Sequence-Dependent Dynamics of TATA-Box Binding Sites

Delphine Flatters and Richard Lavery
Laboratoire de Biochimie Théorique, CNRS UPR 9080, Institut de Biologie Physico-Chimique, Paris 75005, France

ABSTRACT We have carried out two nanosecond-length molecular dynamics simulations on a DNA oligomer, d(GCGTAAAAAAACGC)$_2$, which contains a weak binding site for the TATA-box binding protein. An analysis of the resulting trajectories shows that this oligomer behaves differently from a related oligomer [d(GCGTATATAAAACGC)$_2$] studied earlier using the same protocol (Flatters, D., M. Young, D. L. Beveridge, and R. Lavery. 1997. Conformational properties of the TATA-box binding sequence of DNA. *J. Biomol. Struct. & Dyn.* 14:757–765), and which contains a strong binding site for the same protein. The two basepair mutations that relate these oligomers lead to significant changes in time-averaged structure and in dynamic behavior, which extend over entire length of the oligomer and appear to be compatible with the experimentally observed decrease of binding and functional activity. These results suggest that molecular dynamics simulations, taking into account explicit solvent and counterions, and avoiding the truncation of electrostatic interactions, are a powerful tool for investigating the indirect aspects of protein-nucleic acid recognition.

INTRODUCTION

It is now becoming generally accepted that nucleic acids play a more active role in the formation of protein-nucleic acid complexes than was originally supposed. In the case of DNA, protein recognition goes well beyond the simple location of appropriate hydrogen bonding sites on the bases within the grooves of the double helix. It also involves more subtle and indirect elements, which bring into play the fine structure, flexibility, and dynamic properties of the double helix as a function of its base sequence. These aspects of the recognition process are, however, much more difficult to characterize and to predict than so-called direct sequence readouts.

Today, molecular dynamics simulations in realistic solvent and counterion environments are becoming a powerful technique to address such problems. We have recently used such simulations to investigate the behavior of a DNA oligomer containing a strong binding site for the TATA-box binding protein (TBP). This protein, which is the DNA binding element of the factor TF11D, plays an important role in the first stages of eukaryotic transcription. It is also a good candidate for probing DNA structural properties since a number of crystallographic studies (Kim et al., 1993a, b; Kim and Burley, 1994; Juo et al., 1996) have shown that its interaction with DNA leads to very important deformations. These structural changes involve a local transition to a form close to that of A-DNA, but differing by strong positive basepair inclination (Guzikevich-Guerstein and Shakked, 1996). This is coupled to a dramatic widening of the minor groove at the site of protein binding and to an overall bend of ~90°.

Received for publication 18 February 1998 and in final form 17 April 1998.

Address reprint requests to Dr. Richard Lavery, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, Paris 75005, France. Tel.: 33-1-4325-2609; Fax: 33-1-4329-5645; E-mail: richard@ibpc.fr.

0006-3495/98/07/372/10 $2.00

We have recently shown that these changes can be reproduced by stretching the protein binding region of DNA by its 3′-ends, locally creating a ribbonlike conformation (Lebrun et al., 1997). This implies that the protein can exert a stretching force on the double helix and probe its mechanical response in the same way that has been achieved in nanomanipulation experiments on polymeric DNA (Cluzel et al., 1996; Smith et al., 1996; Lebrun and Lavery, 1996).

To understand more about the intrinsic properties of DNA containing a TBP binding site, we recently carried out nanosecond scale molecular dynamics simulations on 15-nucleotide-long oligomer d(GCGTATATAAAACGC)$_2$. This sequence is related to that used in the crystallographic studies from the Sigler group (Kim et al., 1993b) and contains the TATATAAA binding site at its center. (The duplex region used in the crystallographic study was extended by a total of three GC basepairs to improve the stability of the oligomer during our simulations and to reduce the impact of end effects within the TBP binding region.) This study (Flatters et al., 1997a), was carried out using AMBER version 4.1 (Weiner et al., 1986; Cornell et al., 1995) in the presence of explicit solvent and counterions and using particle mesh Ewald summations to avoid the simple truncation of electrostatic interactions (Darden et al., 1993; Cheatham et al., 1995). It led to unexpected results. Starting from a canonical B-DNA conformation, the center of this oligomer underwent a spontaneous transition to a conformation between that of A-DNA and that of the DNA observed in the TBP-DNA co-crystals. These changes involved both important modifications in sugar pucker and helical parameters and strong transient bending toward the central major groove. A partial transition from B-DNA toward an A-like conformation was particularly unexpected in the light of recent simulations using AMBER 4.1 with very similar protocols, which showed that DNA oligomers with other sequences remained stable in B-like conformations (Duan et al., 1997; Young et al., 1997; Young and

TABLE 1 Experimental properties of the two binding sequences studied

Sequence	Bending (°)	Lifetime (min)	Relative transcription
TATATAA	106	185	0.73 0.77
TAAAAAA	<34	1	0.11 0.01

Relative transcriptional activity was measured with respect to 1.0 for the sequence TATAAAA using yeast TBP (Starr et al., 1995) or yeast TF11D (Wobbe and Struhl, 1990).

Beveridge, 1997), or could even undergo spontaneous A-to-B transitions (Cheatham and Kollman, 1996).

Our study therefore suggested that the oligomer containing a strong TBP binding site had rather special dynamic properties that predisposed it toward binding the protein and undergoing the necessary structural deformations. To test this result, we decided to carry out simulations on a modified sequence that binds TBP much more weakly. We chose to make these modifications small in order to test the sensitivity of the simulation protocol. Our sequence mutations were limited to two TA pairs at positions 6 and 8 of the original oligomer that were modified into AT pairs. This led to the 15-mer d(GCGTAAAAAAAACGC)$_2$, which contains a very weak TBP binding site, TAAAAAAA. Experimentally, as shown in Table 1, this sequence leads to a TBP complex with a lifetime more than two orders of magnitude shorter than that of the strong site, bends DNA by <34°, and leads to an important reduction in transcriptional activity with TBP or TF11D (Starr et al., 1995; Wobbe and Struhl, 1990).

We have performed two-nanosecond-long simulations on this new oligomer and have indeed found that two punctual changes in sequence lead to clearly modified average conformation and dynamic behavior which are, moreover, compatible with weaker TBP binding.

MATERIALS AND METHODS

Starting conformations

Starting B-DNA conformations for the 15-mer d(GCGTAAAAAAAACGC)$_2$ (termed S2, see Table 2) were generated by energy minimization using the internal/helicoidal variable program JUMNA (Lavery et al., 1995). Solvent and counterion electrostatic damping were introduced via a sigmoidal distance-dependent dielectric function and halved phosphate charges (Lavery et al., 1995; Hingerty et al., 1985). After minimization, this structure contained only C2′-endo sugar conformations. Since our previous study of the strong TBP binding sequence d(GCGTATATAAAACGC)$_2$ (S1) showed a tendency to adopt an A-like conformation in the TBP binding site (underlined), we repeated the strategy used earlier and generated a second starting conformation by forcing the sugars of the underlined zone of the S2 oligomer to adopt C3′-endo puckers [as in the crystallographic TBP-DNA complexes, the 3′-terminal sugars of the binding site region were left in C2′-endo conformations (Kim et al., 1993b)].

These starting conformations, termed respectively "B" and "BAB," were then introduced into AMBER and 28 Na^+ counterions were added to neutralize the net phosphate charges. The position of these ions was energetically optimized using a Monte Carlo procedure (Jayaram et al., 1990; Young et al., 1997), again employing a sigmoidal dielectric function. The system was then solvated so that there was a roughly 15-Å-thick layer of water around DNA (ions were initially located roughly 3–11 Å from DNA). This led to boxes of dimension roughly 80 Å × 50 Å × 50 Å containing >4500 TIP3P water molecules (Jorgensen, 1981; Jorgensen et al., 1983, see Table 2 for details). These systems were energy minimized for roughly 2000 cycles using a combination of steepest descent and conjugate gradient algorithms and applying periodic boundary conditions. Electrostatic energy was calculated using the particle mesh Ewald summation with a 1-Å grid spacing and a 10^{-6} convergence criterion (Darden et al., 1993; Cheatham et al., 1995). Lennard-Jones interactions were truncated at 9.0 Å. During the energy minimization, the conformation of the oligomer and the position of the counterions were maintained with harmonic restraints. The force constants used for this purpose were diminished as the minimization progressed. The final 500 cycles of minimization were performed without restraints.

Dynamics simulations

Calculations were performed with the AMBER 4.1 program (Weiner et al., 1986; Pearlman et al., 1995) using the PARM 94 force field (Cornell et al., 1995). It is remarked that, using the same implicit solvent model described above, this new force field leads to very similar conformations for the canonical forms of DNA as those obtained with the FLEX force field of JUMNA (Flatters et al., 1997b).

For each of the minimized systems, heating to 300 K was carried out during 5 ps, rescaling the velocities as necessary (up to 150 K) and then coupling the system to a heat bath using the Berendsen algorithm (Berendsen et al., 1984). During heating, harmonic restraints were imposed on the atomic positions of the oligomer and the sodium counterions. After an equilibration period of 5 ps at 300 K, these restraints were then slowly relaxed over a period of 25 ps until a free system was achieved. From this phase onward the simulation was performed under NPT conditions. A final equilibration of the system was carried out for a further 5 ps, regularly removing the overall translation and rotation of the oligomer (every 0.2 ps). All bond lengths involving hydrogen atoms were restrained using the SHAKE algorithm (Ryckaert et al., 1977; van Gunsteren and Berendsen, 1977) and a time step of 2 fs was used. The total duration of the two simulations was respectively 1.11 ns and 0.96 ns. We also remark that since our earlier publication (Flatters et al., 1997a), the two simulations performed on the S1 oligomer have both been slightly extended to reach 1.0 ns. In discussing the present simulations and comparing them with our earlier results we will use the nomenclature given in Table 2.

TABLE 2 Notation used to describe the dynamic trajectories and details of the systems simulated

Simulation	Sequence	Box size (Å)	No. of Waters	Duration (ns)
S1-B	GCGTATATAAAACGC	80 × 45 × 45	4441	1.00
S1-BAB	"	80 × 48 × 42	4344	1.00
S2-B	GCGTAAAAAAAACGC	80 × 48 × 48	4601	1.11
S2-BAB	"	80 × 45 × 48	4619	0.96

Mixed sugar pucker starting conformations involved C3′-endo sugars for the underlined nucleotides (excluding the terminal 3′ nucleotides).

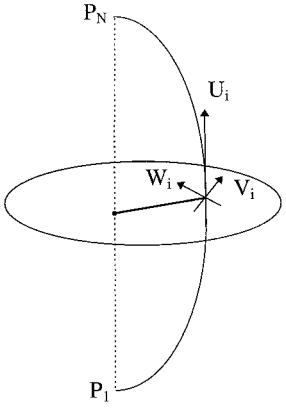

FIGURE 1 Bending magnitude and direction are measured with respect to the pseudodyad vector at basepair level i (pointing into the major groove), using the intersection of the P_1-P_N vector with the V_iW_i plane passing through P_i.

Analysis

Conformations resulting from each simulation were stored at intervals of 1 ps and were analyzed using CURVES (Lavery and Sklenar, 1989). The results presented here include both local variables (which describe the relative position of successive basepair steps) and global variables (which relate the basepairs to an overall, optionally curved, helical axis). The helical axis calculated with CURVES consists of a helical reference system for each basepair level. These systems comprise a reference point P_i and three orthogonal vectors, the helical axis vector U_i, the pseudodyad axis V_i (pointing into the major groove), and a final vector W_i, roughly aligned with the long axis of the basepairs and pointing toward the phosphodiester backbone. Overall helical bending is described by calculating the angle formed between the first and last helical axis vectors of the oligonucleotide. In order to verify that this angle was not strongly influenced by perturbations of the terminal basepairs, we made an alternative calculation based on the angle between the vectors formed by the reference points P_1-P_2 and P_{N-1}-P_N. These two calculations agreed to within ~5°.

To describe bending direction, we constructed a vector running between the terminal reference points P_1 and P_N and then, at each basepair level, calculated the point at which this vector intersected the plane defined by V_i and W_i. The direction of this intersection with respect to the V_i axis (measured in a right-handed sense around U_i) gave the bending direction in the local frame (see Fig. 1). Values close to zero consequently imply bending into the major groove, while values close to 180° imply bending into the minor groove.

RESULTS

We present the results concerning the S2 sequence by looking first at the overall conformation of the oligomer during the two simulations we have carried out, and then by moving to finer conformational details and to deformations such as helix bending. Each section includes a comparison with the corresponding results on the S1 sequence.

RMSD analysis of the trajectories

We begin by considering the overall helical conformation obtained during the simulations of the S2 oligomer. In our earlier article, on the S1 oligomer, we used RMSD calculations with respect to canonical A- and B-DNA conformations to show that the S1-B trajectory drifted away from the B-form and toward the A-form, whereas the S1-BAB trajectory, with C3′-endo sugars in the binding site region, appeared to be more stable and stayed closer to the A form (Flatters et al., 1997a). The corresponding results for the S2 oligomer (and for the extended trajectories of S1) are given in Fig. 2. These results show several important points. For

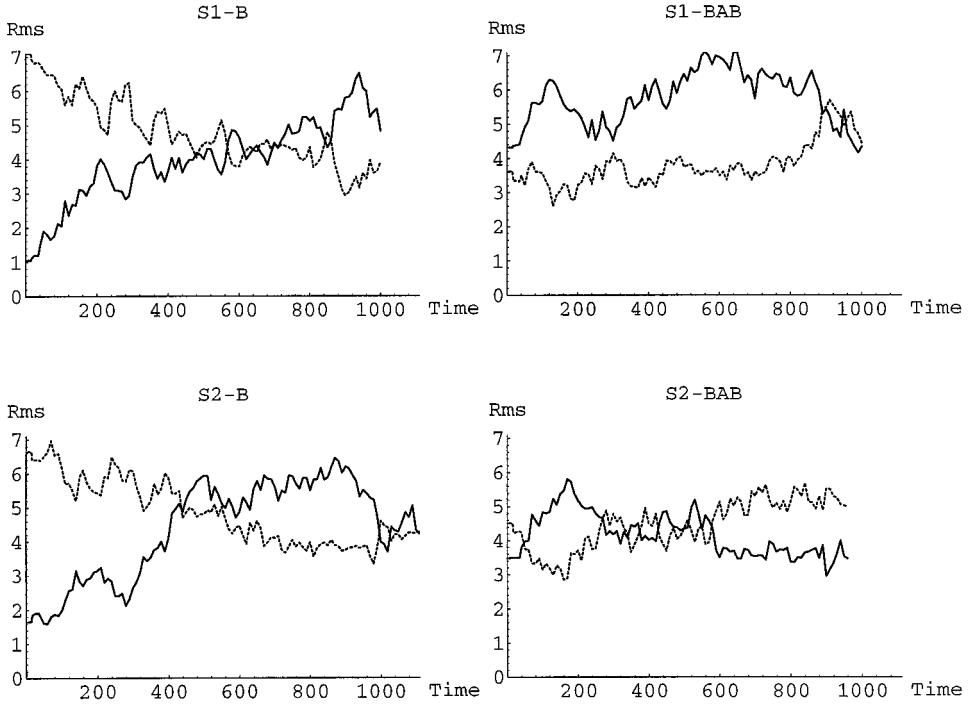

FIGURE 2 RMSD values (in Å and limited to non-hydrogen atoms) along the trajectories of the S1 and S2 oligomers, measured with respect to canonical B-DNA (*solid lines*) and A-DNA forms (*dotted lines*).

S2-B, as for S1-B, there is a relatively uniform drift away from the B form and toward the A form. With both sequences, this leads to conformations roughly midway between the canonical allomorphs after ~500 ps of simulation. After this point, both sequences appear to oscillate slowly within the midway zone. In contrast, starting from mixed B/A conformations, with C3′-endo puckers in the central region of the oligomers, the trajectories are placed immediately between the two canonical forms. S2-BAB again shows slow oscillations between the two canonical forms but, compared to S1-BAB, passes much more time closer to the B canonical form.

This result is substantiated by comparing conformations drawn from the B and BAB trajectories for each sequence (Fig. 3). In the case of the S1 sequence, this leads to a clear correspondence of conformations at the end of the B trajectory with those at the beginning of the BAB trajectory. RMSD values <3 Å are found within a block covering ~400 ps of each trajectory. This notion of continuation from B to BAB is in line with our original assumption that the spontaneous sugar pucker changes seen in the first trajectory would, with time, become more complete in the central region of the oligomer. In the case of S2, the correspondence between the two trajectories is much more extensive and excludes only two relatively limited areas at the beginning and the end of both trajectories.

In order to interpret these trends more deeply, it is necessary to look both at overall changes in backbone geometries and in helical parameters, as well as at local deformations that can influence the fit of short oligomers with the canonical reference conformations.

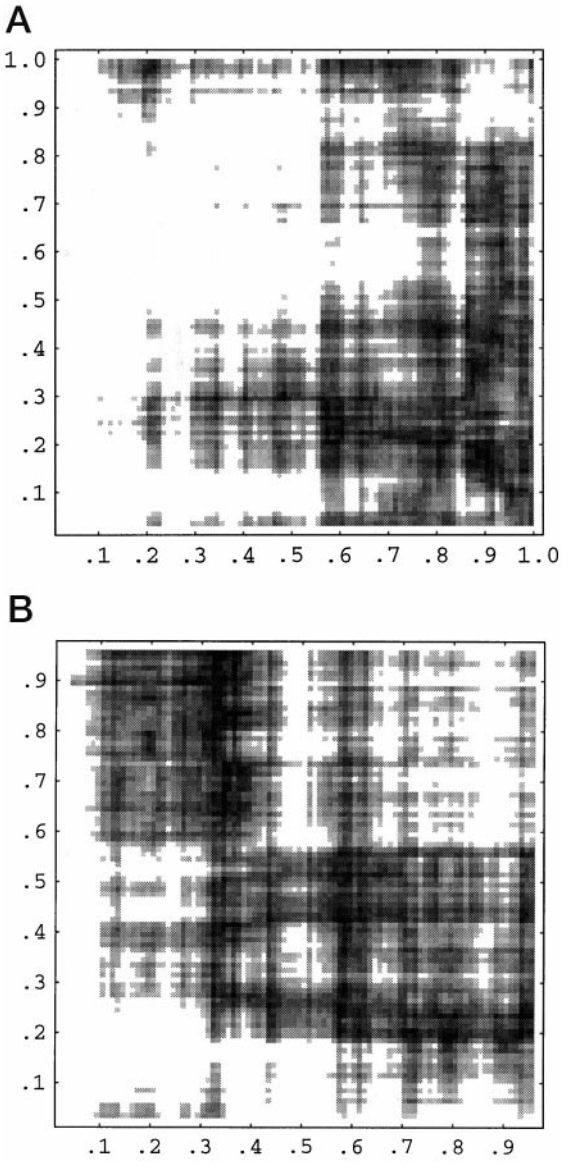

FIGURE 3 RMSD comparison of structures taken from the B (*horizontal axis*) and BAB (*vertical axis*) trajectories at intervals of 10 ps. All shaded areas correspond to RMSD values for all non-hydrogen atoms of better than 3 Å, darker shading indicating better superposition: (*a*) S1 sequence ($RMSD_{min}$ = 1.4 Å), (*b*) S2 sequence ($RMSD_{min}$ = 1.2 Å).

Backbone conformation

During the S2-B trajectory, the sugars make only relatively limited transient excursions into the C3′-endo domain (Fig. 4). This contrasts with the S1-B trajectory where, as we have discussed earlier (Flatters et al., 1997a), several nucleotides, notably in the first strand of the TATA-box region, spend long periods in this domain. This change can easily be seen in a plot of the percentage time each nucleotide spends in the O1′-endo→C3′-endo region (Fig. 5 *a*).

The S2-BAB trajectory also behaves differently from the S1 sequence. In this case, all the thymines in the binding site region repucker from C3′-endo to O1′-endo/C2′endo almost simultaneously after ~150 ps. Beyond this point they only make rare excursions into the C3′-endo domain. A similar repuckering occurs for the adenines, but the change occurs more slowly. It begins between 150 and 200 ps at the 5′- and 3′-ends of the manually repuckered region and then progresses in the 5′→3′ sense for remaining nucleotides. Overall, the transition covers the period of ~600 ps (Fig. 4 *b*). After this conversion, as for the thymines, there are only relatively rare excursions into the C3′-endo domain. The difference with the S1 sequence is again reflected in the percentage of time nucleotides spend in the O1′-endo→C3′-endo region over the latter part of the two trajectories (Fig. 5 *b*). This plot, with limited exceptions at the ends of the binding site region, shows clearly a less A-like character for the S2 sequence.

Concerning the remaining backbone parameters, the only notable deviations from canonical values concern coupled changes in the ϵ (C4′-C3′-O3′-P) and ζ (C3′-O3′-P-O5′) dihedrals from tg^- to g^-t, which move the corresponding phosphate groups from B_I to B_{II} conformations (Privé et al., 1987). Such transitions occur more frequently for the S2 sequence than for S1, as shown in Fig. 6. However, such events typically occur only once or twice for a given dinucleotide step during the nanosecond trajectories. It is thus impossible to conclude on the relative preference of differ-

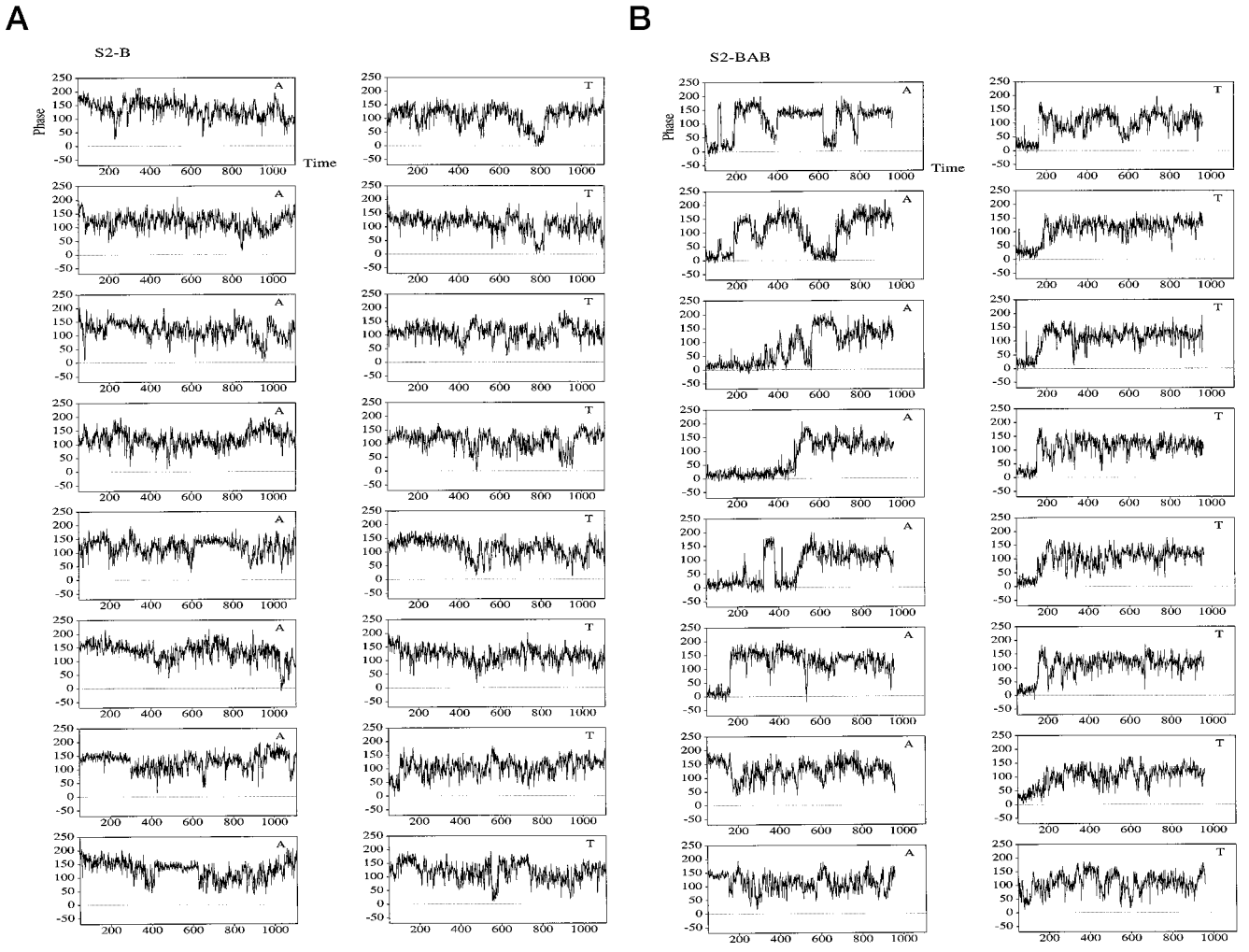

FIGURE 4 Time evolution of the sugar phase angles (°) for the central part of the S2 sequence, during the (*a*) S2-B and (*b*) S2-BAB trajectories. The diagrams are oriented so that the 5′-3′ sense of the first strand points upward.

ent steps for B_{II} conformations. For both the S1 and S2 oligomers there are differences in B_{II} populations between the B and BAB trajectories. For S1, the principal difference involves the terminal CpG step, which stays in a B_{II} conformation for most of the B trajectory. For S2, the B trajectory shows a much more widely spread presence of B_{II} phosphates.

Helical parameters

Overall, the S2 simulations, like those of S1, lie in between the A and B conformations in terms of helical parameters. In both sequences, mean X_{disp} values fall ~ -2 Å and -3 Å. The only clear difference due to the sequence mutations occurs during the latter portion (600–1000 ps) of the BAB trajectories. During this period the S2 sequence moves away from the A-form more rapidly and has a mean X_{disp} of -1.9 Å, compared to a value of -2.7 Å for S1. (Both B trajectories give almost identical results around -2 Å.) Global twist and rise values also lie generally below those of canonical B-DNA, with twist oscillating around 30–34° (with excursions down to 28° during the B trajectory) and rise oscillating around 3.1–3.5 Å. Both of these parameters show small amplitude fast oscillations ($\sim$2° and 0.2 Å, respectively, on a 10-ps time scale), superposed on large amplitude, slow oscillations ($\sim$6° and 0.4 Å, respectively, on a 500-ps time scale). A similar indication of partially A-like structure is given by the average local slide values, which vary between ~ -1 Å and -1.5 Å for both the sequences studied, without showing much sequence dependence. (These values can be compared to 0.1 Å for canonical B-DNA and -1.9 Å for canonical A-DNA.) There is, however, a difference between the two sequences in terms of mean local roll, where the S2-B and S2-BAB trajectories with values of 3.1° and 3.4°, respectively, both have distinctly smaller rolls than the 4.9° measured over the S1-BAB run. (Again, these values should be compared with those of canonical B-DNA 0.9° and canonical A-DNA 11.4°.) Finally, basepair inclination remains small ($\pm$2°) for both sequences studied.

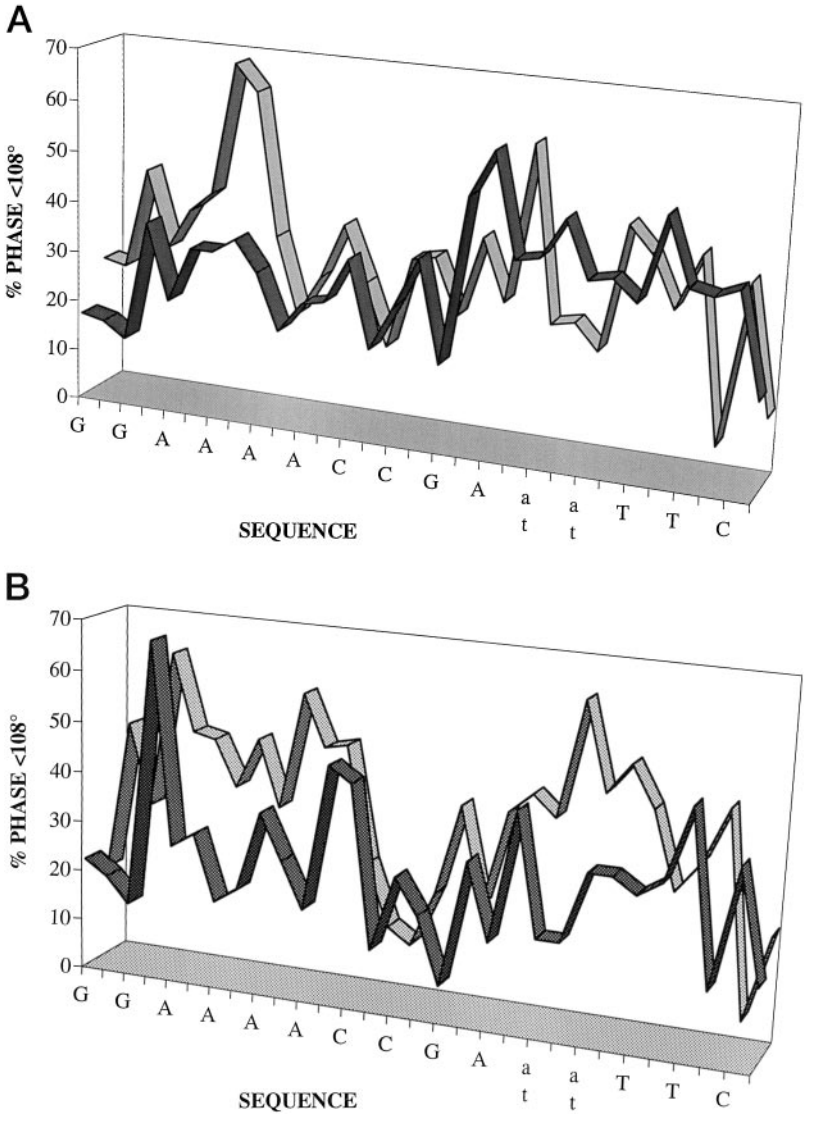

FIGURE 5 Time-averaged percentage of sugar puckers below 108° (i.e., O1′-endo→C3′-endo) for (*a*) the B trajectories and (*b*) the latter part of the BAB trajectories (600–1000 ps). The results for the S1 sequence are in pale shading and those for the S2 sequence in dark shading. The two stands are placed one after another along the abscissa, respectively, in the 5′-3′ and 3′-5′ sense.

It is interesting to compare these values to a CURVES analysis of the DNA within the TBP-DNA co-crystal from the Sigler group (Kim et al., 1993b). This structure, which has a binding site with the S1 sequence, shows an average X_{disp} of −2.1 Å, a twist of 30°, and a rise of 3.4 Å. These values are all similar to those seen with both the S1 and S2 sequences. However, protein binding also causes a major change in basepair inclination, which adopts a mean value of 21° (with a maximum of 31°) and average rolls which become highly positive (mean 16°, maximum 40°). These effects are clearly absent from the isolated DNA simulations performed here.

If we look in more detail at basepair rolls, sequence-dependent effects can be seen. Averaging over the S2 trajectories shows strong positive rolls for YpR steps (CpG, TpA) and small rolls for RpR (ApA) and RpY (GpC, ApT) steps (Fig. 7). This was also the case for the S1 sequence, with the exception of the small roll at the second TpA step during the S1-B trajectory which, however, dramatically increases during the BAB run. There are no such differences between the S2-B and S2-BAB trajectories. It should be noted that, for both the S1 and S2 sequences, ApA steps generally maintain small positive rolls of ~2–3°. There also seems to be a general trend toward smaller rolls at the 3′-end of the A-tracts in both sequences.

Finally, in line with crystallographic (Goodsell and Dickerson, 1994) and earlier modeling results (Zhurkin et al., 1979; Olson and Zhurkin, 1996) local basepair tilts are much smaller than rolls, lying here within a range of ±2° and showing no clear sequence-related behavior.

Helix bending

Fig. 8 summarizes the overall magnitude and direction of bending for the S1 and S2 oligomers, using the analysis technique described in the Methods section. As we have discussed earlier (Flatters et al., 1997a), the S1-BAB trajectory undergoes very strong bending three times during the trajectory. The first two of these periods correspond to

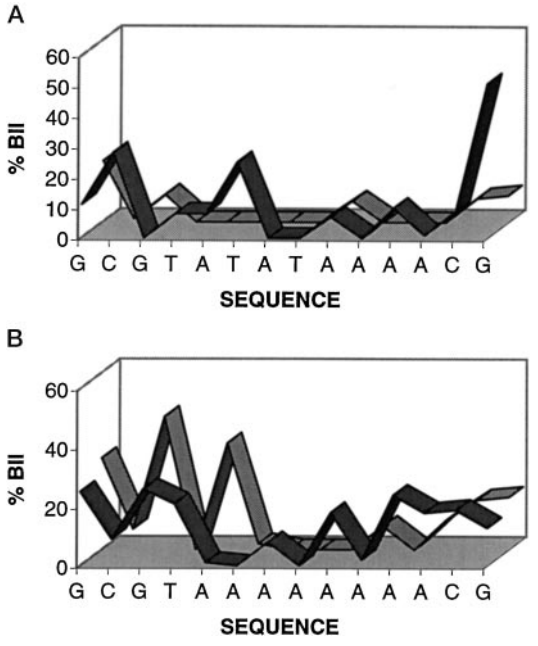

FIGURE 6 Percentages of B_{II} conformations occurring at each basepair step for the (*a*) S1 sequence, and (*b*) S2 sequence. In both diagrams the results for the B trajectory are shown in dark shading and those for the BAB trajectory in pale shading.

bending mainly in the central part of the oligomer and show a bending direction that is remarkably tightly constrained around the major groove pseudodyad vector of the TATA-box. This is almost identical to the direction of bending with the TBP-DNA co-crystals (Kim et al., 1993b). When the central section is weakly bent or when bending occurs mainly at the 5′-end of the oligomer (toward the end of the trajectory) the bending direction is less constrained. In contrast, the S2-BAB trajectory shows both weaker and more uniform magnitudes of bending and, although bending again occurs mainly toward the major groove in the center of the oligomer, the fluctuation in bending direction is clearly greater (very similar results are obtained for the S2-B trajectory).

If we limit our analysis to the central part of each sequence (repeating the CURVES calculations for these segments alone), equivalent results are found. As shown in Fig. 9, the TATATAA tract in S1 is typically bent by 20–30°, while the TAAAAAAA tract in S2 is generally bent by around 10°. It should be noted that the central tracts also show a difference in bending anisotropy, with the more flexible S1 tract bending much more sharply along the central major groove dyad.

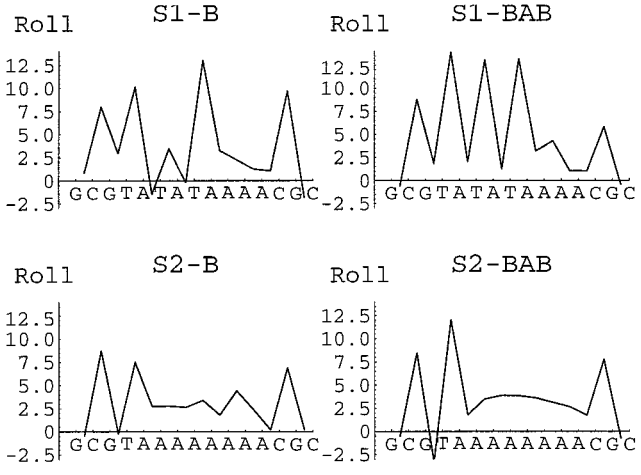

FIGURE 7 Time-averaged local roll angles (°) for each trajectory performed.

Groove geometry

Using CURVES, groove width can be measured continuously along the oligomers (Stofer and Lavery, 1993) and averaged over the dynamics trajectories. The results in Fig. 10 show that, in line with numerous experimental results (see, for example, Drew and Dickerson, 1981; Nadeau and Crothers, 1989), the A-tracts have narrower minor grooves than the surrounding sequences. In addition, the short A-tract in S1 shows a clear decrease in groove width in the 5′-3′ direction, in correlation with the hydroxy radical cleavage experiments from the Tullius group (Burkhoff and Tullius, 1987) and with NMR studies of hydrogen exchange kinetics (Leroy et al., 1988, Katahira et al., 1988). In contrast, with the much longer A-tract in S2, this effect is not visible.

RMSD triangulation

Finally, to return to an overall view of the conformations obtained with each sequence, we have extended our RMSD calculations by including a third, significant reference structure, namely the crystallographic conformation of DNA contained within the TBP-DNA co-crystal (Kim et al., 1993b). (Note that the comparison of S2 with the crystal conformation involved editing the crystal structure to change its base sequence, without changing the position or orientation of the bases concerned.) The results are shown schematically in Fig. 11 for conformations drawn from the BAB trajectories of the S1 and S2 sequences at ~550 ps, when both sequences show relatively strong bending toward the major groove. It is clearly seen that the S2 oligomer lies closer to B-DNA than does S1, and further from both A-DNA and the conformation complexed by TBP (termed "TA-DNA," following the notation proposed by Guzikevich-Guerstein and Shakked, 1996).

DISCUSSION

On the basis of the data presented above, we can conclude that both the S1 and S2 oligomers adopt conformations in between those of A- and B-DNA and share some features of DNA complexed by TBP. This is true both for the negative average X_{disp} values and the low average twist, but not for

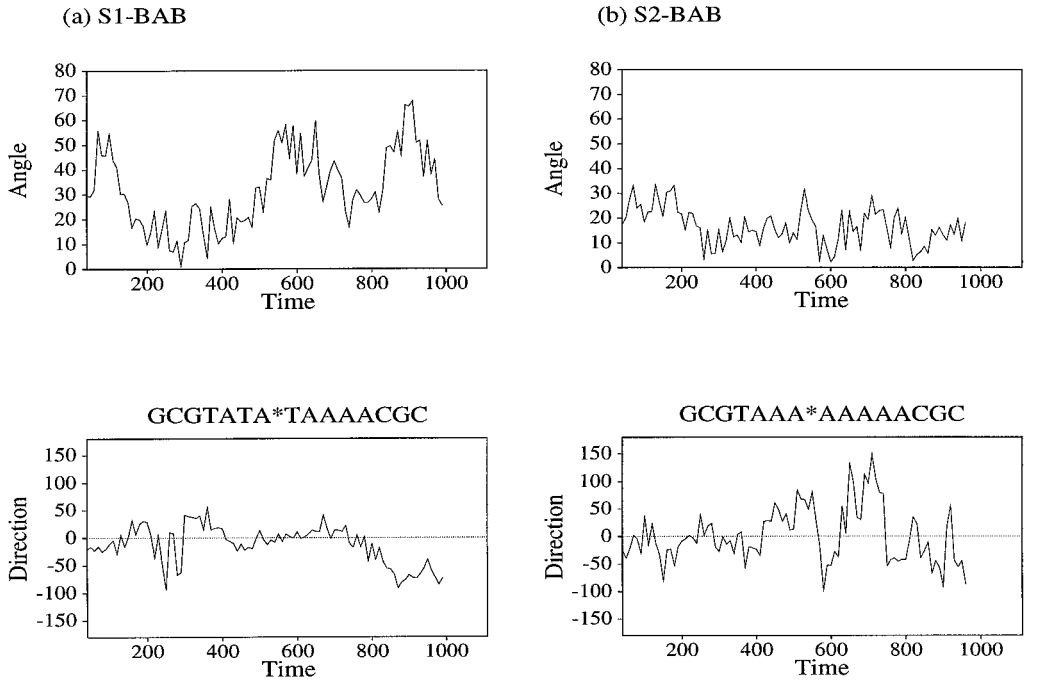

FIGURE 8 Magnitude and direction of overall bending during the (*a*) S1-BAB and (*b*) S2-BAB trajectories. Values are in degrees and direction is measured with respect to the pseudodyad vector at the center of the TBP binding site (i.e., A7pT8 in S1 and A7pA8 in S2).

the basepair inclinations or the rolls, which are much smaller in the isolated DNA's than in the protein complex.

However, the two mutations in sequence involved in passing from S1 to S2 produce significant differences in the dynamic and time-averaged behavior of these oligomers. Overall, S1 drifts closer to the A-form than does S2, and also stays closer to it after the introduction of C3′-endo sugars at the protein binding site. This is particularly clear in the coupled transition of the sugars away from C3′-endo puckers during the first half of the S2-BAB trajectory. The

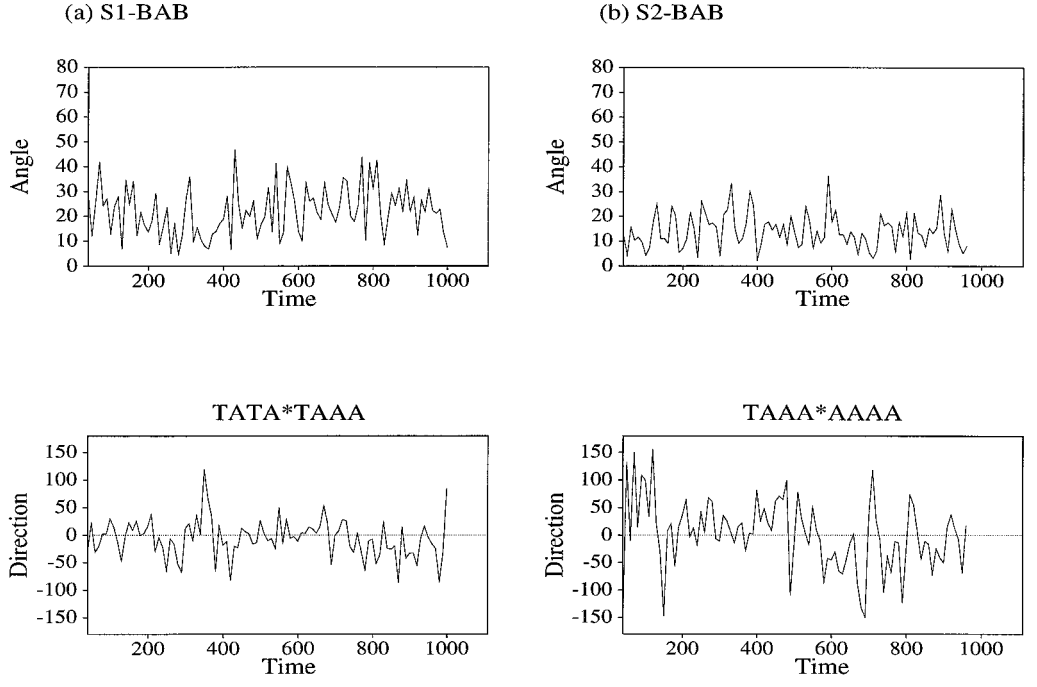

FIGURE 9 Magnitude and direction of bending for the central 8-bp tracts of the (*a*) S1-BAB and (*b*) S2-BAB trajectories. Values are in degrees and direction is measured with respect to the pseudodyad vector at the center of the TBP binding site (i.e., A7pT8 in S1 and A7pA8 in S2).

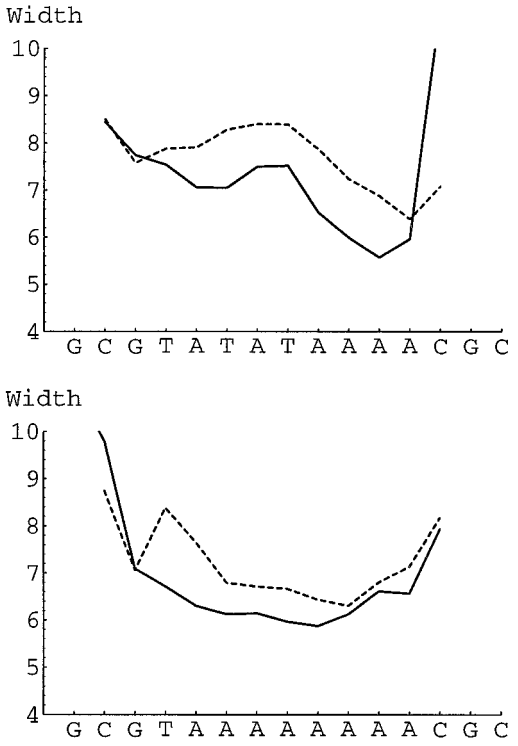

FIGURE 10 Time-averaged minor groove widths (Å) measured with CURVES along the S1 (above) and S2 oligomers (below). Solid lines refer to the B trajectories and dotted lines to the BAB trajectories.

same difference is also reflected in the time-averaged values of the sugar puckers, which are further from B-like puckers for the S1 sequence, in the first strand of the TATA-box during the B trajectories and in both strands of the protein binding site during the BAB trajectories. Additional confirmation of these differences comes from the average helicoidal parameters, which show both less negative X_{disp} and smaller positive rolls for the S2 sequence.

Overall bending occurs with both sequences in the direction of the central major groove, as seen in the TBP-DNA complex. However, it is significant that the S2 oligomer bends much less and also less anisotropically than S1. Bending is mainly linked to large positive rolls at YpR steps (while RpY steps adopt rolls close to zero). This is in line with the early propositions of Zhurkin and collaborators (Zhurkin et al., 1979; Ulyanov and Zhurkin, 1984), with crystallographic (Gorin et al., 1995) and NMR (Ulyanov and James, 1995) data, and with our recent modeling studies of superhelical polymeric DNA using JUMNA (Sanghani et al., 1996). The general properties of the tracts of AT basepairs in these simulations thus appear to be more in line with non-A tract models (Goodsell and Dickerson, 1994; Olson and Zhurkin, 1996) than with dinucleotide ApA wedge models. Lastly, the shorter A-tract in S1 and the much longer A-tract in S2 share narrow minor groove widths and small positive basepair rolls. Groove narrowing in the 5′→3′ sense is visible with the S1 sequence, but absent with S2, where the eight AT pairs presumably come closer to the uniform polymeric conformation of poly(dA)·poly(dT). This hypothesis would also explain the greater bending rigidity of the S2 sequence mentioned above.

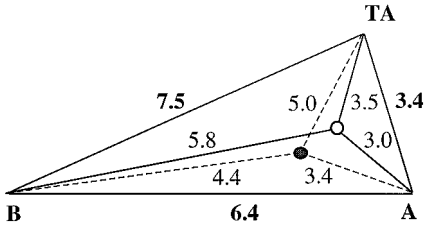

FIGURE 11 Schematic diagram showing the RMSD values (Å) between representative conformations of the S1 (*open circle*) and S2 (*filled circle*) sequences during their BAB trajectories with respect to canonical B-DNA, A-DNA, and "TA-DNA," the conformation of DNA complexed with TBP (Kim et al., 1993b). The comparison is made for using conformations of each oligomer occurring at ~550 ps into the simulation as is limited to the central 12 bp of each sequence.

Concerning the central aim of this study, it has been demonstrated that two point mutations in a TBP binding site lead to easily measurable differences in dynamic and time-averaged properties. In agreement with experimental measurements, these differences suggest that the S2 sequence is less well adapted to TBP binding by being globally less close to an A-like conformation and in exhibiting weaker and less anisotropic bending. These findings suggest that molecular dynamics simulations using explicit water and counterion representations and avoiding electrostatic truncation are already capable of answering questions concerning fine sequence-dependent behavior. We conclude from the present study that the intrinsic properties of TBP binding sites apparently play a significant role in determining the quality of protein binding, subsequent structural deformations, and the corresponding functional changes.

To end with a word of caution, it is clear that nanosecond scale simulations, although informative, cannot give a complete view of the dynamic behavior of DNA oligomers, since many conformational fluctuations occur on slower time scales. It is also clear that the force fields and simulation protocols chosen have important effects. Thus, recent CHARMM simulations of seven DNA dodecamers containing various TBP binding sequences (Pastor et al., 1997) show some global features in common with our simulations (unwinding, negative slide, positive basepair roll, and some evidence of bending toward the major groove), but do not agree on certain sequence dependent effects, showing, for example, strong positive rolls for RpY steps unlike the positive values seen for YpR steps in our simulations. However, differences in the analysis and the absence of any sequence close to S2 make precise comparisons difficult.

The authors would like to thank IDRIS (Institut du Développement et des Ressources en Informatique Scientifique) for the attribution of Cray C98 resources and Matthew Young for many helpful discussions.

REFERENCES

Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, A. DiNiola, and J. R. Hask. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.

Burkhoff, A. M., and T. D. Tullius. 1987. The unusual conformation adopted by the adenine tract of kinetoplast DNA. *Cell*. 48:935–943.

Cheatham, T. E. III, and P. A. Kollman. 1996. Observation of the A-DNA to B-DNA transition during unrestrained molecular dynamics in aqueous solution. *J. Mol. Biol.* 259:434–444.

Cheatham, T. E. III, J. L. Miller, T. Fox, T. A. Darden, and P. A. Kollman. 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.

Cluzel, P., A. Lebrun, C. Heller, R. Lavery, J.-L. Viovy, D. Chatenay, and F. Caron. 1996. DNA: an extensible molecule. *Science*. 271:792–794.

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.

Darden, T. A., D. M. York, and L. G. Pedersen. 1993. Particle mesh Ewald: an N · log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

Drew, H. R., and R. E. Dickerson. 1981. Structure of a B-DNA dodecamer. II. Influence of base sequence on helix structure. *J. Mol. Biol.* 149: 761–786.

Duan, Y., P. Wilkosz, M. Crowley, and J. M. Rosenberg. 1997. Molecular dynamics simulation study of DNA dodecamer d(CGCGAATTCGCG) in solution: conformation and hydration. *J. Mol. Biol.* 272:553–572.

Flatters, D., M. Young, D. L. Beveridge, and R. Lavery. 1997a. Conformational properties of the TATA-box binding sequence of DNA. *J. Biomol. Struct. & Dyn.* 14:757–765.

Flatters, D., K. Zakrzewska, and R. Lavery. 1997b. Internal coordinate modeling of DNA: force field comparisons. *J. Comp. Chem.* 18: 1043–1055.

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.

Guzikevich-Guerstein, G., and Z. Shakked. 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.

Hingerty, B., R. H. Richie, T. L. Ferrel, and T. E. Turner. 1985. Dielectric effects in biopolymers: the theory of ionic saturation revisited. *Biopolymers*. 24:427–439.

Jayaram, B., S. Swaminathan, and D. L. Beveridge. 1990. Monte Carlo simulation studies on the structure of the counterion atmosphere of B-DNA. Variations on the primitive dielectric model. *Macromolecules*. 23:3156–3165.

Jorgensen, W. L. 1981. Transferable intermolecular potential functions for water, alcohols and ethers. Application to liquid water. *J. Am. Chem. Soc.* 103:335–340.

Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–936.

Juo, Z. S., T. K. Chiu, P. M. Leiberman, I. Baikalov, A. J. Berk, and R. E. Dickerson. 1996. How proteins recognize the TATA-box. *J. Mol. Biol*. 261:239–254.

Katahira, M., H. Sugeta, Y. Kyogoku, S. Fujii, R. Fujisawa, and K. Tomita. 1988. One and two dimensional NMR studies on the conformation of DNA containing the oligo(dA) · oligo(dT) tract. *Nucleic Acids Res.* 16:8619–8632.

Kim, J. L., and S. K. Burley. 1994. 1.9 Å resolution refined structure of TBP recognizing the minor groove of TATAAAG. *Nature Struct. Biol*. 1:638–653.

Kim, Y., J. H. Gieger, S. Hahn, and P. B. Sigler. 1993b. Crystal structure of a yeast TBP/TATA-box complex. *Nature*. 365:512–520.

Kim, J. L., D. B. Nikolov, and S. K. Burley. 1993a. Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature*. 365: 520–527.

Lavery, R., and H. Sklenar. 1989. Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. & Dyn.* 6:655–667.

Lavery, R., K. Zakrzewska, and H. Sklenar. 1995. JUMNA (junction minimization of nucleic acids). *Comp. Phys. Commun.* 91:135–158.

Lebrun, A., and R. Lavery. 1996. Modeling extreme stretching of DNA. *Nucleic Acids Res*. 24:2260–2270.

Lebrun, A., Z. Shakked, and R. Lavery. 1997. Local DNA stretching mimics the distortion caused by the TATA-box binding protein. *Proc. Natl. Acad. Sci (USA)*. 94:2993–2998.

Leroy, J.-L., E. Charetier, M. Kochoyan, and M. Gueron. 1988. Evidence from base-pair kinetics for two types of adenine tract structures in solution: their relation to DNA curvature. *Biochemistry*. 27:8894–8898.

Nadeau, J. G., and D. M. Crothers. 1989. Structural basis for DNA bending. *Proc. Natl. Acad. Sci. (USA)*. 86:2622–2626.

Olson, W. K., and V. B. Zhurkin. 1996. Twenty years of DNA bending. *In* Biological structure and dynamics. R. H. Sarma and M. H. Sarma, editors. Adenine Press, New York. 341–370.

Pastor, N., L. Pardo, and H. Weinstein. 1997. Does TATA matter? A structural exploration of the selectivity determinants in its complexes with TATA box-binding protein. *Biophys. J.* 73:640–652.

Pearlman, D. A., D. A. Case, G. W. Caldwell, W. S. Ross, T. E. Cheatham III, D. M. Ferguson, G. L. Seibel, U. C. Singh, P. Weiner, and P. A. Kollman, editors. 1995. AMBER 4.1. UCSF, San Francisco.

Privé, G. G., U. Heinemann, S. Chandrasekharan, L. S. Kan, M. L. Kopka, and R. E. Dickerson. 1987. Helix geometry, hydration and G. A mismatch in a B-DNA decamer. *Science*. 238:498–504.

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. Comp. Phys.* 23: 327–341.

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

Smith, S. B., Y. Cui, and C. Bustamante. 1996. Overstretching DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science*. 271:795–799.

Starr, D. B., B. C. Hoopes, and D. K. Hawley. 1995. DNA bending in an important component of site-specific recognition by the TATA binding protein. *J. Mol. Biol.* 250:434–446.

Stofer, E., and R. Lavery. 1993. Measuring the geometry of DNA grooves. *Biopolymers*. 34:337–346.

Ulyanov, N. B., and T. L. James. 1995. Statistical analysis of DNA duplex structural features. *Methods Enzymol.* 261:90–120.

Ulyanov, N. B., and V. B. Zhurkin. 1984. Sequence-dependent anisotropic flexibility of B-DNA. A conformational study. *J. Biomol. Struct. & Dyn.* 2:361–385.

van Gunsteren, W. F., and H. J. C. Berendsen. 1977. Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.* 34: 1311–1327.

Weiner, S. J., P. A. Kollman, D. T. Nguyen, and D. A. Case. 1986. An all atom force field for simulation of proteins and nucleic acids. *J. Comp. Chem.* 7:230–252.

Wobbe, C. R., and K. Struhl. 1990. Yeast and human TATA-binding proteins have nearly identical DNA sequence requirements for transcription in vitro. *Mol. Cell Biol.* 10:3859–3867.

Young, M. A., and D. L. Beveridge. 1997. A 5-nanosecond molecular dynamics trajectory for B-DNA: analysis of structure, motions, and solvation. *Biophys. J.* 73:2313–2336.

Young, M. A., B. Jayaram, and D. L. Beveridge. 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.

Zhurkin, V. B., Y. P. Lysov, and V. I. Ivanov. 1979. Anisotropic flexibility of DNA and the nucleosomal structure. *Nucleic Acids Res.* 6:1081–1096.