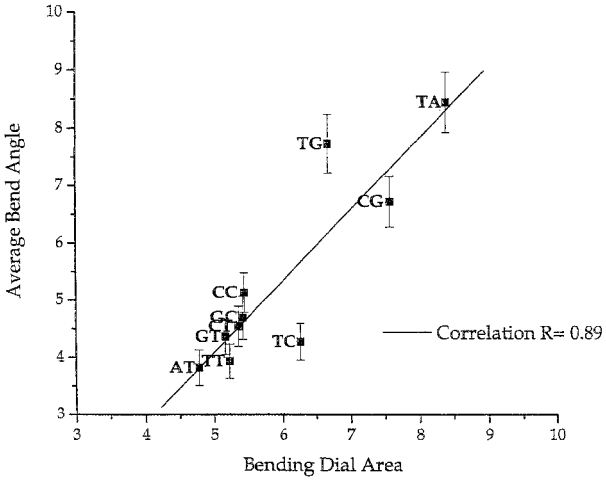

Figure 11. MD calculated bending *versus* MD calculated flexability, by step.

An MD model of DNA bending and flexibility

The wedge, junction and non-A-tract models of DNA bending were postulated *ad hoc* to account for experimentally observed bending data. In MD, the corresponding ''model'' emerges as a result of the MD, not as a postulate. MD structures for the phased A-tract and phased A_4T_4 motifs at the position of maximum curvature are available in Figure 3 of Young *et al.*[6] and Figure 2 of Sprous *et al.*[7] The MD results indicate that A-tract induced curvature originates mainly in the non A-tract region. While the MD studies described here support the non-A-tract bending model, the bending is not distributed randomly in the non-A-tract regions: the major deformations occur at YR steps, mostly CA and CG in the vicinity (often within two or three steps) of a junction between A-tracts and flanking sequences. The NMR dipolar coupling structure described by MacDonald *et al.*[12] for d(GGCAAAAAAACGC) is essentially consistent with these ideas. TpA steps in the 25-mer phased A-track sequence,[6] in T_3A_3 (current study) and T_4A_4 of Sprous *et al.*[7] also exhibit relatively large local bends. The CG steps in the MD model do show an especially large *roll* bend when adjacent to A-tracts, so sequence context here is indicated to be an important factor. The smaller degree of local bending within A-tracts, while not the primary origin of curvature, is locally uncompensated and can be expected to contribute a supplement to the non-A-tract bending, a point noted by Liu & Beveridge in a parallel study using different methodology.[25]

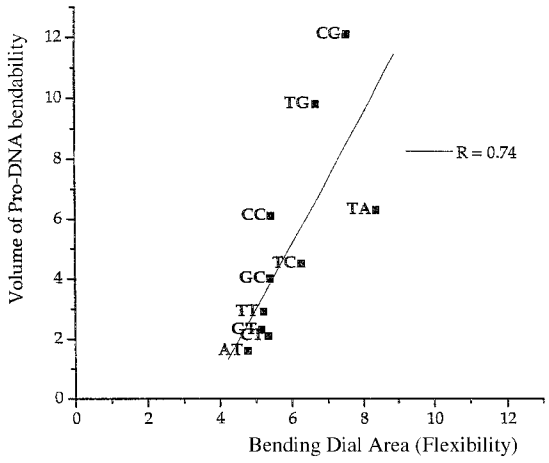

Figure 10. Flexibility for base-pair steps from MD calculations *versus* bendability estimated from crystal structures by Olson *et al.*[89]

In summary, we find that A-tract-induced bending in MD supports the non-A-tract bending paradigm,[47–49,56,90] with the main origin of bending located in deformations of YpR steps toward the major groove in non-A-tract regions. The possibility of a lesser contribution from mild uncompensated ApA deformations is not precluded, but significantly curved A-tracts and the wedge model of DNA bending are not supported by our MD, nor is a tilt bend localized at junctions between A-tracts and flanking sequences. Note, this does not preclude a junction bend induced in a DNA sequence by complexation with protein, as in the TATA-box bound to TBP.[91] One observation that appears at variance with the idea of A-tract-induced bending originating in non-A-tract regions is that of Haran *et al.*,[92] which showed that axis bending was essentially independent of the sequence variation in the non-A-tract region. However, all possible spacers except a G-tract have YpR steps, so that can explain the result. G-tracts *per se* are curved,[93] so that a phased A_5G_5 casette will be bent although lacking in YpR steps and is not a diagnostic test of the YpR origin of bending hypothesis.

On the MPD effect

If curved A-tracts are not involved in the structure, altering A-tracts from bent to a straight form is not a viable explanation of the MPD effect. However, the results of MPD on the reduction of gel retardation for phased A-tracts can be explained readily without the necessity of ''straightening'' ApA steps: all that is required to explain the effect of MPD in reducing gel retardation in phased A-tracts is that it converts *B′*- DNA into normal *B*-DNA, so that there is no anomalous local structure to phase into concerted bending. MPD, ethanol, and other aliphatic alcohols do lower the dielectric constant of the solvent and the water activity. At lower dielectric, phosphate repulsions would be stronger, making the sequence more *B*-like. However, the *A*-form of DNA has reduced phosphate separations yet is well known to be stable at lower water activity, so increased phosphate repulsion effect does not appear to be the explanation.

Dickerson[23] has advanced the idea that the effect of additives such as MPD is not specific but generic, and that the lower water activity as a consequence of MPD and other such additives converts *B′*-DNA to *B*-DNA as a consequence of reduced rigidity. We add to this the possibility that this occurs as a consequence of reduced hydrophobic pressure on base-pair stacking at lower water activity. Dickerson[23] has made an analogy between the MPD effect and the premelting transition of *B′* to *B*-DNA with increasing temperature. While the effect may be the same, the mechanism must be different. Temperature would alter the relative population of *B′* and *B* forms of A-tracts, with the *B′* form having a higher statistical weight at low temperature and the *B*-form favored at high temperature. This would be the case if the potential well for *B′* was lower than *B* in energy, but the *B* state has a broader potential and a more favorable entropy, which is quite likely to be the case. Adding MPD or some such agent and lowering the water activity does not simply alter the population of states, but changes the system: the potential wells *per se* are altered, effectively eliminating the *B′* form as a local minimum on the potential surface.

Summary and Conclusions

MD on nine A-tract and related oligonucleotide sequences ranging in length from dodecamers to 30-mers have been carried out, analyzed in detail and compared with available experimental data and ideas about the structure of A-tracts in the literature. We find the MD structures for A-tracts in solution to agree at ca 2 Å rmsd with the structures obtained for corresponding sequences in crystals, i.e. essentially straight and relatively rigid. The dynamical structure for A-tracts from MD in other respects conforms closely to that discussed in the literature as *B′*- DNA: high propeller, a narrow minor groove with a *width* that decreases 5′ to 3′ along an A-tract. The MD structure of A-tracts is not stabilized by a bifurcated hydrogen bond across the major groove, but is not inconsistent with an interaction of this type in the minor groove. Direct evidence for an interaction across the major groove has been reported from resonance Raman studies, and this discrepancy between experiment and theory on this point is noted. However, MD succeeds in reproducing the characteristics expected of *B′*-DNA without the necessity of a bifurcated hydrogen bond across the major groove. From the collective of MD results, bending and bendability are calculated by step. YR steps are predicted to be most bent and bendable. The calculated flexibility correlates well with the ligand-induced bendability derived independently from the protein DNA crystal structures. The MD results indicate that bending and flexibility at base-pair steps in DNA are highly correlated, i.e. steps that show the most intrinsic deformation from *B*-form DNA are also the most deformable. Analysis of the MD structures with respect to *A*-tract-induced axis bending shows most consistency with the non-A-tract, general-sequence model of DNA bending,[47–49] but the possibility of sequence effects in the non-A-tract region and a supplementary contribution to curvature from uncompensated small deformations at ApA steps is not precluded.

Materials and Methods

For electroneutrality, one Na^+ was included for each for each anionic phosphate group. Following placement of the counterions about a canonical *B* form of the DNA structure, each system was configured into a rectangular prism and hydrated by TIP3P water molecules in a pre-equilibrated configuration. The box dimensions were truncated to achieve a minimum distance of approximately 13 Å beyond all DNA atoms in all directions, resulting in a box size of approximately 47 Å × 47 Å × 70 Å for the 12-mer and 47 Å × 47 Å × 130 Å for the 30-mers, solvating the DNA with approximately 4200 and 7700 water molecules, respectively. The effective concentration of the DNA sample within the periodic box is approximately 10 mM (calculated by volume). In all other matters, the simulations were configured as similar as possible to each other, and to those of Young *et al*. and Sprous *et al*. Molecular interactions are described by the force-field reported by Cornell *et al*.,[4] using the TIP3P model for water.[94] Periodic boundary conditions are employed with long-range interactions calculated *via* the PME method.[3] All MD were performed using AMBER 5.0.[5] Starting DNA configurations for each case were taken as the corresponding canonical *B*-form DNA double helix.[95]

Prior to MD, the starting configuration was subjected to two rounds of energy minimization to relieve any local atomic clashes. All PME calculations were carried out with a 9 Å cutoff for direct space nonbonded calculations and a 0.000001 Ewald convergence tolerance for the long-range electrostatic interactions. Each MD was initiated with a 50 step (25 steepest descent/25 steps conjugate gradient) unrestrained minimization followed by 250 steps (100 steepest decent/150 steps conjugate gradient) of restrained minimizations followed by heat-

ing, equilibration and unrestrained (T,P,N) ensemble MD. Harmonic restraints of 25 kcal/mol per Å were placed on DNA atoms and ion positions during the second round of minimization. Heating and initial system equilibration was performed in a parallel semi-restrained fashion as follows. First, 10 ps of heating was performed on the constant volume system while restraining the DNA and ion atom locations at a 1 fs time-step. This was followed by reducing the restraints on the ions more quickly than on the DNA atoms up to 25 ps. Unrestrained constant volume dynamics were continued for another 5 ps, at which time constant temperature, constant pressure MD (isothermal-isobaric ensemble) was initiated, utilizing the Berendsen algorithm for temperature bath coupling[96] and a 2 fs time-step. The system energy was stable when coupled with a SHAKE constraint of 0.00001 Å on all covalent bonds involving hydrogen atoms. To maintain the alignment of the DNA within the center of the rectangular box, the Sander module of AMBER 5.0 was modified to remove translations and rotations of the solute center of mass at every 200 steps of MD. The center of mass energy removed by this procedure was replaced by appropriately scaling the resulting solute atom velocities. The MD simulations were performed on the SGI Origin2000 array at NCSA, and locally at Wesleyan on SGI workstations.

The direct comparison of calculated and observed results is based primarily on rmsd of heavy (non-hydrogen) atoms after the structures are placed in maximal alignment with respect to centers of mass and principle axes of the moment of inertia tensors. Conformational and helicoidal parameters were calculated using the program CURVES[97] with the local parameters option. Note that there are various coordinate frames that have been defined for calculation of structural properties of DNA, but it has been established recently that there is a high degree of correlation between the various methods, with the exception of base-pair *rise*. For a fuller discussion of these issues, see X.-J. Lu & Olson[98] and Lavery & Zakrezewska.[99] The morphological properties major groove *width*, minor groove *width* are calculated using the procedures incorporated in CURVES 5.3 involving a spline fit to the sugar-phosphate backbone. There are several possible ways of describing local deformation at a step, including the wedge angle $\sqrt{(\rho^2 + \tau^2)}$ and the angle formed between successive segment vectors in CURVES analysis, henceforth referred to as the local bending angle. We rely on both measures in this analysis, the latter being the more sensitive with respect to the magnitude of deformation compared with wedge angle. The bending dials tool[60] is a useful presentation device; for alternative options, see Dickerson[23] and Strahs & Schlick.[55]

Acknowledgments

The participation of Dr Matthew A. Young, Dr Dennis Sprous and Professor B. Jayaram is gratefully acknowledged. We acknowledge as well helpful conversations with Professors Loren Williams, Steve Harvey, Richard Lavery, Peter Kollman, Helen Berman, Nina Pastor, Z. Shakked and Mr Haribabu Arthanari. This research is supported by NIH grant # GM37909 (to D.L.B.) and an NIH Traineeship in Molecular Biophysics, grant # GM 08271 (to K.J.M.).

References

1. Dickerson, R. E., Bansal, M., Calladine, C. R., Diekmann, S., Hunter, W. N. *et al*. (1989). Definitions and nomenclature of nucleic acid structural parameters. *EMBO J.* **8**, 1-4.
2. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr, Ferguson, D. M. *et al*. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179-5197.
3. Darden, T. A., York, D. M. & Pedersen, L. G. (1993). Particle mesh Ewald: an *N* log (*N*) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089-10092.
4. Beveridge, D. L. & McConnell, K. J. (2000). Nucleic acids: theory and computer simulation, Y2 K. *Curr. Opin. Struct. Biol.* **10**, 182-196.
5. Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., III, Ross, W. S., Simmerling, C. *et al*. (1997). *AMBER: Version 5* 5.0 edit. University of California, San Francisco.
6. Young, M. A. & Beveridge, D. L. (1998). Molecular dynamics simulations of an oligonucleotide duplex with adenine tracts phased by a full helix turn. *J. Mol. Biol.* **281**, 675-687.
7. Sprous, D., Young, M. A. & Beveridge, D. L. (1999). Molecular dynamics studies of axis bending in d(G_5-(GA_4T_4C)$_2$-C_5) and d(G_5-(GT_4A_4C)$_2$-C_5): effects of sequence polarity on dna curvature. *J. Mol. Biol.* **285**, 1623-1632.
8. Dlakic, M., Parks, K., Griffith, J. D., Harvey, S. C. & Harrington, R. E. (1996). The organic crystallizing agent 2-methyl-2,4-pentanediol reduces DNA curvature by means of structural changes in A-tracts. *J. Biol. Chem.* **271**, 17911-17919.
9. Sprous, D., Zacharias, W., Wood, Z. A. & Harvey, S. C. (1995). Dehydrating agents sharply reduce curvature in DNAs containing A-tracts. *Nucl. Acids Res.* **23**, 1816-1821.
10. Young, M. A., Srinivasan, J., Goljer, I., Kumar, S., Beveridge, D. L. & Bolton, P. H. (1995). Structure determination and analysis of local bending in an A-tract DNA duplex: comparison of results from crystallography, nuclear magnetic resonance, and molecular dynamics simulation on d(CGCAAAAATGCG). *Methods Enzymol.* **261**, 121-144.
11. Tjandra, N., Tate, S.-I. , Ono, A., Kainosho, M. & Bax, A. (2000). The NMR structure of a DNA dodecamer in an aqueous dilute liquid crystalline phase. *J. Am. Chem. Soc.* **122**, 6190-6200.
12. MacDonald, D., Herbert, K., Zhang, X. T. P. & Lu, P. (2001). Solution structure of an A-tract DNA bend. *J. Mol. Biol.* **307**, 1081-1098.
13. Duguid, J., Bloomfield, V. A., Benevides, J. & Thomas, G. J., Jr (1993). Raman spectroscopy of DNA-metal complexes. I. Interactions and conformational effects of the divalent cations: Mg, Ca, Sr, Ba, Mn, Co, Ni, Cu, Pd, and Cd. *Biophys. J.* **65**, 1916-1928.
14. Chan, S. S., Austin, R. H., Mukerji, I. & Spiro, T. G. (1997). Temperature-dependent ultraviolet resonance Raman spectroscopy of the premelting state of dAdT DNA. *Biophys. J.* **72**, 1512-1520.
15. Costanzo, G., Di Mauro, E., Salina, G. & Negri, R. (1990). Attraction, phasing and neighbour effects of histone octamers on curved DNA. *J. Mol. Biol.* **216**, 363-374.

16. El Hassan, M. A. & Calladine, C. R. (1998). Two distinct modes of protein-induced bending in DNA. *J. Mol. Biol.* **282**, 331-343.
17. Koo, H.-S., Wu, H.-M. & Crothers, D. M. (1986). DNA bending at adeninethymine tracts. *Nature,* **320**, 501-506.
18. Crothers, D. M. & Shakked, Z. (1999). DNA bending by adenine-thymine tracts. In *Oxford Handbook of Nucleic Acid Structure* (Neidle, S., ed.), pp. 455-469, Oxford University Press, Oxford.
19. Sundaralingam, M. & Sekharudu, Y. C. (1988). Sequence directed DNA bending and curvature. An overview. In *Structure and Expression* (Olson, W. K., Sarma, M. H., Sarma, R. H. & Sundaralingam, M., eds), vol. 3, pp. 9-23, Adenine Press, New York.
20. Olson, W. K. & Zhurkin, V. B. (1996). Twenty years of DNA bending. In *Biological Structure and Dynamics* (Sarma, R. H. & Sarma, M. H., eds), pp. 341-370, Adenine Press, Albany, NY.
21. Hagerman, P. J. (1990). Sequence-directed curvature of DNA. *Annu. Rev. Biochem.* **59**, 755-781.
22. Sinden, R. R. (1994). *DNA Structure and Function*, Academic Press, San Diego, London.
23. Dickerson, R. E., Goodsell, D. & Kopka, M. L. (1996). MPD and DNA bending in crystals and in solution. *J. Mol. Biol.* **256**, 108-125.
24. Friedman, R. A. & Honig, B. (1995). A free energy analysis of nucleic acid base stacking in aqueous solution. *Biophys. J.* **69**, 1528-1535.
25. Liu, Y. & Beveridge, D. L. (2001). A refined prediction method for gel retardation of DNA oligonucleotides from dinucleotide step parameters:reconciliation of DNA bending models with crystal structure data. *J. Biomol. Struct. Dynam.* **18**, 505-526.
26. Coll, M., Frederick, C. A., Wang, A. H. J. & Rich, A. (1987). A bifurcated hydrogen-bonded conformation in the d(AT) base-pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc. Natl Acad. Sci. USA,* **84**, 8385-8389.
27. Nelson, C. M. H., Finch, J. T., Luisi, B. F. & Klug, A. (1987). The structure of an oligo(dA).oligo(dT) tract and its biological implications. *Nature,* **330**, 221-226.
28. Shatzky-Schwatz, M., Arbuckle, N., Eisenstein, M., Rabinovich, D., Bareket-Samish, A., Haran, T. E. *et al*. (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.
29. Ghosh, A. & Bansal, M. (1999). C-H…O hydrogen bonds in minor groove of A-tracts in DNA double helices. *J. Mol. Biol.* **294**, 1149-1158.
30. Michalczyk, R., Silks, L. A., III & Russu, I. M. (1996). Heteronuclear ^{15}N and ^{1}H magnetic resonance study of a DNA dodecamer containing an A3T3 tract. *Magn. Reson. Chem.* **34(Special Issue)**, S97-S104.
31. Chuprina, V. P. (1985). Regularities in formation of the spine of hydration in the DNA minor groove and its influence on the DNA structure. *FEBS. Letters,* **186**, 98-102.
32. Young, M. A., Jayaram, B. & Beveridge, D. L. (1997). Intrusion of counterions into the spine of hydration in the minor groove of *B*-DNA: fractional occupancy of electronegative pockets. *J. Am. Chem. Soc.* **119**, 59-69.
33. Shui, X., McFail-Isom, L., Hu, G. G. & Williams, L. D. (1998). The *B*-DNA dodecamer at high resolution reveals a spine of water on sodium. *Biochemistry,* **37**, 8341-8355.
34. Hud, N. V., Sklenar, V. & Feigon, J. (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.
35. Hud, N. V. & Feigon, J. (1997). Localization of divalent metal ions in the minor groove of DNA A-tracts. *J. Am. Chem. Soc.* **119**, 5756-5757.
36. Shui, X., Sines, C. C., McFail-Isom, L., VanDerveer, D. & Williams, L. D. (1998). Structure of the potassium form of CGCGAATTCGCG: DNA deformation by electrostatic collapse around inorganic cations. *Biochemistry,* **37**, 16877-16887.
37. McConnell, K. J. & Beveridge, D. L. (2001). DNA structure: what's in ''charge''? *J. Mol. Biol.* **304**, 803-820.
38. Denisov, V. P. & Halle, B. (2000). Sequence-specific binding of counterions to *B*-DNA. *Proc. Natl Acad. Sci. USA,* **97**, 629-633.
39. McFail-Isom, L., Sines, C. C. & Williams, L. D. (1999). DNA structure: cations in charge? *Curr. Opin. Struct. Biol.* **9**, 298-304.
40. Williams, L. D. & Maher, L. J. (2000). Electrostatic mechanisms of DNA deformation. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 497-521.
41. Rouzina, I. & Bloomfield, V. A. (1998). DNA bending by small mobile, multivalent cations. *Biophys. J.* **74**, 3152-3164.
42. Bonvin, A. M. (2000). Localisation and dynamics of sodium counterions around DNA in solution from molecular dynamics simulation. *Eur. Biophys. J.* **29**, 57-60.
43. Stellwagen, N. C., Magnusdottir, S., Gelfi, C. & Righetti, P. G. (2001). Preferential counterion binding to A-tract DNA oligomers. *J. Mol. Biol.* **305**, 1025-1033.
44. Bolshoy, A., McNamara, P., Harrington, R. E. & Trifonov, E. N. (1991). Curved DNA without A-A: experimental estimation of all 16 DNA wedge angles. *Proc. Natl Acad. Sci. USA,* **88**, 2312-2316.
45. DeSantis, P., Pelleschi, A., Savino, M. & Scipioni, A. (1990). Validity of the nearest neighbor approximation in the evaluation of electrophoretic manifestations of DNA curvature. *Biochemistry,* **29**, 9269-9273.
46. Ulanovsky, L. E. & Trifonov, E. N. (1987). Estimation of wedge components in curved DNA. *Nature,* **326**, 720-722.
47. Calladine, C. R., Drew, H. R. & McCall, M. J. (1988). The intrinsic curvature of DNA in solution. *J. Mol. Biol.* **201**, 127-137.
48. Goodsell, D. S., Kaczor, Grzeskowiak M. & Dickerson, R. E. (1994). The crystal structure of C-C-A-T-T-A-A-T-G-G. Implications for bending of *B*-DNA at T-A steps. *J. Mol. Biol.* **239**, 79-96.
49. Maroun, R. C. & Olson, W. K. (1988). Base sequence effects in double-helical DNA III average properties of curved DNA. *Biopolymers,* **27**, 585-603.
50. Trifonov, E. N. (1980). Sequence-dependent deformational anisotropy of chromatin DNA. *Nucl. Acids Res.* **8**, 4041-4053.
51. Ulanovsky, L., Bodner, M., Trifonov, E. N. & Choder, M. (1986). Curved DNA: design, synthesis, and circularization. *Proc. Natl Acad. Sci. USA,* **83**, 862-866.
52. Wu, H.-M. & Crothers, D. (1984). The locus of sequence-directed and protein-induced DNA bending. *Nature,* **308**, 509-513.

53. Zinkel, S. S. & Crothers, D. M. (1987). DNA bend direction by phase sensitive detection. *Nature,* **328**, 178-181.
54. Koo, H. S. & Crothers, D. M. (1988). Calibration of DNA curvature and a unified description of sequence-directed bending. *Proc. Natl Acad. Sci. USA,* **85**, 1763-1767.
55. Strahs, D. & Schlick, T. (2000). A-tract bending: insights into experimental structures by computational models. *J. Mol. Biol.* **301**, 643-663.
56. Goodsell, D. S. & Dickerson, R. E. (1994). Bending and curvature calculations in *B*-DNA. *Nucl. Acids Res.* **22**, 5497-5503.
57. Hardwidge, P. R. & Maher, L. J. (2001). Experimental evaluation of the Liu-Beveridge dinucleotide step model of DNA structure. *Nucl. Acids Res.* **29**, 2619-2625.
58. Harvey, S. C., Dlakic, M., Griffith, J., Harrington, R., Park, K., Sprous, D. & Zacharias, W. (1995). What is the basis of sequence-directed curvature in DNAs containing A tracts? *J. Biomol. Struct. Dynam.* **13**, 301-307.
59. Ganunis, R. M., Hong, G. & Tullius, T. D. (1996). Effect of the crystallizing agent 2-methyl-2,4-pentanediol on the structure of adenine tract DNA in solution. *Biochemistry,* **35**, 13729-13732.
60. Young, M. A., Ravishanker, G., Beveridge, D. L. & Berman, H. M. (1995). Analysis of local helix bending in crystal structures of DNA oligonucleotides and DNA-protein complexes. *Biophys. J.* **68**, 2454-2468.
61. Chiu, T. K., Kaczor-Grzeskowiak, M. & Dickerson, R. E. (1999). Absence of minor groove monovalent cations in the crosslinked dodecamer C-G-C-G-A-A-T-T-C-G-C-G. *J. Mol. Biol.* **292**, 589-608.
62. Olson, W. K., Marky, N. L., Jernigan, R. L. & Zhurkin, V. B. (1993). Influence of fluctuations on DNA curvature: a comparison of flexible and static wedge models of intrinsically bent DNA. *J. Mol. Biol.* **232**, 530-554.
63. Ulyanov, N. B. & Zhurkin, V. B. (1984). Seqence-dependent anisotropic flexibility of *B*-DNA: a conformational study. *J. Biomol. Struct. Dynam.* **2**, 361-385.
64. Zhurkin, V. B. (1985). Sequence-dependent bending of DNA and phasing of nucleosomes. *J. Biomol. Struct. Dynam.* **2**, 785-804.
65. Suzuki, M., Amano, N., Kakinuma, J. & Tateno, M. (1997). Use of a 3D structure data base for understanding sequence-dependent conformational aspects of DNA. *J. Mol. Biol.* **274**, 421-435.
66. Sanghani, S. R., Zakrzewska, K., Harvey, S. C. & Lavery, R. (1996). Molecular modelling of $(A4T4NN)_n$ and $(T4A4NN)_n$: sequence elements responsible for curvature. *Nucl. Acids Res.* **24**, 1632-1637.
67. Cheatham, T. E., III, Miller, J. L., Fox, T., Darden, T. A. & Kollman, P. A. (1995). Molecular dynamics simulations on solvated biomolecular systems: the particle mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. *J. Am. Chem. Soc.* **117**, 4193-4194.
68. Young, M. A., Ravishanker, G. & Beveridge, D. L. (1997). A 5-nanosecond molecular dynamics trajectory for *B*-DNA: analysis of structure, motions and solvation. *Biophys. J.* **73**, 2313-2336.
69. Bevan, D. R., Li, L., Pedersen, L. G. & Darden, T. A. (2000). Molecular dynamics simulations of the d(CCAACGTTGG)2 decamer: influence of the crystal environment. *Biophys. J.* **78**, 668-682.
70. Feig, M. & Pettitt, B. M. (1999). Modeling high-resolution hydration patterns in correlation with DNA sequence and conformation. *J. Mol. Biol.* **286**, 1075-1095.
71. Young, M. A., Jayaram, B. & Beveridge, D. L. (1998). Local dielectric environment of *B*-DNA in solution: results from a 14 nanosecond molecular dynamics trajectory. *J. Phys. Chem. ser. B,* **102**, 7666-7669.
72. Sherer, E. C., Harris, S. A., Soliva, R., Orozco, M. & Laughton, C. A. (1999). Molecular dynamics studies of DNA A-tract structure and flexibility. *J. Am. Chem. Soc.* **121**, 5981-5991.
73. Cheatham, T. E., III & Kollman, P. A. (1996). Observation of the *A*-DNA to *B*-DNA transition during unrestrained molecular dynamics in aqueous solution. *J. Mol. Biol.* **259**, 434-444.
74. Sprous, D., Young, M. A. & Beveridge, D. L. (1998). Molecular dynamics studies of the conformational preferences of a DNA double helix in water and in an ethanol/water mixture: theoretical considerations of the *A/B* transition. *J. Phys. Chem.* **102**, 4658-4667.
75. Fritsch, V. & Westhof, E. (1991). Three-center hydrogen bonds in DNA: molecular dynamics of poly(dA).poly(dT). *J. Am. Chem. Soc.* **113**, 8271-8277.
76. Pastor, N., MacKerell, A. D., Jr. & Weinstein, H. (1999). TIT for TAT: the properties of inosine and adenosine in TATA box DNA. *J. Biomol. Struct. Dynam.* **16**, 787-810.
77. Flatters, D., Young, M. A., Beveridge, D. L. & Lavery, R. (1997). Conformational properties of the TATA box binding sequence of DNA. *J. Biomol. Struct. Dynam.* **14**, 1-9.
78. Flatters, D. & Lavery, R. (1998). Sequence-dependent dynamics of TATA-box binding sites. *Biophys. J.* **75**, 372-381.
79. Pastor, N., Pardo, L. & Weinstein, H. (1997). Does TATA matter? A structural exploration of the selectivity determinants in its complexes with TATA box-binding protein. *Biophys. J.* **73**, 640-652.
80. Burkhoff, A. M. & Tullius, T. D. (1987). The unusual conformation adopted by the adenine tracts in kinetoplast DNA. *Cell,* **48**, 935-943.
81. Burkhoff, A. M. & Tullius, T. D. (1988). Structural details of an adenine tract that does not cause DNA to bend. *Nature,* **331**, 455-456.
82. Hagerman, P. J. (1986). Sequence-directed curvature of DNA. *Nature,* **321**, 449-450.
83. Saenger, W. (1984). *Principles of Nucleic Acid Structure*, Springer Verlag, New York.
84. Moe, J. G., Folta-Stogniew, E. & Russu, I. M. (1995). Energetics of base-pair opening in a DNA dodecamer containing an A3T3 tract. *Nucl. Acids Res.* **23**, 1984-1989.
85. Michalczyk, R. & Russu, I. M. (1999). Rotational dynamics of adenine amino groups in a DNA double helix. *Biophys. J.* **76**, 2679.
86. Hizver, J., Rozenberg, H., Frolow, F., Rabinovich, D. & Shakked, Z. (2001). DNA bending by an adenine-thymine tract and its role in gene regulation. *Proc. Natl Acad. Sci. USA,* **98**, 8490-8495.
87. Sines, C. C., McFail-Isom, L., Howerton, S. B., VanDerveer, D. & Williams, L. D. (2000). Cations mediate *B*-DNA conformational heterogeneity. *J. Am. Chem. Soc.* **122**, 11048-11056.
88. Dickerson, R. E. & Chiu, T. K. (1997). Helix bending as a factor in protein/DNA recognition. *Biopolymers,* **44**, 361-403.

89. Olson, W. K., Gorin, A. A., Lu, X. J., Hock, L. M. & Zhurkin, V. B. (1998). DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl Acad. Sci. USA,* **95**, 11163-11168.
90. Goodsell, D. S., Kopka, M. L., Cascio, D. & Dickerson, R. E. (1993). Crystal Structure of CATGGCCATG and its implications for A-tract bending models. *Proc. Natl Acad. Sci. USA,* **90**, 2930-2934.
91. Guzikevich-Guerstein, G. & Shakked, Z. (1996). A novel form of the DNA double helix imposed on the TATA-box by the TATA-binding protein. *Nature Struct. Biol.* **3**, 32-37.
92. Haran, T. E., Kahn, J. D. & Crothers, D. M. (1994). Sequence elements responsible for DNA curvature. *J. Mol. Biol.* **244**, 135-143.
93. Milton, D. L., Caspar, M. L., Wills, N. M. & Gesteland, R. F. (1990). Guanine tracts enhance sequence dependent DNA bends. *Nucl. Acids Res.* **18**, 817-820.
94. Jorgensen, W. L. (1981). Transferable intermolecular potential functions for water, alcohols and ethers. Application to liquid water. *J. Am. Chem. Soc.* **103**, 335-340.
95. Arnott, S. & Hukins, D. W. L. (1972). Optimized parameters for *A* and *B* DNA. *Biochem. Biophys. Res. Commun.* **47**, 1504-1510.
96. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F. & DiNola, A. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684-3690.
97. Lavery, R. & Sklenar, H. (1989). Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dynam.* **6**, 655-667.
98. Lu, X.-J. & Olson, W. K. (1999). Resolving the discrepancies among nucleic acid conformational analyses. *J. Mol. Biol.* **285**, 1563-1575.
99. Lavery, R. & Zakrzewska, K. (1999). Base and base-pair morphologies, helical parameters, and definitions. In *Oxford Handbook of Nucleic Acid Structure* (Neidle, S., ed.), pp. 39-76, Oxford University Press, Oxford.

Edited by B. Honig

(Received 11 October 2001; received in revised form 28 June 2001; accepted 2 July 2001)