Molecular Dynamics Simulations Find That 3′ Phosphoramidate Modified DNA Duplexes Undergo a B to A Transition and Normal DNA Duplexes an A to B Transition

Piotr Cieplak,[†] T. E. Cheatham, III,[‡] and Peter A. Kollman*[,‡]

Contribution from the Department of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland, and Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94143

Received November 12, 1996[⊗]

Abstract: We present unrestrained molecular dynamics studies on a deoxyribose dodecamer duplex d(CGCGAATTCGCG)$_2$ and its phosphoramidate (replacing O3′ by NH) analog using particle mesh Ewald electrostatics[1] and the Cornell *et al*. force field.[2] The simulations were carried out beginning in both canonical A and B forms, which differ in RMS position by ∼6.5 Å. These simulations, which were carried out for ∼1−1.5 ns, are consistent with experiment in that the unmodified dodecamer, whether started in A or B DNA, converges to a B structure which is closer to the observed X-ray structure (9bna) than canonical B. On the other hand, the phosphoramidate modified duplex, whether started in A or B forms, converges to an A form. In this duplex, however, the structure from the simulation begun in the B form has a dependence on the initial location of the N−H hydrogen. In one of our simulations, we find base pair opening *and* closing at the end of the duplex. Specifically, in the simulation of d(CGCGAATTCGCG)$_2$ begun in the A form, base pair opening of the terminal G−C base pair occurs at 400 ps, and then the bases remain unpaired for 700 ps before reclosing.

Introduction

Synthetic oligonucleotides and their analogs are of considerable interest, because they may lead to highly specific therapeutic agents and could be potentially used as diagnostic tools.[3] The normal phosphodiester-linked oligonucleotides are usually degraded relatively easily by cellular nucleases and, because of that, cannot be used as in vivo therapeutics. Thus, it is of interest to develop modified antisense oligonucleotides that are degradation resistant. An example of such analogs are oligonucleotides with a phosphorothioate modified backbone, which have been shown to be among the most promising antisense agents. Presently several oligonucleotide phosphorothioates are in clinical trials against variety of targets.[4] Unfortunately, duplexes and triplexes formed by phosphorothioate strands are less stable than the corresponding phosphodiesters.

Recently, Gryaznov *et al*.[5−11] have initiated studies on another, very promising, modification of the oligonucleotide backbone, where the O3′−P bonds are replaced by N3′−P linkages. Such phosphoramidate analogs form very stable duplexes with single-stranded DNA, RNA, and with themselves. The phosphoramidate analogs are able also to form stable triplexes with double-stranded DNA and RNA under near-physiological conditions.[8,9] They have been tested as antisense agents,[10] have proved much more digestion resistant than phosphorothioates, and have good water solubility.[6,8] In terms of melting temperatures, the duplex stability is enhanced by 2.2−2.6 °C per modified linkage compared with normal phosphodiesters.[8]

Phosphoramidate oligonucleotide analogs demonstrate very interesting conformational properties. Based on CD and NMR spectroscopy, it has been shown that complexes formed by phosphoramidates correspond to an A form of DNA,[8,11] whereas standard deoxyoligonucleotide duplexes adopt the B-form.[12] The observed sugar deoxyribose puckering phases are in agreement with the data provided by Thibaudeau *et al*.[13] According to them, the N/S sugar puckering equilibrium is closely connected to the electronic nature of the 3′ substituent. For the 2′-deoxyribonucleosides a 3′NH$_2$ has almost the same N/S ratio as 3′H, whereas 3′OH has a much smaller ratio. Also, based on the experimental findings of Gryaznov *et al*.,[8] the H atom on the N3′−H group is very important. Substituting that hydrogen by a methyl group leads to complete suppression of the binding to single stranded DNA at room temperature and thus a dramatic reduction of the thermal stability of duplexes formed by the modified 10-mer with single stranded DNA.

[†] University of Warsaw.

[‡] University of California, San Francisco.

[⊗] Abstract published in *Advance ACS Abstracts,* June 15, 1997.

(1) Darden, T. A.; York, D. M.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(2) Cornell, W.; Cieplak, P.; Bayly, C. I.; Gould, I.; Merz, Jr., K. M.; Ferguson, D.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(3) Crooke, S. T.; Lebleu, B. *Antisense research and Applications*; CRC Press: Boca Raton, FL, 1993.

(4) Bayever, E.; Iversen, P. L.; Bishop, M. R.; Sharp, J. G.; Teasary, H. K.; Arneson, M. A.; Pirruccello, S. J.; Auddon, R. W.; Kessinger, A.; Zon, G.; Armitage, J. O. *Antisense Res. Dev.* **1993**, *3*, 382−390.

(5) Gryaznov, S. M.; Sokolova, N. I. *Tetrahedron Lett.* **1990**, *31*, 3205−3208.

(6) Gryaznov, S.; Chen, J.-K. *J. Chem. Soc.* **1994**, *116*, 3143−3144.

(7) Chen, J.-K.; Schultz, R. G.; Lloyd, D. H.; Gryaznov, S. M. *Nucleic Acid Res.* **1995**, *23*, 2661−2668.

(8) Gryaznov, S.; Lloyd, D. H.; Chen, J.-K.; Schultz, R.; DeDionisio, L. A.; Ratmeyer, L.; Wilson, W. D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5798−5802.

(9) Escude, C.; Giovannangeli, C.; Sun, J. S.; Lloyd, D. H.; Chen, J.-K.; Gryaznov, S. M.; Garestier, T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 4365−4369.

(10) Gryaznov, S.; Skorski, T.; Cucco, C.; Nieborowska-Skorska, M.; Chiu, C. Y.; Lloyd, D.; Chen, J.-K.; Koziolkiewicz, M.; Calabretta, B. *Nucleic Acids Res.* **1996**, *24*, 1508−1514.

(11) Ding, D.; Gryaznov, S. M.; Lloyd, D. H.; Chandrasekaran, S.; Yao, S.; Ratmeyer, L.; Pan, Y. P.; Wilson, W. D. *Nucleic Acid Res.* **1996**, *24*, 354−360.

(12) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York 1984.

(13) Thibaudeau, C.; Plavec, J.; Garg, N.; Papchikin, A.; Chattopadhyaya, J. *J. Am. Chem. Soc.* **1994**, *116*, 4038−4043.

S0002-7863(96)03909-1 CCC: $14.00

One of the experimental studies has been devoted to the phosphoramidate modified sequence of d(CGCGAATTCGCG)$_2$.[11] That sequence has been chosen because its deoxyribose form has been widely investigated in detail by X-ray crystallography and NMR methods.[14−16] It has been also the subject of many theoretical analysis involving molecular mechanics and molecular dynamics simulations.[17,18]

In this paper we will describe the results of our unconstrained molecular dynamics simulation using the AMBER 4.1[19] molecular modeling package and the Cornell *et al*. force field[2,20] to investigate the structural preference of d(CGCGAATTCGCG)$_2$ and its phosphoramidate derivative. In all our simulations we have applied the Ewald[21] method for an accurate treatment of the long range electrostatic interactions, which is especially important for highly charged systems such as nucleic acids. The particle mesh Ewald (PME) method[1,22] has been recently implemented in the AMBER programs,[19] and the method has been proved to be very efficient and necessary to simulate proper behavior of normal oligonucleotides in crystals and aqueous solution.[23−27]

By carrying out simulations on both duplexes starting in both A and B structural forms we have been able to show the preference of the modified d(CGCGAATTCGCG)$_2$ containing the N3′−P5′ phosphoramidate linkages to be in the A form and the same sequence of the deoxyribose dodecamer to be more stable in the B form. The results of our simulations have shown, consistent with experiment, that obtaining an A → B transition for normal DNA in water solution using the Cornell *et al*. force field is not sequence dependent, since that kind of behavior was successfully simulated for another sequence and shorter oligonucleotide, d(CCAACGTTGG)$_2$, by Cheatham *et al*.[26] as well. We have also been able to characterize an example of a CG base pair opening at the end of the normal duplex.

Methods

Five different molecular dynamics simulations have been performed. Two "control" simulations have been done for standard dodecamers d(CGCGAATTCGCG)$_2$, one starting from the canonical B form (1000 ps long) and the other starting from the canonical A form[12] (1500 ps). Three other simulations have been performed on the modified dodecamer containing the N3′−P5′ linkages between all nucleosides. Since the accurate structure of such modified dodecamer is not yet available from crystallography or NMR measurements, we created our initial structures from standard canonical A and B forms by replacing the O3′ atom by an N−H group. In each case the N−H group can adopt different orientation which we denote as orientations: "NH in" (characterized by the value of the C4′−C3′−N3′−H angle to be in the range 40−80°) and "NH out" (C4′−C3′−N3′−H angle in the range of −40 to −90°). We will show that the exposure of the N−H to the water solvent plays critical role on the observed conformational behavior of dodecamers during molecular dynamic simulations. As can be seen in Figure 1 the NH in and NH out orientations are substantially different in the case of B-form, with respect to the exposure of the hydrogen atom to the water solution. Thus, we decided to perform two different molecular dynamics simulations, one starting with all hydrogens in the NH in orientations (1000 ps) and another starting with all hydrogens in the NH out orientations (1000 ps). In the case of the A form structure in either N−H location, the N−H hydrogens are fairly well exposed to water. In the course of preparation of the initial structure for the molecular dynamics simulations, beginning with the N−H out, after the full minimization of the A form in water solution, 10 (about half of the total) N−H bonds were positioned in the NH in orientations. For this structure, we performed 1500 ps of a molecular dynamics simulations.

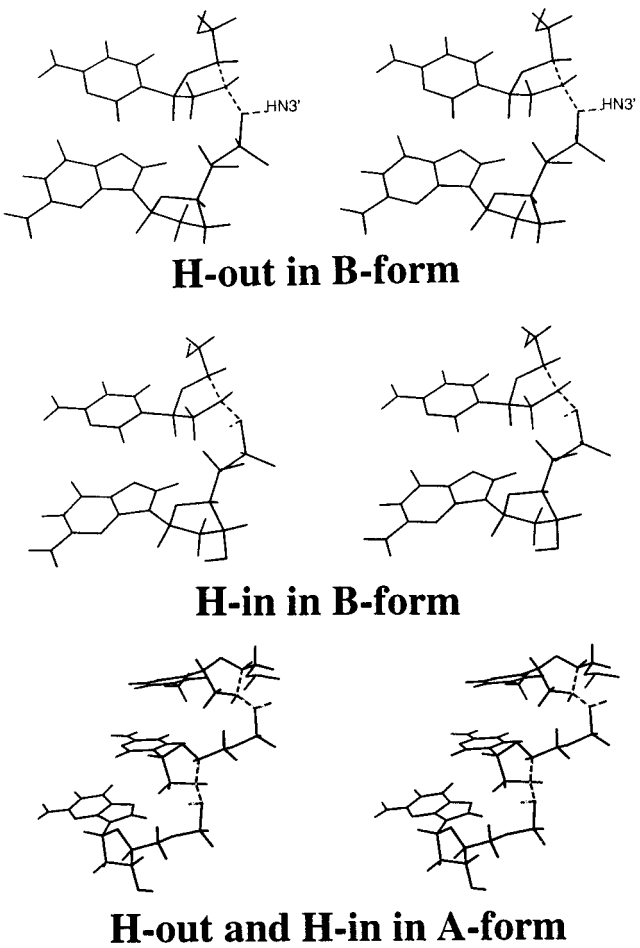

Figure 1. Stereoview illustrated hydrogen position of H3′ of N3′−H3′ in B (top two figures) and A (bottom figure). In A, the top nucleoside has N−H out, the second has N−H in.

In all cases the initial canonical structures of A and B-DNA were created using NUCGEN module of AMBER 4.1.[19] Counterions were placed next to each phosphate group to keep the whole simulated system neutral, and then the whole DNA was immersed in the rectangular box of TIP3P[28] water molecules. The water box extended 11 Å away from any solute

(14) Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratin, A. V.; Kopka, M. L. *Science* **1982**, *216*, 475−485.
(15) Westhof, E. *J. Biomol. Struct. Dyn*. **1987**, *5*, 581.
(16) Chou, S. H.; Flynn, D.; Reid, B. R. *Biochemistry* **1989**, *28*, 2422−2435.
(17) Beveridge, D. L.; Ravishanker, G. *Current Opinion Str. Biol*. **1994**, *4*, 246−255, and references therein.
(18) Miaskiewicz, K.; Osman, R.; Weinstein, H. *J. Am. Chem. Soc*. **1993**, *115*, 1526−1537.
(19) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; DeBolt, S. E.; Ferguson, D. M.; Seibel, G. L.; Kollman, P. A. Comput. Phys. Commun. **1995**, *91*, 1−41.
(20) Cieplak, P.; Cornell, W.; Bayly, C.; Kollman, P. A. *J. Comput. Chem*. **1995**, *16*, 1357−1377.
(21) Ewald, P. *Ann. Phys.* (Leipzig) **1921**, *64*, 253−264.
(22) Essman, U.; Perera, L.; Berkowitz, M. L.; Darden, T. A.; Lee, H.; Pedersen, L. G. *J. Chem. Phys*. **1996**, *103*, 8577−8593.
(23) Darden, T. A.; York, D.; Pedersen, L. G. *J. Chem. Phys*. **1993**, *98*, 10089−10092.
(24) York, D. M.; Yang, W. T.; Lee, H.; Darden, T.; Pedersen, L. G. *J. Am. Chem. Soc*. **1995**, *117*, 5001−5002.
(25) Cheatham III, T. E.; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. *J. Am. Chem. Soc*. **1995**, *117*, 4193−4194.
(26) Cheatham III, T. E.; Kollman, P. A. *J. Mol. Biol*. **1996**, *259*, 434−444.
(27) Duan, Y.; Wilkosz, P. A.; Crowley, M.; Rosenberg, J. M. The Ewald Summation Improves the Fidelity of Molecular Dynamics Simulations of the DNA Dodecame (d(CGCGAATTCGCG). Submitted to *J. Mol. Biol.*

atom. This yielded ~3800 water molecules used for solvation. The simulations were run with SANDER module of AMBER using SHAKE[29] on all hydrogen atoms and a 2 fs time step. The 9 Å cutoff has been applied to the Lennard-Jones interactions, and the simulations have been performed at the temperature of 300 K and 1 atm using the Berendsen algorithm for temperature coupling ($\tau_T = 0.2$ ps) and pressure coupling ($\tau_P = 0.2$ ps).[30] The PME charge grid spacing was approximately 1.0 Å, and the charge grid was interpolated using a cubic B-spline with the direct sum tolerance of 0.000 01 at the 9 Å direct space cutoff.

Before beginning the molecular dynamics simulations the DNA, counterions and water were minimized using the following procedure. First all hydrogen atoms were minimized, next all water molecules, and then water with counterions. After that the whole system was minimized using 5000 steps of the conjugate gradient method.

After each picosecond of molecular dynamics, coordinates were written out for further analysis. A comparison of different trajectories was also based on the calculations of the RMS coordinate deviation (Å) between average structures obtained from the last 500 ps of the MD trajectory, based on the all DNA atoms frames taken at 1 ps intervals. The average structures were obtained using the CARNAL module of AMBER 4.1.[19] Before calculating rms deviations, the reference coordinates were minimized using positional restraints of 5 kcal/mol Å^2 for all DNA heavy atoms.

Additional molecular mechanical parameters were developed based on the *ab initio* calculations for small molecule systems using the Gaussian94[31] program and its 6-31G* basis set. The resultant parameters are presented in Tables 1 and 2.

Charges for the C3′ N−H substituted deoxyribose were developed using our recent RESP approach[20,32] applied to the 1′,2′-dideoxy-ribo-3′-methylophosphoramidate molecule. Dihedral angle terms for CT−NS−P−OS and NS−P−OS−CT phosphoramidate linkage atoms were developed based on geometry optimization and relative torsional energies of the $CH_3NHPO_2^-OCH_3$ molecule. The parameters resulting from those calculations are presented in Tables 1 and 2. Thibadeau[11] had noted that a 3′NH_2 had the same sugar pucker ratio as 3′H and very different from 3′OH. In the Cornell *et al*. force field, there is a 2-fold (V_2) OS−CT−CT−OS torsion[18] to represent the gauche tendency of such torsions. Given the results in ref 11, no such torsion was included for NS−CT−CT−OS.

Table 1. Nonstandard AMBER Force Field Parameters Used in the DNA Simulation

Bond Parameters

atom types	K_B[a] (kcal/mol Å^2)	R_0[b] (Å)
CT−NS	320.0	1.440
H−NS	434.0	1.000
NS−P	230.0	1.710

Angle Parameters

atom types	K_θ[a] (kcal/mol deg^2)	Θ_0[d] (deg)
CT−CT−NS	50.0	109.50
H1−CT−NS	50.0	109.50
CT−NS−H	35.0	109.50
CT−NS−P	100.0	118.00
H−NS−P	40.0	109.50
NS−P−OS	45.0	100.50
O2−P−NS	100.0	108.23
H1−CT−NS	50.0	109.50
H−NS−H	35.0	109.50

Dihedral Parameters

atom types	N_{tor}[e]	$V_{n/2}$[f] (kcal/mol)	γ[g] (deg)	n[h]
X−NS−P−X	6	0.75	0.0	3
X−CT−NS−X	6	1.00	0.0	3
NS−P−OS−CT	1	0.25	0.0	3
NS−P−OS−CT	1	0.60	0.0	2
CT−NS−P−OS	1	0.125	0.0	3
CT−NS−P−OS	1	1.20	0.0	2

[a] Bond stretching force constant. [b] Equilibrium bond length [c] Bond angle bending force constant. [d] Equilibrium angle. [e] Number of bond paths involving central atom pair − the total torsional energy is divided by this value. [f] Torsional energy (see eq 1 in (ref 18). [g] Phase angle. [h] Fourier component − note both 2- and 3-fold terms for some bond types. [i] Additional atom type: nitrogen in phosphoramidate linkage is NS. NS nonbonded Lennard-Jones parameters are the same as other sp^3 nitrogens: R* = 1.875 and $\epsilon = 0.17$.

Table 2. RESP Charges for Deoxyribose Linked to Phosphoramide Unit[a]

atom	charge	atom	charge	atom	charge
C4′	0.063	H2′1	0.035	O2P	−0.776
H4′	0.118	H2′2	0.035	O5′	−0.495
O4′	−0.420	N3′	−0.849	C5′	−0.007
C3′	0.333	HN3′	−0.377	H5′1	0.075
H3′	0.049	P	1.166	H5′2	0.075
C2′	−0.126	O1P	−0.776		

[a] Other charges (e.g., for C1′, H1′, and nucleoside's bases) are the same as in standard AMBER force field.

Results

Molecular Dynamics Simulations on d(CGCGAATTCGCG)$_2$. Young and Beveridge[33] and Duan and Rosenberg[27] have carried out molecular dynamics simulations on this sequence using the Cornell *et al*. force field and PME long range electrostatics, starting with the crystal structure of this sequence and with the structure of this sequence complexed with the protein Eco-RI. We have carried out simulations on the sequence starting in canonical B and A structures. As had been found by Cheatham and Kollman[26] for the duplex d(CCAACGTTGG)$_2$ both A and B structures converge to a common structure, as can be seen in Table 3 where the average structures of the last 500 ps of the two trajectories (Bav and Aav) have an rms deviation of only 1.2 Å. Also, these structures are 2.7−3.1 Å (all atoms) and 1.9−2.2 Å (central eight residues) from the X-ray structure (9bna) of this sequence and closer to the X-ray structure than to canonical A or B DNA. In Figure 2, we display two views of canonical B, Bav, canonical A, and Aav, which demonstrate that both Aav and Bav have the qualitative features of B-DNA. In Table 4, we present the structure parameters of 9bna, canonical A, canonical B, Aav, and Bav. The dihedral angles and sugar pucker parameters for these structures are presented in the first nine rows; as one can see, α, β, γ, ϵ, and δ for Bav and Aav are all very similar to 9bna and/or canonical B. As also found by Cheatham and

(28) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(29) Ryckaert, J.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(30) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(31) Gaussian 94, Revision D.3; Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian, Inc.: Pittsburgh, PA, 1995.
(32) Bayly, C.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(33) Young, M. A.; Jayaram, B.; Beveridge, D. L. *J. Am. Chem. Soc.* **1997**, *119*, 59−69.

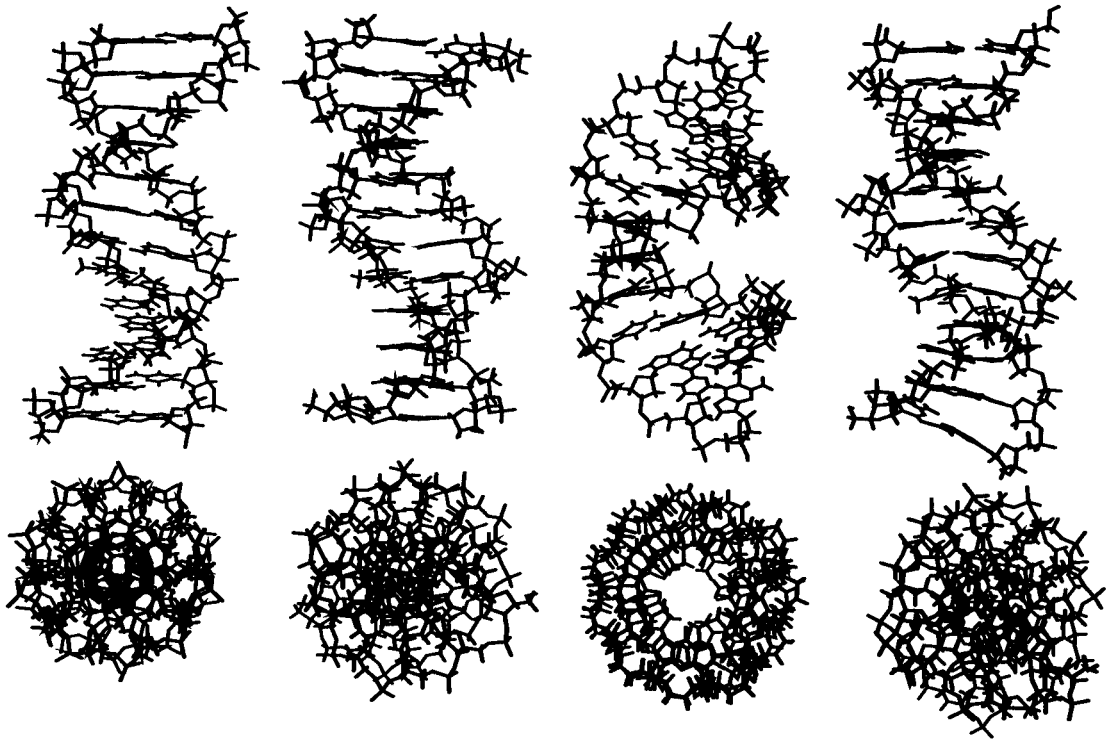

Figure 2. End on the side views of (from left to right) canonical B DNA; average structure starting from B; canonical A DNA and average structure starting from A. The average structures were derived using the final 500 ps of simulation: normal DNA comparison.

Table 3. RMSD Values in Å between Different Structures for Normal d(CGCGAATTCGCG)$_2$[a]

	9bna X-ray	B canonical	A canonical	A-start av last 500 ps	B-start av last 500 ps
9bna X-ray		1.6	6.1	2.7	3.1
B canonical	1.4		6.4	3.1	3.3
A canonical	4.4	4.7		4.7	4.4
A-start av last 500 ps	1.9	2.3	3.6		1.2
B-start av last 500 ps	2.2	2.3	3.6	1.0	

[a] The upper triangle is a comparison using all DNA atoms, and the lower is for the middle eight base pairs.

Table 4. Average Geometrical Parameters for Different Structures for Normal d(CGCGAATTCGCG)$_2$[a]

	9bna X-ray	A canonical	B canonical	B-start av last 500 ps	A-start av last 500 ps
alpha (deg)	292.2	286.1	299.9	290.8	291.7
beta (deg)	172.1	191.1	217.5	171.0	169.7
gamma (deg)	54.4	46.5	39.3	55.6	55.7
delta (deg)	126.7	80.9	149.8	115.5	117.0
epsilon (deg)	193.6	186.3	154.8	192.5	192.7
zeta (deg)	251.7	300.8	272.1	264.8	263.0
chi (deg)	245.8	204.9	262.1	234.6	237.7
W (deg)	138.5	12.3	190.7	120.0	124.3
amplitude (10^2 Å)	33.7	40.1	35.1	40.6	40.4
twist (deg)	35.7	32.7	36.2	31.2	31.1
roll (deg)	1.5	0.4	0.0	6.0	5.1
tilt (deg)	0.0	0.0	0.0	−0.3	−2.8
propellent (deg)	−15.4	3.9	3.8	−13.0	−13.0
rise (deg)	3.33	2.56	3.38	3.37	3.39
x-disp (Å)	−0.38	−5.38	−0.71	−1.68	−1.54
y-disp (Å)	0.09	0.00	0.00	0.12	0.61
tip (deg)	−0.3	0.0	0.0	−0.8	−5.1
buckle (deg)	0.5	0.0	0.0	−0.6	−1.0
opening (deg)	−0.5	−6.4	−5.2	1.9	1.7
inclination (deg)	−1.2	19.2	−6.2	−1.6	−2.0
shear (Å)	0.1	0.0	0.0	0.0	0.0
stretch (Å)	0.0	−0.6	0.0	0.2	0.2
stagger (Å)	0.1	0.1	0.1	−0.1	0.0
shift (Å)	0.1	0.0	0.0	0.0	0.0
slide (Å)	−0.0	0.0	0.0	−0.2	−0.2

[a] Averages for the last structure over the last 500 ps A-start (last column) were calculated without taking into account the last oligonucleotide step (G12-C13), which undergoes opening and closing.

Kollman, χ is calculated to be somewhat smaller for Aav and Bav than for 9bna and the sugar pucker parameters (delta, pucker phase (W), and amplitude (q) reflect sugar pucker phases that correspond to an equilibrium between C2′endo (∼80%) and C3′endo (∼20%) (data for individual sugar puckers not shown in detail but discussed by Cheatham and Kollman).[26]

The remaining rows of Table 4 describe the helicoidal parameters for these models. The most important one which distinguish A and B DNA are twist, x-disp, and inclination. x-disp represents the location of the nucleic acid bases relative to the helix axis; values near zero correspond to structures with no hole in the middle (Figure 2) (B-DNA), and those with a hole (A-DNA) have significant negative values near −4 to −5 Å. As one can see, Aav and Bav are near −1 to −2 Å, similar to values formed by Cheatham and Kollman and close to 9bna and canonical B. Similarly, for inclination, which represents the base inclination relative to the helix axis, both Bav and Aav have small negative values, close to 9bna. As noted by Cheatham and Kollman for the decamer sequence, the twist values are smaller than 9bna and canonical B and close to canonical A. However, the highest resolution NMR data suggest twist values of 30−35°, and the standard deviations for the calculated twist found both by Cheatham and Kollman and here are in the range of ±5. Also, end effects may be operative both in fibers and in the crystal, so smaller duplexes may not have as large a twist, even for B-family structures, as in the fibers or crystals. Unfortunately, NMR is not accurate enough to definitively establish the twist for this sequence to ±5°, and the simulations have a standard deviation in that range.

In Figure 3 we present the average structural properties as a function of the simulation begun in the A form. We do not show the comparable plot for the B form since the average

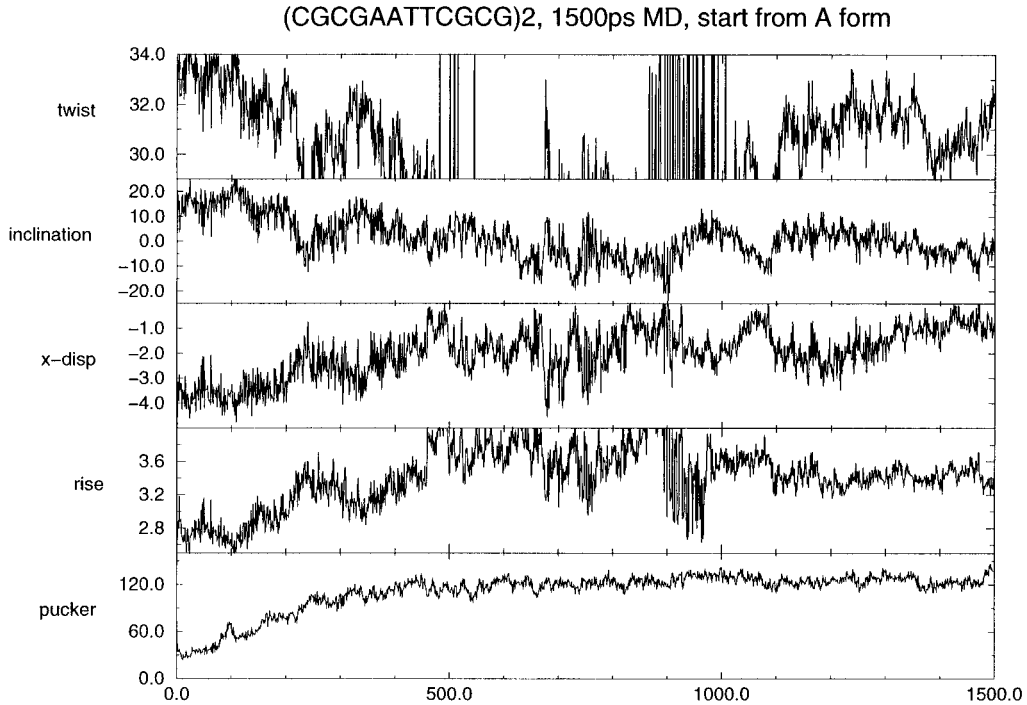

Figure 3. The time dependence of twist (degrees), base inclination (degrees), x-displacement (A), rise (AG) and sugar pucker phase as a function of time for the simulation begun in canonical A. The scale of time is picoseconds.

Table 5. RMSD Å Values between Different Structures for Modified d(CnpGnpCnpGnpAnpAnpTnpTnpCnpGnpCnpG)$_2$[a]

	B canonical (NH out)	A canonical	A-start av last 500 ps	B-start (NH out) av last 500 ps	B-start (NH in) av all 500 ps
B canonical (NH out)		6.4	7.0	6.1	4.6
A canonical	4.7		1.7	2.3	6.6
A-start av last 500 ps	5.1	1.4		1.9	6.9
B-start (NH out) av last 500 ps	4.3	2.0	1.5		5.6
B-start (NH in) av last 500 ps	2.5	4.2	4.3	3.4	

[a] The upper triangle is a comparison using all DNA atoms and the lower is for the middle eight base pairs.

properties remain at comparable values of these variables during the simulation. As one can see, it takes 350−500 ps for the structural parameters to change the A-like values for inclination, pucker, x-disp, and rise to those characteristic of B-like values. The twist shows the widest range and goes off the scale during the period from 499−1000 ps during which one of the end GC base pairs breaks. It has the largest variation even during the last 500 ps of the simulation, which was used to determine the average structures, whose properties were described in Tables 3 and 4.

Molecular Dynamics Simulations of d(CnpGnpCnpGnpAnpAnpTnpTnpCnpGnpCnpG)$_2$

Three molecular dynamics simulations were carried out for this sequence (NPmod). Starting in the A form and in the B form with the N−H group pointing out leads to structures whose average properties are similar and are in the "A" family. In Table 5, these structures are 1.9 Å apart (1.5 Å for the core, the central eight residues). These structures are also 1.7−2.3 Å (1.4−2.0 Å for the core) from canonical A, whereas they are 6.4−7.0 Å (3.0 Å−5.0 Å for the core) from canonical B DNA. In Table 6, the average structural parameters of these two

Table 6. Average Geometrical Parameters for Different Structures for Modified d(CnpGnpCnpGnpAnpAnpTnpTnpCnpGnpCnpG)$_2$

	A canonical	B canonical	A-start av last 500 ps	B-start (NH out) av last 500 ps	B-start (NH in) av last 500 ps
alpha (deg)	286.1	299.9	287.4	284.5	263.2
beta (deg)	191.1	217.5	174.6	178.3	150.3
gamma (deg)	46.5	39.3	57.2	62.5	85.6
delta (deg)	80.9	149.8	92.5	88.0	125.4
epsilon (deg)	186.3	154.8	202.2	195.5	231.6
zeta (deg)	300.8	272.1	274.9	288.5	250.0
chi (deg)	204.9	262.1	218.5	208.4	253.7
pucker (deg)	12.3	190.7	54.1	43.5	133.2
amplitude (10^2 Å)	40.1	40.1	40.2	40.1	41.2
twist (deg)	32.7	36.2	29.1	28.7	24.3
roll (deg)	0.4	0.0	8.4	8.7	2.9
tilt (deg)	0.0	0.0	−0.5	1.0	0.2
propeller (deg)	3.9	3.8	−5.9	−7.2	−3.7
rise (Å)	2.56	3.38	3.24	3.11	3.58
x-disp (Å)	−5.38	−0.71	−2.45	−3.30	−0.99
y-disp (Å)	0.00	0.00	0.18	−0.19	−0.04
tip (deg)	0.0	0.0	−1.3	2.0	−0.6
buckle (deg)	0.0	0.0	−0.5	0.8	0.3
opening (deg)	−6.4	−5.2	3.0	3.0	2.7
inclination (deg)	19.2	−6.2	1.5	2.5	16.4
shear (Å)	0.0	0.0	0.0	0.0	0.0
stretch (Å)	−0.6	0.0	0.3	0.3	0.1
stagger (Å)	0.1	0.1	−0.3	−0.3	0.2
shift (Å)	0.0	0.0	0.0	0.0	0.0
slide (Å)	0.0	0.0	−0.4	−0.4	−0.1

structures (Aav and Boutav) are presented. Here, there is no X-ray structure or fiber structure for this molecular so we use canonical A and B structures from normal DNA. Examining the dihedral angles and the sugar pucker, we see that the χ dihedral and sugar pucker are much closer to those characteristic of A than B structures. The remaining dihedrals do not distinguish a structure. Examining the helicoidal parameters in Table 6, we see that the x-disp is more A-like than found for the corresponding simulated structures for unmodified DNA but not as negative as canonical A-DNA. The inclination is intermediate between A and B as well.

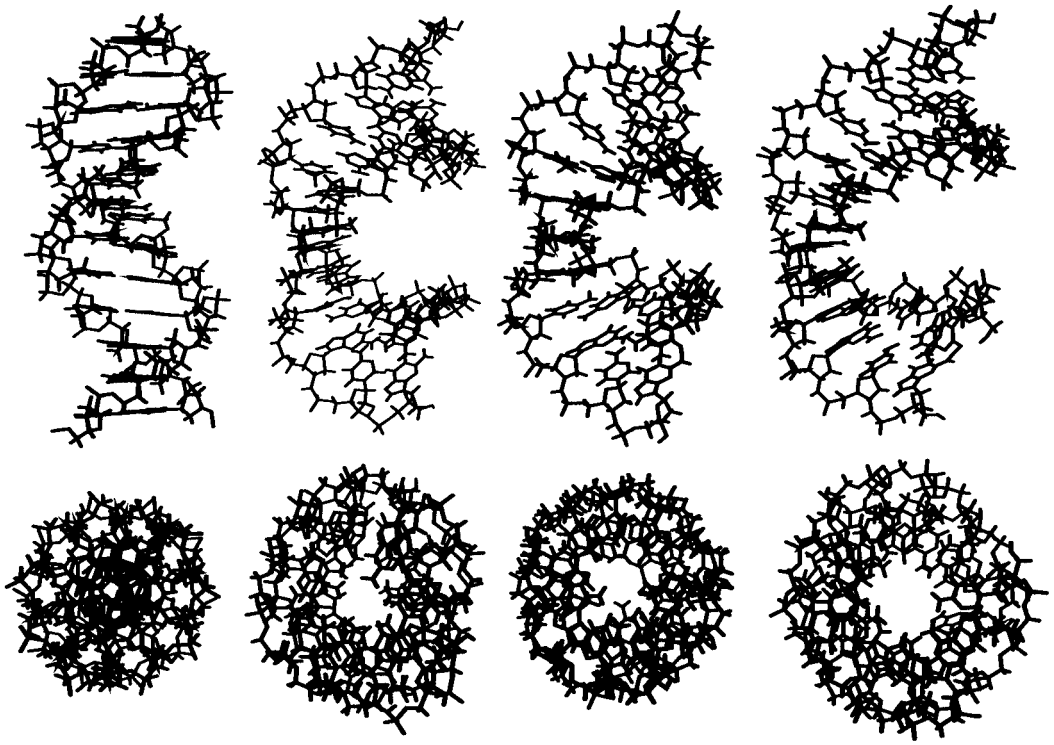

Figure 4. Same as Figure 2 for N3′−H3′ modified DNA (N−P modified DNA comparison).

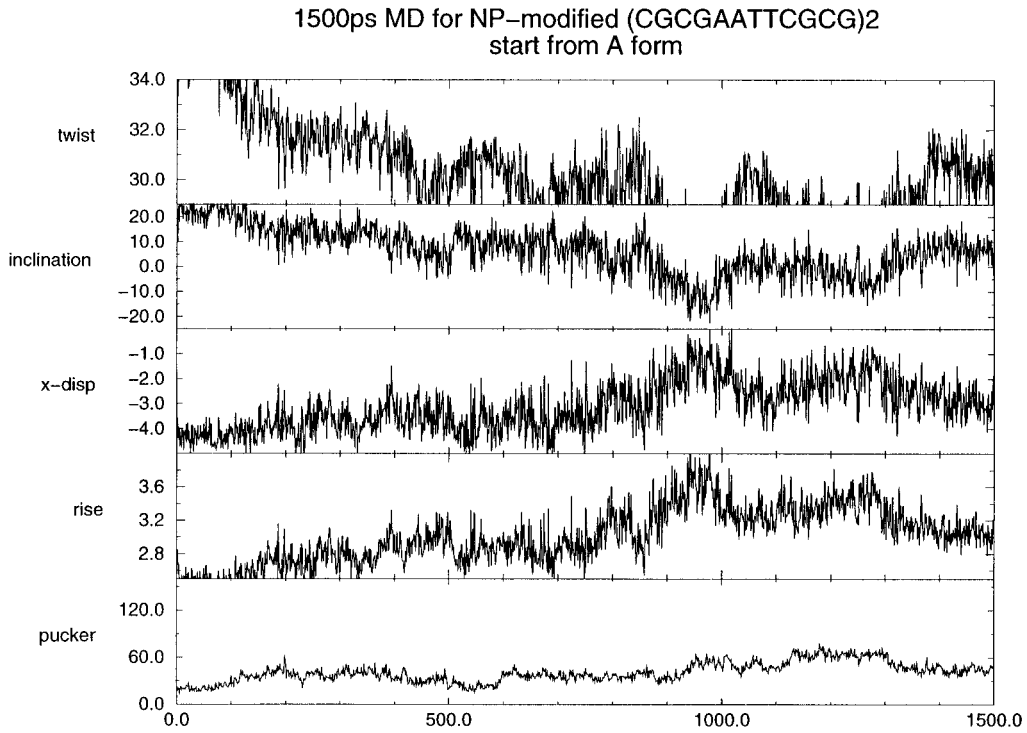

Figure 5. Same as Figure 3 for N3′−N3′ modified DNA starting in B form (N3′−H3′ out).

The twist is closer to canonical A and B and, as shown in Figure 4, the Aav and Boutav structures have a shape and a "hole in the middle" that are closer to the A family than B. In Figures 5 and 6, the dependence of the crucial geometrical variables as a function of time are presented for the simulations beginning in the A and Bout forms. As one can see, there is a significant range of standard deviation in all the variables—twist, inclination, x-disp, rise, and pucker—but all are more characteristic of A than B. For example, x-disp generally has a value in the range of −3 Å but spends a significant amount of time near −4 Å, and inclination in the end of both simulations is in the range of +10°, close to canonical A. In contrast, in normal DNA (Figure 2), in the last 500 ps, x-disp is almost never as low as −3 Å and inclination is nearer 0° rather than 10°. The sugar pucker in NPmod averages near −40°, which corresponds to mainly C3′endo sugar puckers, i.e., ∼80% C3′-endo and 20% C2′-endo.

From Table 6 and Figure 4, it is clear that the B(NH-in) structure does not convert from B to A, even in 1 ns, although its inclination becomes more A-like and it becomes a longer, thinner helix, with a rise even 0.2 Å larger than found in both B-DNA and our simulation on the normal dodecamer.

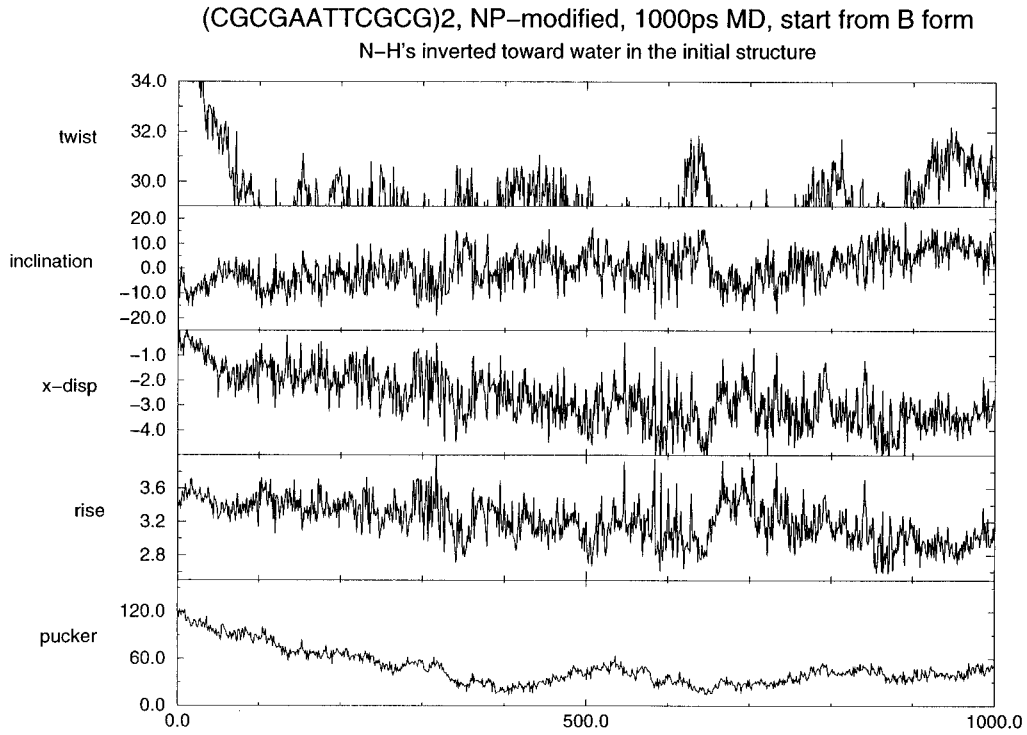

Figure 6. Same as Figure 3 for N3′−H3′ modified DNA starting in A form.

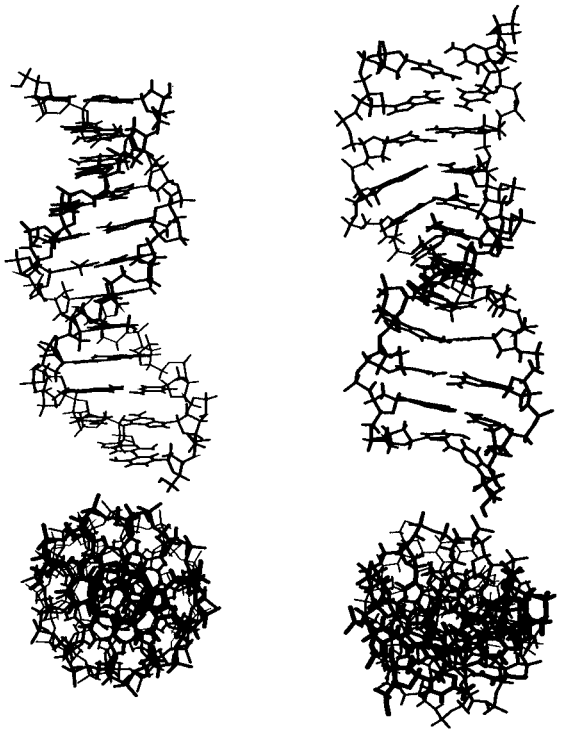

Figure 7. End on end side view of BDNA start and average structure for B(N3′−H3′, starting with N−H in). The average structure was derived from the final 500 ps of simulations: (left) N−P modified DNA comparison and (right) N−H oriented, inside DNA.

Base Pair Opening in DNA Duplexes

In the simulation of the DNA dodecamer beginning in the A form, we have observed that the end base pair opens at ∼400 ps and reforms at ∼1100 ps. There are features of this base pair opening and closing that are of interest.[34−36] In Figure 7, we display stereoviews of the GC base pair opening, which begins at 377 ps with the base pair still intact and ends at 480 ps, with it broken. As one can see at 377 ps, two water molecules are hydrogen bonding to the N3 and O2 of cytosine from below and, at 378 ps, replace the guanine N1−H and N2−H as proton donors. This is followed by the cytosine O2 twisting to form a bifurcated hydrogen bond with the guanine N1−H and N2−H at 381 ps, with a water forming a bridging hydrogen bond between guanine O6 and cytosine N3. The bifurcated H-bond remains until 458 ps and breaks between 458 and 459 ps, with the views at 476 and 480 ps showing a fully exposed cytosine. It is intuitively reasonable that the guanine, which can form a stronger van der Waals interaction, stays well stacked with the preceding base pair, while the cytosine swings out into solution. In Figure 8, we begin by showing a stereoview at 880 ps, with the cytosine very far from its base partner and fully solvated. At 1088 ps, it is still well solvated, but its carbonyl O2 has formed a hydrogen bond with the 5′OH, bringing it into closer proximity to its guanine partner. This hydrogen bond remains as the cytosine is moving to hydrogen bond with guanine, first with its carbonyl bifurcating the N3−H and N2−H, which was the "jumping off point" for the base pair to break (up to 1094 ps) and then, at 1099 ps and continuing, with the base pair reformed.

In Figures 10 and 11 (Supporting Information) we display the distances of key H-bonding atoms for the CG base pair that breaks (Figure 10) and the dyad related base pair (Figure 11) that stays intact. In the case of the hydrogen bond that breaks, we show the time dependence of key atom−atom distances from 350−1130 ps. One can see here what was evident in the stereoviews: the N1−N3 and O6−N4 hydrogen bonds break first, with N2−O2 remaining near 3 Å and N1−O2 shortening from 4 to 3 Å. Then, beginning at 450−460 ps, all the bonds break. At around 1090 ps, one can see the N2−O2 and N1−O2 distances shorten, followed by the N1−O2 distance slightly lengthening. First the N1−N3 distance shortens, then, finally, the O6−N4 distance shortens to a value characteristic of a hyrogen bond. For the remaining time from 1100−1500 ps (not shown) the hydrogen bonds remain intact.

In Figure 10 (Supporting Information) we see that the dyad related hydrogen bond remains intact, with two interesting transient openings of the hydrogen bonds at ∼200 and ∼1000

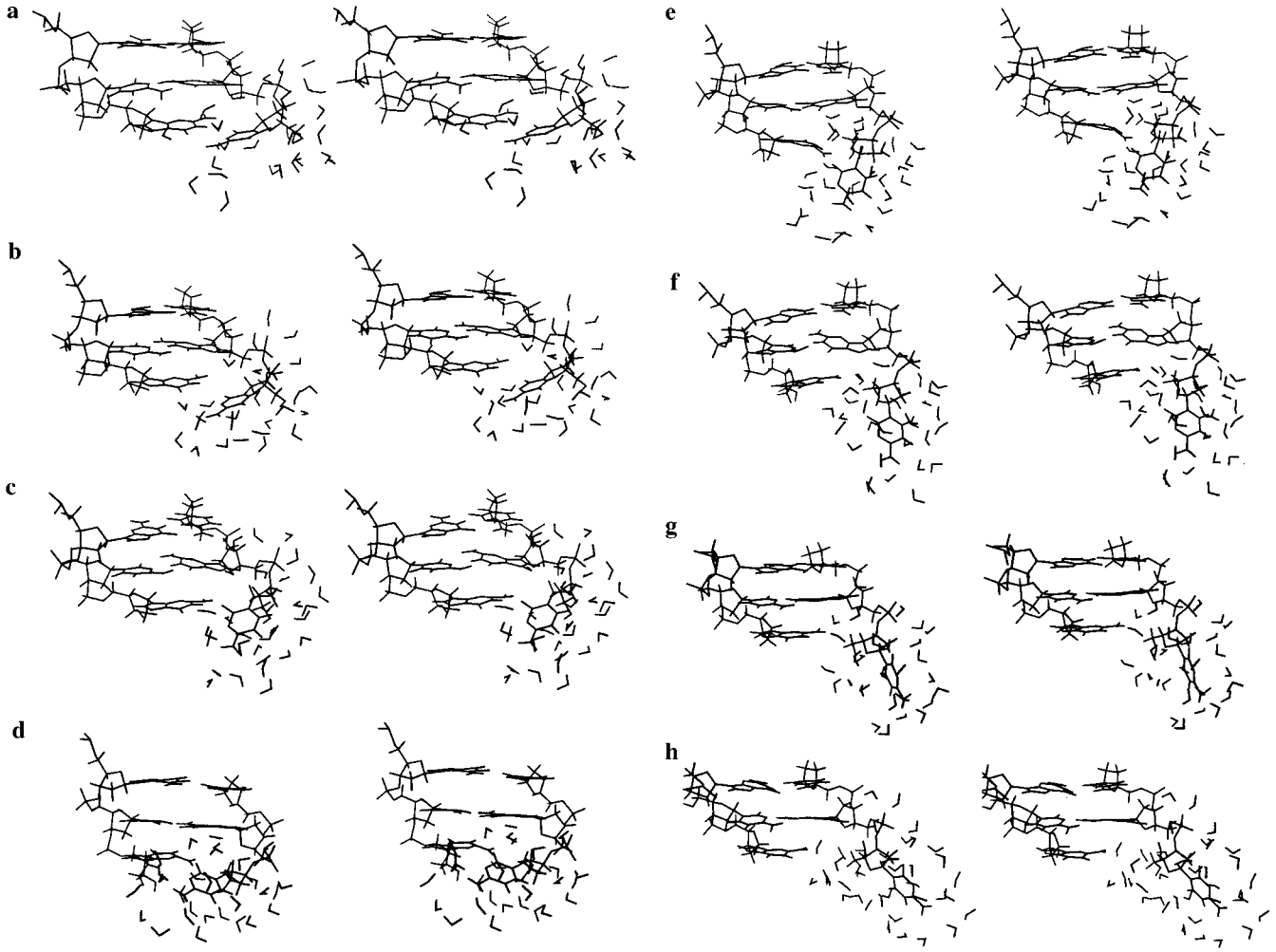

Figure 8. (a)−(h): Stereoviews of end three base pairs as a function of time, including those waters within 4 Å of the terminal cytosine, which breaks its H-bond with the terminal G: (a) 377, (b) 378, (c) 381, (d) 457, (e) 458, (f) 459, (g) 476, and (h) 480 ps.

ps. It is also (Figure 11 is Supporting Information) of some interest to monitor the χ value of the cytosine and guanine whose hydrogen bond breaks (G12−G13) and that which remains intact (C1−G24) (Figure 12 in Supporting Information). Even for the intact base pair, there are considerable fluctuations in the base pair, albeit it stops in the anti ($\chi \sim 250°$) region. C13, on the other hand, undergoes transitions to the syn conformation as the base pair breaks but does not stay for a long time at syn until ∼620 ps. Surprisingly, it is syn during the base pair reforming at ∼1100 ps and returns to anti, while the base pair is intact at ∼1300 ps. Interestingly, G12 as reflected in χ is actually much less mobile when its hydrogen bonds to C13 are broken, presumably because it can satisfy its starting interaction without "competition" from its base pair hydrogen bonding.

Discussion and Conclusions

We have presented the application of unrestrained molecular dynamics simulations to model A → B and B → A DNA transitions in water solution for normal and N3′−P modified DNA, respectively.

It is encouraging that both the "B" preference of the normal dodecamer and the "A" preference of the N3′ modified dodecamer are found here in unrestrained simulations. Nonetheless, the dependence of the N3′ modified simulation structure on starting geometry of the N3′−H is disconcerting. With the N3′−H "in", the structure starting with B is still more B like than A-like after 1 ns, in contrast to the simulations starting in A or B with the N3′−H "out". The convergence of the latter two to an A-like structure suggests that this is a lower energy, but one cannot prove that. In any case, these results emphasize the importance of using multiple starting geometries for a given system to more completely evaluate model predictions.[37] The time scale for the B → A transition in N3′ modified DNA is comparable (300−500 ps) to that found for the A → B transition in normal DNA.

It should be emphasized that we used both *ab initio* data (to derive partial charges and the torsional parameters in CH_3-$NHPO_2^-OCH_3$) and experimental data (the relative sugar pucker preference of 3′H, 3′NH_2, and 3′OH substituted sugars, which suggested that no 2-fold torsion of NS−CT−CT−OS was necessary, in contrast, to OS−CT−CT−OS) to derive the molecular mechanical parameters used here, as we have done previously for unmodified nucleic acids.[2]

This work is nicely complementary to theoretical studies of Veal and Brown,[38] who studied 3′S and 5′S substituted nucleosides using a combination of *ab initio* and molecular mechanics/dynamics calculation. They concluded that 5′S has unfavorable steric and electrostatic interactions, and 3′S is

(34) Ramstein, J.; Lavery, R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7231−7235.
(35) Briki, F.; Ramstein, J.; Lavery, R.; Genest, D. *J. Am. Chem. Soc.* **1991**, *113*, 2490−2493.
(36) Pohorille, A.; Ross, W. S.; Tinoco, I., Jr. *Int. J. Supercomp. App.* **1990**, *4*, 81−96.
(37) Aufinger, P.; Louise-May, S.; Westhof, E. *J. Am. Chem. Soc.* **1995**, *117*, 6720−6726.
(38) Piskur, J.; Rupprecht, A. *FEBS Lett.* **1995**, *375*, 174−178.

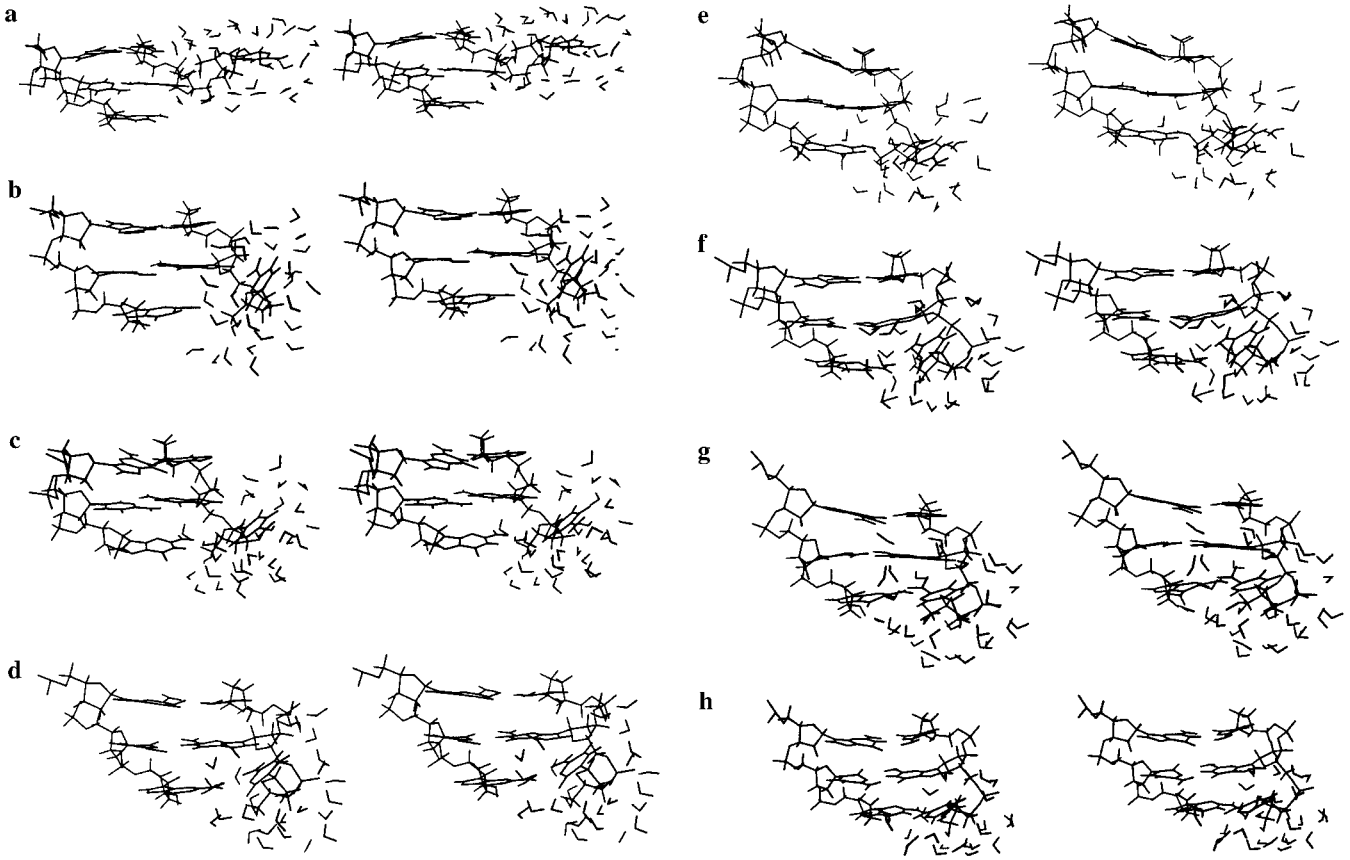

Figure 9. (a)−(h): Same as Figure 7, during the reformation of the base pair: (a) 880, (b) 1088, (c) 1089, (d) 1090, (e) 1092, (f) 1093, (g) 1094, and (h) 1099 ps.

detrimental to B DNA but tolerated in A DNA structures. Their result for 3′S is quite consistent with that found here for 3′NH, and their result for 5′S ties in nicely with the experimental observation, noted above, that 5′ phosphoramidates do not form duplexes.

It makes sense that, given the preference of 3′OH to favor C2′ endo more than 3′NH_2 (3′H), O3′ (normal DNA) containing nucleotides would favor B more than N3′, because the stability of A DNA requires C3′ endo sugar puckers. Our studies also allow us to rationalize why O5′ → N5′−H does not favor the A form of DNA,[9] since this change will not effect sugar puckering (both A and B DNA have γ(C4′−C5′) in the g^+ conformation).

Because we have not calculated the relative free energy of the double → single strand transition in normal vs phosphoramidate modified DNA, we cannot address the greater intrinsic stability of the phosphoramidate duplexes (in the A form) compared to the comparable sequences of normal DNA (in the B form). RNA duplexes do generally melt higher (are more stable) than DNA, despite their greater rigidity. Of course, it is the relative rigidity of single and double stranded forms that is key, and we cannot easily assess the effect on stability of the 2′OH group. Nonetheless, a possible interpretation of the greater stability of phosphoramidate than normal duplexes is that the nonblocked properties of A helices are intrinsically more stable than B and the generally ubiquitous observation of BDNA duplexes is due to the sugar pucker preference for C2′ endo of ∼1 kcal/mol per sugar. This sugar pucker preference can be "overcome" if, for example, the DNA is dehydrated in high salt or mixed alcohol/water solutions, where the DNA helix becomes A[39] and our calculations, indeed, find an A structure stable in 85% ethanol/water for >3 ns.[40] Nonetheless, it would be very valuable to be able to calculate the relative free energies of A and B structures, and we are currently attempting to do this for A and B RNA,[41] where the larger barrier in RNA allows each conformational family to be stable for >1 ns.

Acknowledgment. These calculations were carried out using the facilities of the UCSF Computer Graphics Lab (T. Ferrin, supported by NIH-RR-1081) and the Pittsburgh Supercomputer Center (MCA935017P). We are also grateful for research support from the NIH (CA-25644 to P.A.K.). P.C. was partially supported by The Polish Committee for Scientific Research, Grant KBN-3-T09A-096-10. We are grateful to S. Gryaznov for sending us preprints prior to publication.

Supporting Information Available: Figures of H-bonding distances involving G12 and C13 and over entire range of trajectory for C1−G24 and plot of χ for C13 (4 pages). See any current masthead page for ordering and Internet access instructions.

JA963909J

(39) Veal, J.; Brown, F. *J. Am. Chem. Soc*. **1995**, *117*, 1873−1880.
(40) Cheatham, T.; Crowley, M.; Fox, T.; Kollman, P. A Molecular Level Picture of the Stabilization of A-DNA in Mixed Ethanol-water Solutions. *Proc. Natl. Acad. Sci. U.S.A.* In press.
(41) Cheatham, T.; Duan, Y.; Rosenberg, J.; Crowley, M.; Kollman, P. A. Studies in progress.