Accurate DNA Dynamics without Accurate Long-Range Electrostatics

Alexey K. Mazur

Contribution from the Laboratoire de Biochimie Théorique, CNRS UPR9080, Institue de Biologie Physico-Chimique, 13, rue Pierre et Marie Curie, Paris 75005, France

Received May 1, 1998. Revised Manuscript Received July 13, 1998

Abstract: Molecular dynamics (MD) simulations of the DNA dodecamer d(CGCGAATTCGCG)$_2$, with the bulk solvent represented implicitly but the minor groove filled with explicit water, demonstrate the excellent stability of both the duplex in B form and the water shell and converge to the same final state starting from different initial conformations. This opens the way to MD simulations of much longer sequences. The final state appears surprisingly close to the experimental B-DNA structures, much closer than in earlier, more expensive simulations involving explicit counterions, periodic boundary conditions, and particle mesh Ewald evaluations of Coulomb forces. This paper presents the first application of a new internal coordinate molecular dynamics (ICMD) method to nucleic acids. Two ICMD trajectories computed with a 10-fs time step exhibit the same qualitative features and converge to the same structure as the traditional Cartesian coordinate MD.

Introduction

DNA dynamics arguably play the key role in molecular biology, and since the first pioneering simulations,[1] a large number of original computational studies have been published (comprehensive surveys of the literature can be found in the recent reviews[2−4]). Only recently, however, it has become possible to obtain in calculations dynamically stable and reasonably accurate DNA structures. This long-awaited progress is generally attributed to improved force field parameters,[5,6] to the new particle mesh Ewald (PME) algorithm for evaluation of Coulomb forces,[7] and finally, to the growing computer power. The last factor is essential because the PME approach is computationally expensive. When convergence is necessary, it requires supercomputer resources even for small oligomers because the solute molecule must be placed in a large enough water box to accommodate all neutralizing counterions without exceeding reasonable levels of effective DNA and salt concentrations.

One of the most promising new methods that can significantly reduce the computational cost of molecular dynamics (MD) trajectories is the internal coordinate molecular dynamics (ICMD)[8−11] (see ref 11 for a historical review). Unlike traditional MD, it employs torsions and, if desired, valence angles and bond lengths as generalized coordinates in the equations of motion. ICMD originates from the Euler−Lagrange−Hamilton formalism of classical mechanics and makes possible modeling of polymers as chains of rigid bodies, which automatically eliminates the most severe time step limitations characteristic for Newtonian MD.

Nucleic acids present a special and perhaps the most difficult case for ICMD because this method is generally applicable only to polymers with tree topologies.[12] The latter property is broken by any closed ring of chemical bonds if the whole group is not treated as rigid. In nucleic acids, however, pseudorotation of furanose rings is largely responsible for their polymorphism and structural flexibility; therefore, any approach that considers sugars frozen would have a limited value, if any. We have recently studied the possible ways to allow for the ring flexibility in ICMD.[13] It has been shown that a rational solution of the problem exists, and a few practical algorithms have been proposed.

The advantages of ICMD should be most significant when the price of one step is very high, like in PME calculations of DNA dynamics. There are many important domains, however, where the PME approach can hardly be applied, in the foreseeable future, even with 10 times larger steps, for instance, dynamics of long linear DNA fragments or plasmids. It is interesting, therefore, to try less demanding alternatives, and here, we report the results of such an attempt.

We show that a combination of a partial explicit hydration with long-known approximate approaches involving reduction of phosphate charges and a distance-dependent dielectric function solves the problem. Quite unexpectedly, however, long time trajectories converge to conformations that are remarkably closer to experimental data than ever before. We consider here dynamics of a DNA duplex with the sequence d(CGCGAAT-

* Corresponding author. Fax: (33-1) 43.29.56.45. E-mail: alexey@ibpc.fr.

(1) Levitt, M. *Cold Spring Harb. Symp. Quant. Biol.* **1983**, *47*, 251−262.
(2) Jayaram, B.; Beveridge, D. L. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 367−394.
(3) Ravishanker, G.; Auffinger, P.; Langley, D. R.; Jayaram, B.; Young, M. A.; Beveridge, D. L. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1997; pp 317−372.
(4) Auffinger, P.; Westhof, E. *Curr. Opin. Struct. Biol.* **1998**, *8*, 227−236.
(5) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(6) MacKerell, A. D., Jr; Wiórkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.
(7) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(8) Mazur, A. K.; Abagyan, R. A. *J. Biomol. Struct. Dyn.* **1989**, *6*, 815−832.
(9) Jain, A.; Vaidehi, N.; Rodriguez, G. *J. Comput. Phys.* **1993**, *106*, 258−268.
(10) Rice, L. M.; Brünger, A. T. *Proteins: Struct., Funct., Genet.* **1994**, *19*, 277−290.
(11) Mazur, A. K. *J. Comput. Chem.* **1997**, *18*, 1354−1364.
(12) Abagyan, R. A.; Mazur, A. K. *J. Biomol. Struct. Dyn.* **1989**, *6*, 833−345.
(13) Mazur, A. K. *J. Chem. Phys.*, submitted.

10.1021/ja981498w CCC: $15.00
Published on Web 10/08/1998

$TCGCG)_2$. This dodecamer was the first to crystallize in the B form of DNA,[14,15] and since then, it has become perhaps the most studied DNA fragment both experimentally and theoretically.[1,16−21] In the recent years it has been often used in benchmark tests of new force fields and algorithms.[22−29] We demonstrate that the new ICMD technique makes possible simulations with a 10-fs time step for DNA in explicit water. For comparison, one of the long ICMD trajectories has been repeated by using a traditional Cartesian coordinate MD technique. It appears that additional constraints employed in ICMD have no significant effect upon the results.

Methods and Simulation Protocols

The equations of motion and the numerical algorithm used for ICMD calculations are described in detail elsewhere.[11,13] This method makes possible arbitrary freezing of internal degrees of freedom, but here, we try to reduce their number to a reasonable minimum. Therefore, except for sugar rings, the geometry of the nucleotides was fixed, that is the bases were rigid and all other bond lengths and bond angles were fixed. It is well-known, however, that five-membered rings remain flexible only if bond angles vary;[30] therefore, torsional dynamics of nucleic acids would mean dynamics with frozen sugars. Given the necessity to vary intracyclic bond angles, we preferred to keep flexible all other adjacent valence angles, which means that sugars in our simulations had the same degrees of freedom as in Cartesian MD with fixed bond lengths. The fastest motions and time step limitations for this model system are considered in a special section below.

In all calculations, AMBER94[5] parameters were used with rigid TIP3P water[31] and no cutoff schemes. The following mixed strategy was employed for modeling solvent effects. On one hand, phosphate charges were reduced by 0.5 eu and a linear distance-dependent dielectric function $\epsilon = r$ was used. On the other hand, a certain number of explicit water molecules were added by the following procedure. The DNA molecule was first covered by a 5 Å thick water shell. After that, cylinder-like volumes around each strand were built from spheres centered at phosphorus atoms with radii of 12 Å. All water molecules that appear outside the intersection area of the two volumes have been removed. The solvent remaining was next relaxed by energy minimization, first with the solute held rigid and then with all degrees of freedom. This procedure gives a partially hydrated duplex, with the minor groove completely filled and a few solvent molecules in the major groove. Implicit modeling of solvent effects has been long known in conformational analysis of nucleic acids.[32−34] To our knowledge, however, the mixed strategy used here has never been tried.

Stability and step size limits are evaluated by the method specifically designed for the leapfrog integrator,[35] with the same protocols as before.[36] The test trajectory is repeatedly calculated with different time steps, and appropriate system averages are compared with "ideal" values, i.e., the same parameters evaluated with a very small time step. One of the appropriate averages[3] is the total energy computed as, $E = \bar{U} + \bar{K}$, where $\bar{U}$ and $\bar{K}$ are the average potential and kinetic energies computed for integer steps and half-steps, respectively. A deviation of $0.2D[U]$, where $D[.]$ denotes operator of variance, is taken as an upper acceptable level. The step size maximum determined is denoted as h_c and called "characteristic". Virtually harmonic conditions are simulated by reducing the temperature down to 1 K with the same equilibration protocol as before.[36] The duration of the test trajectory was 10 ps. Spectral densities of autocorrelation functions of generalized velocities were calculated from separate 50-ps trajectories at 1 K.

The two long ICMD trajectories of a partially hydrated duplex $d(CGCGAATTCGCG)_2$ were obtained as follows. The first trajectory, below referred to as TJA, was calculated in internal coordinates with constraints outlined above and the following initialization procedure. The canonical B form was first constructed with program JUMNA[34] with Arnott B73 parameters.[37] It was first energy minimized without water, which gives a conformation with the root mean square deviation (rmsd) of 2.6 Å from the canonical B form. Next, the above hydration procedure was applied, which added 134 water molecules and reduced the rmsd to 2.3 Å. The trajectory was initiated with zero solvent temperature by giving the solute a kinetic energy corresponding to 300 K, and the system was equilibrated during 7 ps as described earlier.[11,36]

The second trajectory, denoted TJB, was calculated with the same degrees of freedom but starting from a "kinked" duplex conformation in the *Eco*RI endonuclease complex. This conformation[38] (file PDE001 in Nucleic Acids Database[39]) has been hydrated as above without preminimization to keep the "kinked" form intact. The number of remaining water molecules was 114.

Finally, the Cartesian trajectory, referred to as TJC, was computed from the TJA starting state but with a completely free solute model and atom Cartesian coordinates as dynamic variables. Water molecules remained rigid, and their motion was calculated in the same way as in TJA.

All three trajectories had the same duration of 5 ns. In production runs the temperature was maintained at 300 K by the Berendsen algorithm[40] with a large relaxation time of 10 ps. Conformations were saved with a 2.5-ps interval. Structures from the last nanosecond were used for computing an average conformation referred to as the final MD state of the corresponding trajectory. Cartesian coordinates of all nonhydrogen DNA atoms were superimposed and averaged, and thus obtained structures were used in further computations without additional corrections of chemical geometry.

The average conformational and helicoidal parameters were computed from outputs of Curves procedure[41] applied to the final MD states. The rms deviations of the corresponding local values are referred to as sequence variances. They are distinguished from time variances which characterize fluctuations of similar average structural parameters during

(14) Drew, H.; Wing, R.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Proc. Natl. Acad. U.S.A.* **1981**, *78*, 2179−2183.
(15) Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Nature* **1980**, *287*, 755−758.
(16) Rao, S. N.; Kollman, P. *Biopolymers* **1990**, *29*, 517−532.
(17) Srinvasan, J.; Withka, J. M.; Beveridge, D. L. *Biophys. J.* **1990**, *58*, 533−547.
(18) Swaminathan, S.; Ravishanker, G.; Beveridge, D. L. *J. Am. Chem. Soc.* **1991**, *113*, 5027−5040.
(19) Miaskiewicz, K.; Osman, R.; Weinstein, H. *J. Am. Chem. Soc.* **1993**, *115*, 1526−1537.
(20) Kumar, S.; Duan, Y.; Kollman, P. A.; Rosenberg, J. M. *J. Biomol. Struct. Dyn.* **1994**, *12*, 487−525.
(21) McConnell, K. J.; Nirmala, R.; Young, M. A.; Ravishanker, G.; Beveridge, D. L. *J. Am. Chem. Soc.* **1994**, *116*, 4461−4462.
(22) York, D. M.; Yang, W.; Lee, H.; Darden, T.; Pedersen, L. G. *J. Am. Chem. Soc.* **1995**, *117*, 5001−5002.
(23) Yang, L.; Pettitt, B. M. *J. Phys. Chem. B* **1996**, *100*, 2564−2566.
(24) MacKerell, A. D., Jr *J. Phys Chem. B* **1997**, *101*, 646−650.
(25) Duan, Y.; Wlkosz, P.; Crowley, M.; Rosenberg, J. M. *J. Mol. Biol.* **1997**, *272*, 552−572.
(26) Young, M. A.; Ravishanker, G.; Beveridge, D. L. *Biophys. J.* **1997**, *73*, 2313−2336.
(27) Young, M. A.; Jayaram, B.; Beveridge, D. L. *J. Am. Chem. Soc.* **1989**, *119*, 59−69.
(28) Cieplak, P.; Cheatham, T. E., III; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 6722−6730.
(29) Sprous, D.; Young, M. A.; Beveridge, D. L. *J. Phys. Chem. B* **1998**, *102*, 4658−4667.
(30) Gō, N.; Scheraga, H. A. *Macromolecules* **1970**, *3*, 178−187.
(31) Jorgensen, W. L.; Chandreskhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys* **1991**, *79*, 926−935.
(32) Tidor, B.; Irikura, K. K.; Brooks, B. R.; Karplus, M. *J. Biomol. Struct. Dyn.* **1983**, *1*, 231−251.
(33) Mazur, J.; Jernigan, R. L. *Biopolymers* **1991**, *31*, 1615−1629.
(34) Lavery, R.; Zakrzewska, K.; Sklenar, H. *Comput. Phys. Commun.* **1995**, *91*, 135−158.
(35) Mazur, A. K. *J. Comput. Phys.* **1997**, *136*, 354−365.
(36) Mazur, A. K. *J. Phys. Chem. B* **1998**, *102*, 473−479.
(37) Arnott, S.; Hukins, D. W. L. *Biochem. Biophys. Res. Commun.* **1972**, *47*, 1504−1509.
(38) McClarin, J. A.; Frederick, C. A.; Wang, B.-S.; Greene, P.; Boyer, H. W.; Grable, J.; Rosenberg, J. M. *Science* **1986**, *234*, 1526−1541.
(39) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S. H.; Srinivasan, A. R.; Schneider, B. *Biophys J.* **1992**, *63*, 751−159.
(40) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(41) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63−91.

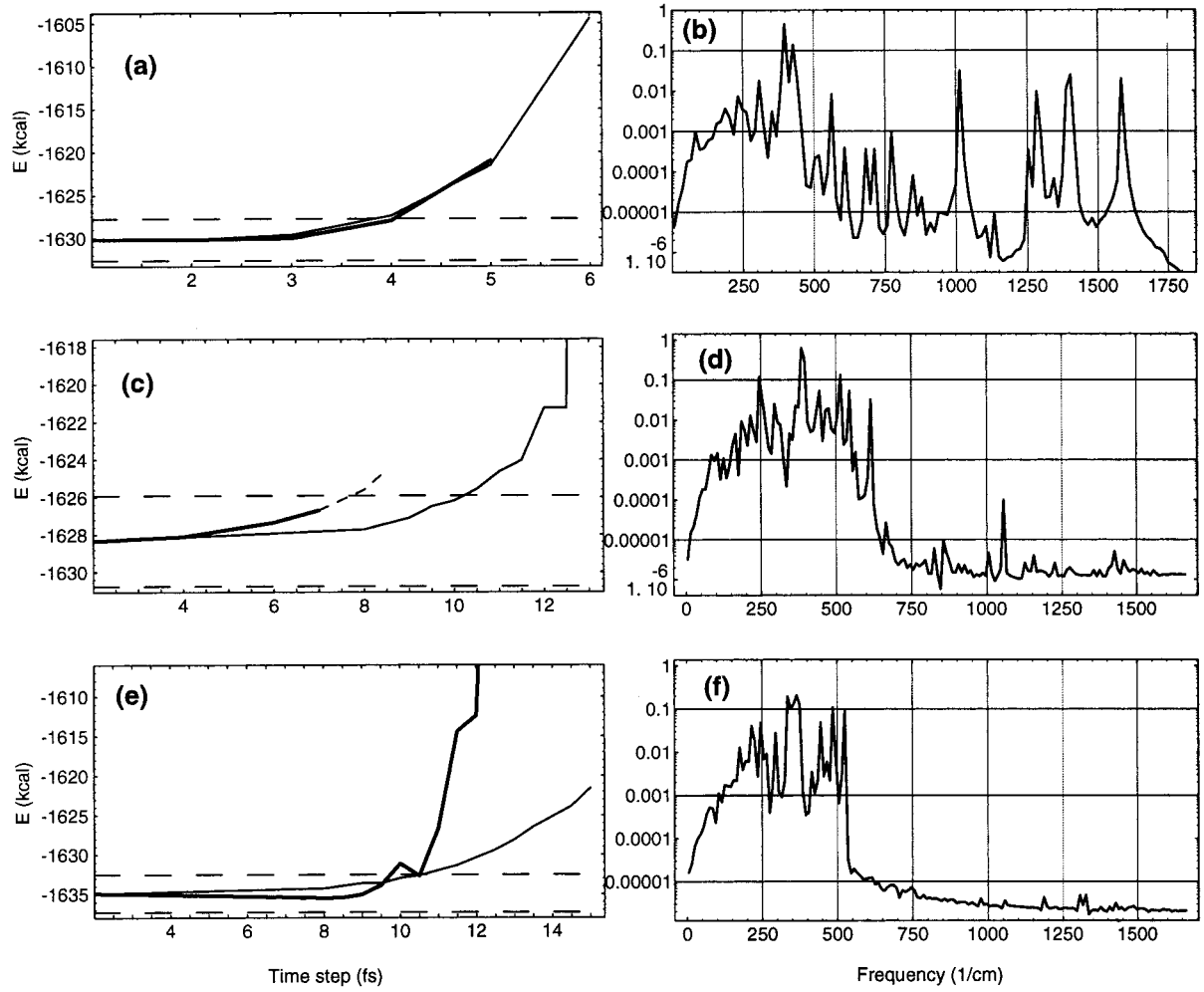

Figure 1. Time step dependencies of the average total energy (a, c, e) and corresponding spectral densities (b, d, f) for three models of the partially hydrated DNA duplex. Thinner traces in panels a, c, and e show results for virtually harmonic conditions when temperature was lowered down to 1 K. The DNA molecule has fixed bond lengths, rigid bases, and fixed valence angles except intra- and extracyclic bond angles in sugars. (a, b) No modifications of inertia. Signals in panel b with frequencies beyond 600 cm^{-1} correspond to bond bending of hydrogens. (c, d) Inertia of hydrogen-only rigid bodies modified as explained in the text. (e, f) Inertia of sugar ring atoms also modified as described in the text. Spectral densities are shown for a C2′−C1′−O4′ angle in one of the sugars.

MD. The latter are obtained by applying Curves[41] to individual states in trajectories. The time variance is the rms fluctuation of an instantaneous value from its time average.

Canonical A and B forms used for all comparisons were constructed with NUCGEN utility of AMBER.[42] The reference X-ray conformation[14] corresponds to file 1bna in Protein Database.[43]

Results and Discussion

Fast Motions and Time Step Limitations. Figure 1 shows representative results of time step tests for the model system used in TJA. As already noted, we preferred to keep several fast bond angle bending modes around sugar rings, rather than violate the symmetry by freezing exocyclic angles. The fastest such motion is the scissors H−C−H mode with a frequency around 1600 cm^{-1} which according to theory[35] should limit h_c to approximately 3.6 fs. Figure 1a,b confirms this a priori estimate. Note that h_c does not change with temperature indicating the harmonic nature of the limitation.

To raise h_c to the level of 10 fs, which has been found optimal, in a certain sense, for in-water simulations of proteins,[36] we modify inertias of hydrogen-only rigid bodies as has been proposed in our previous studies.[11,36] Additional moments of inertia of 9 and 4 amu·Å^2 are applied to C−H bonds and other hydrogen-only rigid bodies, like thymine methyls, respectively. Water molecules are given additional inertia of 5 amu·Å^2. With these modifications the time step testing gives results shown in Figure 1c. Note that the low-temperature h_c is increased to the desired level and, accordingly, all signals beyond 600 cm^{-1} in the spectrum in Figure 1b have disappeared. A few remaining small peaks apparently are overtones since the h_c value obtained is close to the harmonic limit for a frequency of 600 cm^{-1}.[35] This frequency is close to the fastest pseudorotation normal mode.[44]

The normal temperature trace in Figure 1c, however, clearly indicates that the limiting motions are no longer harmonic. At 300 K the test trajectory could be finished only for the time steps below 7.5 fs. The dashed prolongation of the 300 K trace shows results for trajectories lasting several picoseconds but interrupted by a rapid temperature growth. We attribute this

(42) Pearlman, D. A.; Case, D. A.; Caldwell, J. C.; Ross, W. S.; Cheatham, T. E., III; Ferguson, D. M.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 4.1*; University of California: San Francisco, CA, 1995.

(43) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535−542.

(44) Lin, D.; Matsumoto, A.; Gō, N. *J. Chem. Phys* **1997**, *107*, 3684−3690.

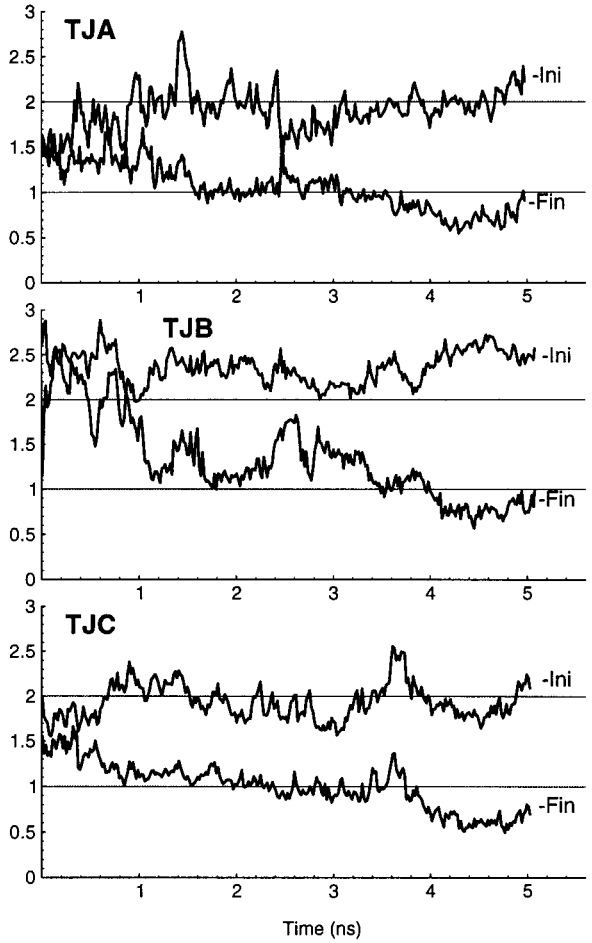

Figure 2. Time dependencies of rmsd of atom coordinates from initial conformations and from the final MD states in three different trajectories. The final states were obtained by averaging Cartesian coordinates of superimposed conformations taken with a 2.5-ps interval during the last nanosecond.

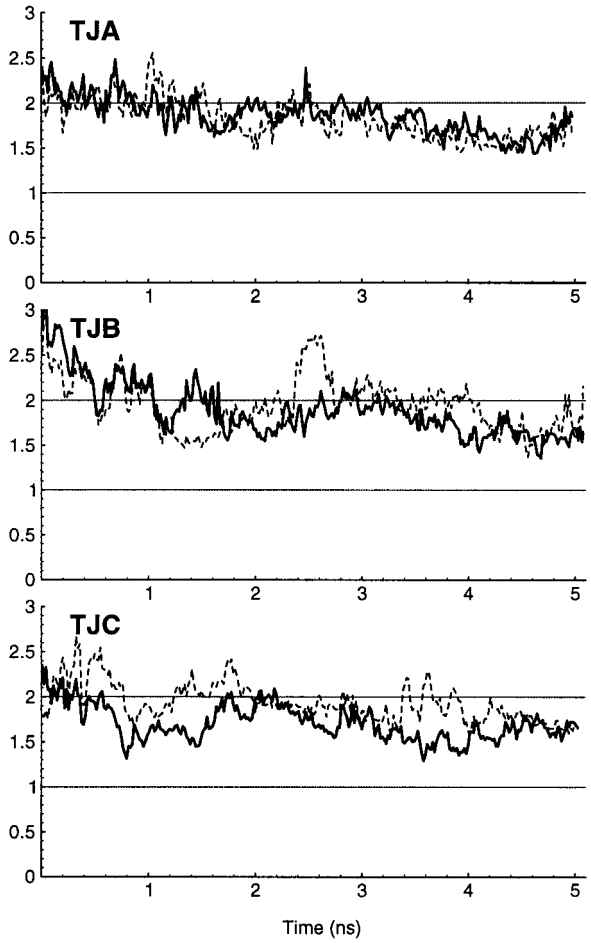

Figure 3. Time dependencies of rmsd of atom coordinates from the crystal conformation in three different trajectories. Dashed traces correspond to rmsd computed with the alternative assignment of strands in the experimental structure.[14]

effect to fast collisions of non-hydrogen atoms in sugar rings with their neighbors. With unfavorable collision angles, non-hydrogen ring atoms may have a considerably smaller effective inertia than similar atoms in already studied models where bond angles were fixed.[36]

Upon the basis of this interpretation, we additionally increase moments of ring C−C and C−O bonds by 15 and 5 amu·Å^2, respectively. This should make the inertias of all ring bonds similar and approximately equal to that of a water molecule, with a 50% increase for C−C bonds. Testing of the resulting model system is shown in Figure 1e,f. We see that the low-temperature trace is not significantly changed, but the normal temperature h_c value has increased to the desired 10-fs level. The upper border of the spectrum is also somewhat lowered, and there are much fewer overtone signals than in Figure 1d.

The two ICMD trajectories, TJA and TJB, were both computed with a step size of 10 fs in the conditions corresponding to Figure 1e,f. The Cartesian coordinate trajectory, TJC, was obtained with a step size of 2 fs. Time step limits in unconstrained models have been analyzed in our recent studies.[35,36] They are harmonic, and for h_c of 2 fs, the highest frequency in the system should be less than 3000 cm^{-1}. With AMBER94 parameters only two bond stretching modes in unconstrained dodecamer exceed this limit, namely, OH stretching of terminal hydroxyl groups and NH stretching in bases. These two frequencies have been lowered by reducing the corresponding force constants from 553 and 434 kcal/(mol·Å^2), respectively, to 400 kcal/(mol·Å^2). The terminal OH groups were discharged to prevent proton dissociation in hydrogen-bonded complexes. The possibility of the latter side effect is caused by the zero van der Waals radii of hydroxyl hydrogens in AMBER94 potentials.[5] With OH bond stretching constant reduced, the energy barrier for such transitions is also somewhat lowered.

Convergence and Stability of Trajectories

Figures 2−4 show the time evolution of rmsd of the duplex conformation from reference structures. Each point in these figures corresponds to a 15-ps interval and shows an averaged rmsd value. Figure 2 exhibits traces of the rmsd from the initial and final MD states. A slow drift of the structure is observed in all three trajectories and duration longer than 5 ns is apparently desirable to guarantee the ultimate stabilization. In TJA, for instance, the rmsd from the initial state grows almost steadily during the second half of the trajectory. The drop of rmsd from the final state during the last nanosecond is an expected effect, but during the first four nanoseconds, this value on average also decreases.

Figure 3 shows the traces of the rmsd from the crystal conformation. In the crystal,[14,15] the duplex molecules are linked in a head-to-tail manner and the two opposite ends are not symmetrical even though the sequence is. This asymmetry is significant, and by exchanging the assignment of strands, one gets two sometimes very different values of rmsd. Figure 3

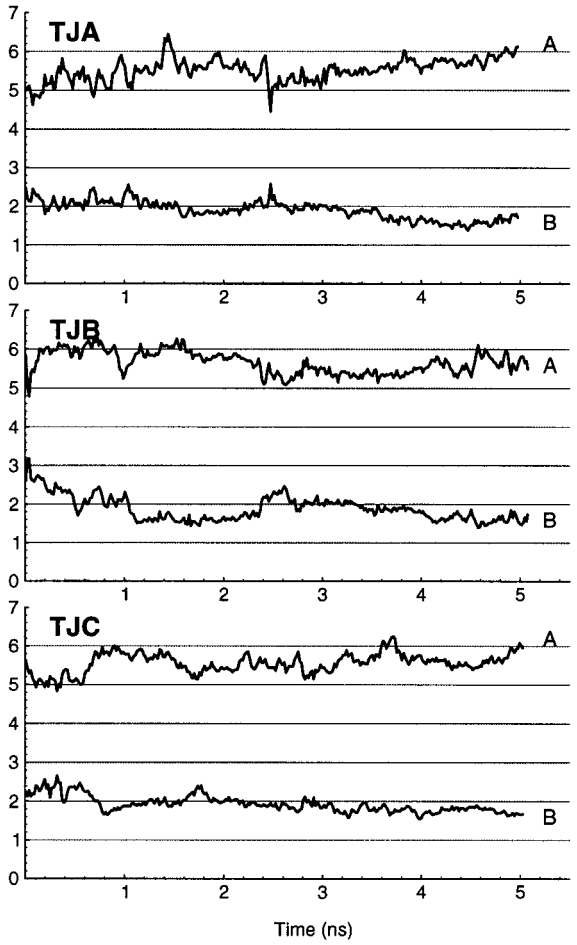

Figure 4. Time dependencies of rmsd of atom coordinates from canonical A- and B-DNA forms in three different trajectories. Canonical DNA forms[37] were constructed with NUCGEN procedure of AMBER.[42]

exhibits the traces of both, and they sometimes diverge considerably, indicating that, in dynamics, asymmetrical conformations sometimes appear spontaneously and can persist during nanosecond intervals. The most remarkable, however, is the fact that all three trajectories approach the crystal structure, with the final rmsd reaching the lowest values ever observed in free MD simulations. Similarly, the trace of rmsd from the canonical B-DNA also falls significantly below the 2-Å level, which is shown in Figure 4. At the same time, the rmsd from A-DNA grows and reaches the level corresponding to the difference between the canonical A and B forms.

The above results demonstrate that AMBER94 parameters with the semiexplicit solvent representation used here provide a perfectly stable B-DNA structure. A more detailed analysis shows that Watson−Crick base paring as well as stacking interactions are well maintained throughout the simulations. The water shell remained stable and attached to the duplex. Some solvent molecules diffused for more than 20 Å along the grooves, but only six (two in each trajectory) completely dissociated and moved away. The same number of "hot" molecules is a coincidence; no correlation in their initial positions has been found. Five of them have separated within the first 100 ps.

Figure 5 shows the time dependence of the total potential energy in the three trajectories. Similar to Figures 2−4 every point in Figure 5 corresponds to a 15-ps average. In TJA and TJC, the potential energy decreased by approximately 150 kcal/mol during the first nanosecond and then continued to decrease,

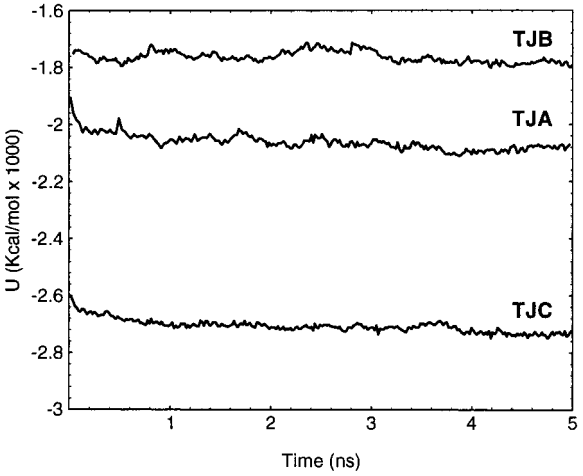

Figure 5. Time dependencies of the total potential energy in three different trajectories.

but slower. For TJB the decrease is less significant, but it is seen that the trace of the potential energy generally repeats those for the rmsd from the crystal structure and the canonical B form. Thus, it is seen that the final states in all three trajectories are favored by the potential energy, and the drift of the structure apparently results from slow annealing of the system in a better minimum.

We believe that the observed decrease in the potential energy should be mainly attributed to water reorganization in the minor groove and formation of a DNA−water interface. As we noted, the initial duplex conformation in TJA and TJC was first minimized without water starting from the canonical B form. Energy minimization studies of in these conditions[45] suggest that this starting point should be in the basin of the global minimum. The last structure in TJA has an rmsd of 2.8 Å from this minimum and cannot be driven back just by minimization because some phosphate groups are in B_{II} conformation. Reminimization without water reduces, however, rmsd to 1.6 Å, but the final energy is nearly 20 kcal/mol higher. Thus, the DNA structure is driven out from its own energy minimum closer to the experimental B-DNA conformations by the minor groove water, which plays a decisive structural and not just stabilizing role. In agreement with this interpretation, the net decrease of the potential energy in TJA and TJC is similar even though the potential energy trace of the latter is significantly shifted downward by additional nonbonded intrabase interactions.

Conformational Fluctuations

The range of conformational fluctuations can be estimated from the "Dials and Windows"[46] representation of TJA and TJC shown in Figure 6. Plots are shown for the same DNA strand, but these two duplex models have 646 and 2264 internal degrees of freedom, respectively. Despite this large difference the two sets of dials in Figure 6 are qualitatively indistinguishable. A similar degree of similarity is observed for fluctuations of helical parameters not shown here. Apparently, the ICMD model, which is allowed to move only along narrow paths in the full unconstrained configurational space, still keeps enough low-energy tracks to sample from the main areas defined by a given temperature of 300 K.

(45) Flatters, D.; Zakrzewska, K.; Lavery, R. *J. Comput. Chem.* **1997**, *18*, 1043−1055.

(46) Ravishanker, G.; Swaminathan, S.; Beveridge, D. L.; Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *6*, 669−699.

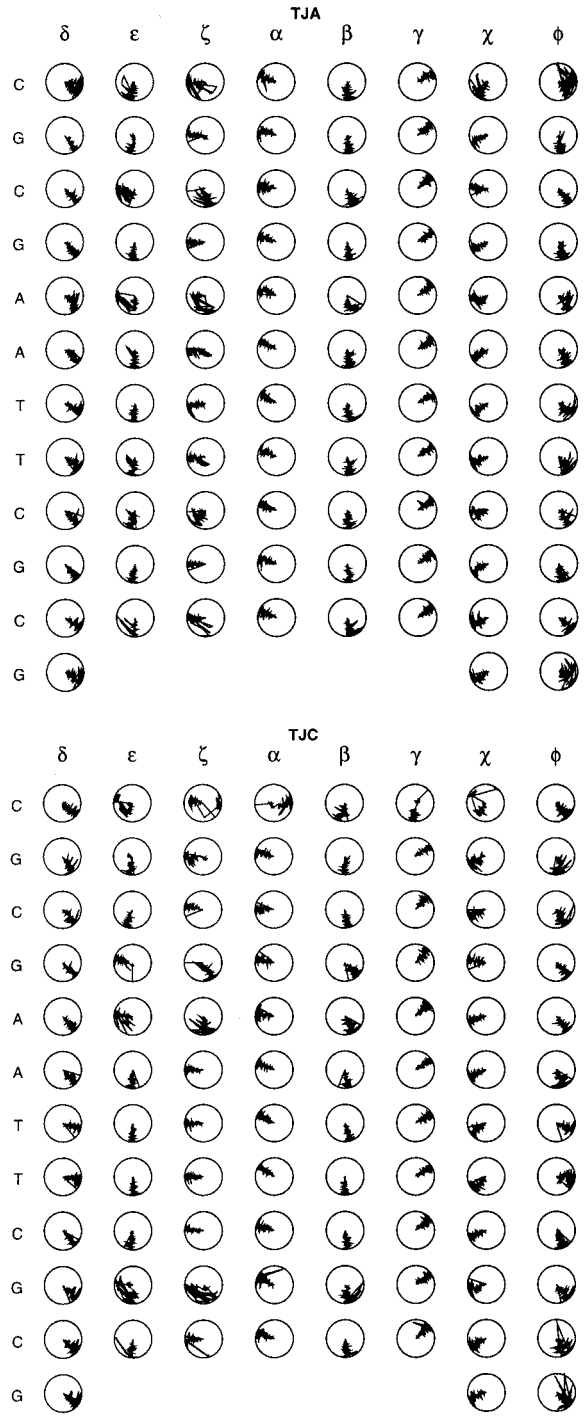

Figure 6. Conformational dials[46] for the same DNA strand in two trajectories. The radical coordinate is the time axis, with zero at the center and 5 ns at the circumference. The time interval between the points used for this analysis was 500 ps.

All conformational parameters fluctuate within the characteristic intervals of B-DNA structures.[47,48] Many distinct B_I-B_{II} transitions occur which are seen as correlated $(t,g^-) \leftrightarrow (g^-,t)$ jumps in ϵ and ζ dials. In the initial state of TJB, eight pairs of α and γ angles have nonstandard values corresponding to a "crankshaft" transition from the standard B-DNA backbone conformation. Only one of them has rotated back to the standard orientation, and this was the only such transition observed in the three trajectories. In PME calculations,[26] larger conformational activity is observed for α, β, γ, and χ angles, and therefore, one may note that dynamics in Figure 6 look somewhat "cooler". Figure 6 evidences, however, that any such difference cannot be attributed to additional constraints in ICMD but results from a different solvent treatment. A more detailed analysis of the differences between ICMD and Cartesian MD is generally interesting, but it is beyond this paper. The above results indicate that the differences are not readily apparent, and probably for many applications, they are not essential.

There are two physical factors in the present model that tend to reduce dynamic fluctuations with respect to PME calculations. First, the distance-dependent electrostatic function strengthens hydrogen bonds, and Guenot and Kollman[49] showed that this results in a 10% increase of the bulk water density for the TIP3P model. They observed reduction of fluctuations in analogous models of partially hydrated proteins but found that modification of atom charges to correct the energy of hydrogen bonding does not remove this effect and, therefore, it should rather be attributed to decreasing the long-range interactions.

We believe, however, that the specific contribution of the electrostatic treatment is minor here. Simulations with partial hydration, but full phosphate charges and $\epsilon = 1$ are also possible for double helices of this lengths. They generally give elongated conformations, but otherwise, the dynamics are qualitatively similar. Reduction of fluctuations should be expected when a complete water shell is replaced by a minor groove cloud because some deformations that need exchanging of hydrogen-bonding partners are penalized. This should reduce the range of sampled conformations and, accordingly, decrease the long time dynamic fluctuations.

On the other hand, comparison of thermal fluctuations of individual atoms suggests that the DNA molecule in our model is not much less flexible than in PME calculations. Figure 7 compares experimental crystallographic *B* factors with theoretical values computed from atom position fluctuations during the last nanosecond of TJA. Similar data have been reported by Duan et al.[25] for a PME simulation of the same molecule. This figure shows that there is a certain qualitative agreement with experimental data, notably *B* factors are systematically higher for the backbone atoms. Note that a 10-Å^2 level of *B* factors for central base pairs is two times lower than in experiments and that the computed fluctuations tend to increase toward the ends of the helix. Both these characteristic features have been observed in PME calculations.[25] The increase of fluctuations toward the ends is stronger here, and the overall pattern resembles that computed by normal-mode analysis,[44] which indicates that the slow harmonic modes are less damped in our model than in PME calculations.

Table 1 shows the averages of conventional DNA conformational parameters for the final MD states and the reference A- and B-DNA structures. Note that the values shown in brackets here and in tables below are sequence rather than time variances. It is seen that for TJA and TJC all values are close to the crystal structure and generally correspond to the "average DNA" as derived from NMR solution data.[47] For TJB, very large sequence variances of α, β, and γ angles are seen, which is due to the seven nonstandard local backbone conformations remaining from the initial state. Interestingly, a significant proportion of such groups has no sensible effect upon the rmsd of atom coordinates from the reference structures.

(47) Ulyanov, N. B.; James, T. L. *Methods Enzymol.* **1995**, *261*, 90−120.

(48) Hartmann, B.; Lavery, R. *Q. Rev. Biophys.* **1996**, *29*, 309−368.

(49) Guenot, J.; Kollman, P. A. *Protein Sci.* **1992**, *1*, 1185−1205.

Table 1. Average Conformational Parameters for Different Structures of d(CGCGAATTCGCG)$_2$[a]

	α	β	γ	δ	ϵ	ζ	χ	phase	amp
A	276(0)	208(0)	45(0)	84(0)	179(0)	311(0)	206(0)	13(0)	39(0)
B	313(0)	214(0)	36(0)	156(0)	155(0)	265(0)	262(0)	192(0)	36(0)
Xray	297(8)	172(14)	55(8)	123(18)	189(25)	254(34)	243(13)	130(25)	46(4)
TJA	291(6)	166(12)	54(4)	133(9)	200(27)	242(34)	248(9)	138(15)	39(5)
TJB	235(91)	181(29)	92(63)	140(9)	221(33)	226(50)	244(17)	142(12)	41(5)
TJC	291(11)	165(12)	64(48)	133(27)	204(35)	239(49)	247(15)	137(31)	39(12)

[a] All values are given in degrees.

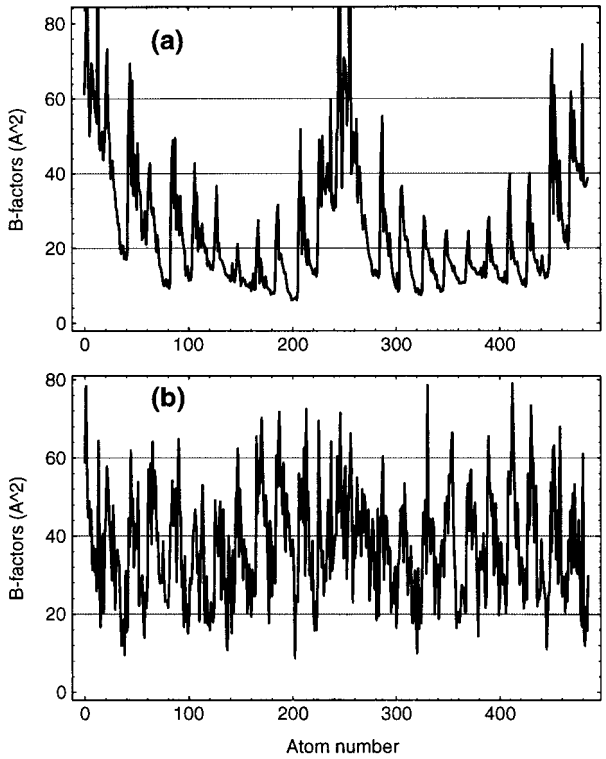

Figure 7. Comparison of calculated *B* factors (a) with experimental values[14] (b). The rms *B* factors were estimated from atom position fluctuations in the superimposed duplex conformations from the last nanosecond of TJA using $B = (8\pi^2/3)\langle\Delta r^2\rangle$. Backbone atoms were similarly grouped together in both figures.

Table 2. Nonhydrogen Atom rmsd (Å) between Different Structures for d(CGCGAATTCGCG)$_2$[a]

	A	B	X-ray[b]	X-ray[b]	TJA	TJB	TJC
A	0	6.22	6.02	6.02	5.77	5.64	5.63
B	5.46	0	1.35	1.35	1.45	1.44	1.65
X-ray	5.22	1.21	0	1.77	1.45	1.38	1.51
X-ray	5.22	1.21	1.41	0	1.45	1.49	1.62
TJA	5.14	1.14	1.28	1.17	0	1.20	1.13
TJB	4.93	1.40	1.24	1.34	1.03	0	1.22
TJC	5.02	1.44	1.27	1.38	1.04	1.14	0

[a] The upper and the lower triangles show results for all and for the middle 10 base pairs, respectively. [b] The two X-ray conformations correspond to the same structure with alternative assignments of palindromic strands.

Final MD States

The final states of the three trajectories are compared in terms of rmsd with each other and with the reference structures in Table 2. Four points should be noted here. *First*, the last six structures in this table form a group of B-DNA conformations that are approximately equally separated from A-DNA. *Second*, within this group, the three MD structures form a subfamily: they are closer to each other than to other B-DNAs. *Third*, the largest difference in the B-DNA family in Table 2 is found between the two crystal structures with inversely assigned strands. Thus, the crystal deformations, which may be considered as an upper estimate of the conformational variability in water, are larger than the difference between the experimental and calculated structures. *Finally*, the rmsd values between the computed and experimental structures are significantly lower than in all earlier free MD simulations.

Table 3. Global Base Pair: Axis Parameters

	Xdisp	Ydisp	inclin	tip
A	−5.43(0.00)	0.00(0.00)	19.12(0.01)	0.00(0.01)
B	−0.70(0.00)	0.00(0.00)	−5.98(0.02)	0.00(0.01)
X-ray	−0.54(0.24)	0.12(0.12)	0.19(3.52)	−0.41(2.70)
TJA	−1.10(0.24)	0.09(0.26)	2.06(2.91)	0.90(2.50)
TJB	−0.95(0.22)	−0.02(0.12)	0.62(3.52)	0.25(1.84)
TJC	−1.40(0.21)	0.10(0.14)	4.04(4.86)	0.69(1.80)
Δ_t	0.27	0.18	2.10	1.83

Tables 3−5 compare the final MD states and the reference structures in terms of the average helicoidal parameters. They all correspond to the characteristic B-DNA values.[48] The bottom lines show the time variances of averages during the last nanosecond of TJA. The time variances were similar for all three trajectories, and they give an appropriate measure for comparing computed and experimental averages in Tables 3−5. There are a few cases when the deviations are beyond this range. Namely, Xdisp in Table 3 is somewhat more negative, but still much less than in A-DNA. Stagger in Table 4 and slide in Table 5 are both slightly more negative. Also, the buckle and propeller angles computed for TJC slightly differ from the rest in Table 4. Computation of the last parameters, however, is sensitive to geometry of bases, and this difference should be rather attributed to strong deformations of bases in the averaged TJC structure. In all above examples, however, the computed and experimental values are almost within the range of dynamic fluctuations, and these "bad" cases rather emphasize a good general agreement.

Inter-base-pair parameters shown in Table 5 should be distinguished from the rest. These are "true" helical parameters in the sense that they characterize the symmetry transformation of a helical step. They are approximately additive, and therefore, are most critical for the size and the overall shape. For instance, the average rise roughly characterizes the total length, while the average twist, the net winding angle of the duplex. Note that the time variances in Table 5 are much smaller than the corresponding sequence variances, which suggests that, in dynamics, the corresponding local fluctuations are anti-correlated and tend to compensate each other. As should be expected for straight helices, all averages in Table 5 are close to zero except for twist and rise. Figure 8 shows the time dependencies of the last two averages in all three trajectories. Note that the ranges of fluctuations are similar and well characterized by the Δ_t values in Table 5. There is a distinguishable slow trend toward the final values, the latter being closer to experimental data than in the initial structures.

We noted already that the present model apparently exhibits a somewhat "cooler" dynamics for "crankshaft" backbone

Table 4. Global Base:Base Parameters

	shear	stretch	stagger	buckle	propel	opening
A	0.00(0.00)	−0.45(0.00)	0.19(0.00)	0.00(0.09)	13.70(0.02)	−4.62(0.02)
B	0.00(0.01)	0.03(0.01)	0.06(0.00)	0.00(0.07)	3.87(0.01)	−3.93(0.19)
X-ray	−0.04(0.28)	−0.07(0.09)	0.02(0.21)	0.31(5.88)	−13.45(6.77)	0.84(4.27)
TJA	0.00(0.21)	−0.06(0.03)	−0.13(0.13)	0.68(7.50)	−12.54(5.11)	−0.84(1.13)
TJB	0.02(0.15)	−0.07(0.03)	−0.11(0.20)	−0.79(9.59)	−12.80(4.45)	−0.79(1.45)
TJC	0.05(0.32)	0.03(0.13)	−0.14(0.32)	3.01(15.03)	−16.30(5.77)	1.07(2.60)
Δ_t	0.06	0.03	0.08	2.48	1.42	0.91

Table 5. Global Inter Base Pair Parameters

	shift	slide	rise	tilt	roll	twist
A	0.00(0.00)	0.00(0.00)	2.56(0.00)	0.00(0.02)	0.00(0.04)	32.70(0.02)
B	0.00(0.01)	0.00(0.00)	3.38(0.01)	0.00(0.01)	0.00(0.04)	36.01(0.12)
X-ray	0.05(0.50)	0.00(0.30)	3.37(0.15)	−0.36(2.67)	−0.40(5.48)	35.89(3.82)
TJA	0.07(0.41)	−0.09(0.45)	3.31(0.22)	0.56(2.99)	1.19(5.94)	35.67(5.25)
TJB	−0.01(0.78)	−0.01(0.36)	3.37(0.23)	0.31(5.72)	−0.14(4.86)	36.36(2.32)
TJC	0.01(0.52)	−0.04(0.32)	3.36(0.26)	−0.29(3.56)	0.74(6.71)	34.69(6.95)
Δ_t	0.04	0.03	0.05	0.78	0.83	0.52

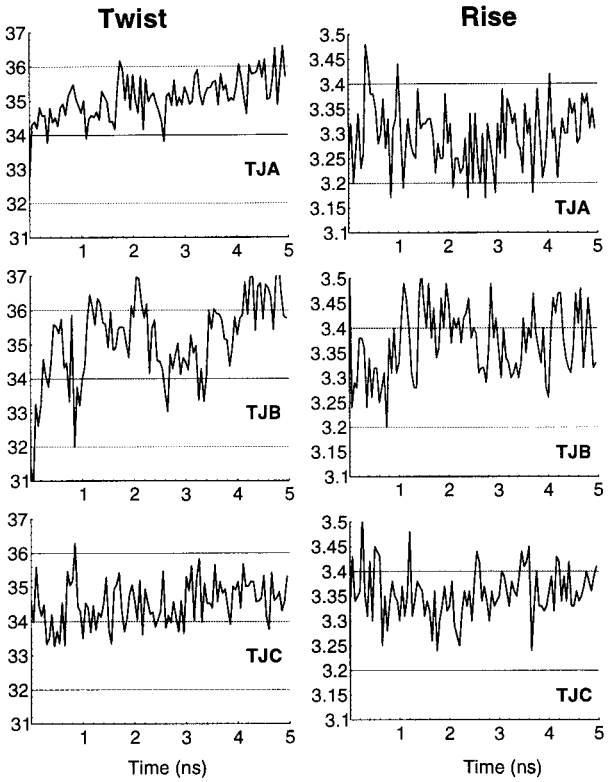

Figure 8. Time dependencies of the average twist and rise in three different trajectories. Helicoidal parameters were computed with Curves program.[41] The time interval between the points used for this analysis was 500 ps. No smoothing was applied.

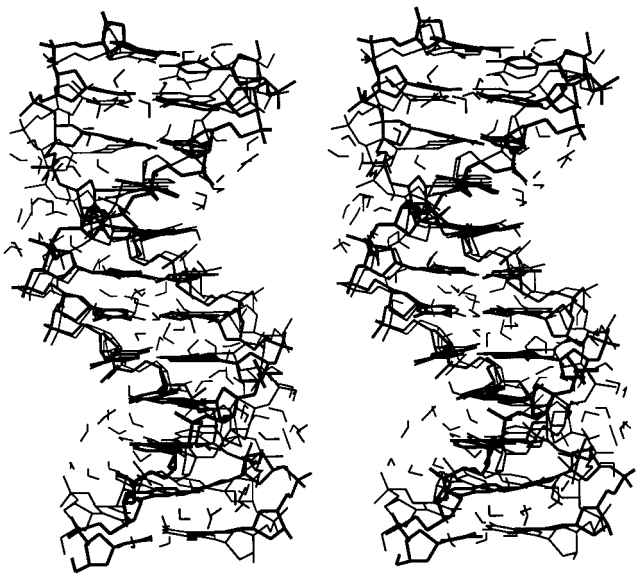

Figure 9. Superposition of the crystal conformation with the average structure from the final TJA state. The latter have been obtained by fitting the standard geometry model to averaged Cartesian coordinates by minimization with distance restraints (the final rmsd is 0.04 Å). The water molecules represent a snapshot from the last nanosecond of the same trajectory.

transitions. The same effect is suggested by the generally small Δ_t values in Tables 3−5, with the time variances of global inter-base-pair parameters 10 times smaller than in the recent PME calculations.[50] The last paper is the only reported PME simulation where the time variances were analyzed, but unfortunately, it considered a shorter DNA fragment with a different sequence. Both the length and the sequence of the DNA fragment may affect the amplitudes of fluctuations.

Figure 9 shows superposition of the crystal conformation with the TJA final state. Note the absence of the growing deviation of base pairs toward the ends of the duplex, which is very clear in similar superposition figures from PME calculations.[22,25] The correspondence is high both near the ends and in the central part, where one can notice a characteristic narrowing of the minor groove. Additional studies are necessary, however, to make sure that this is a reproducible sequence effect. In the TJC and TJB final states, for instance, the minor groove width was similarly small in the center, but it widened toward one end only for TJC and was evenly narrow for TJB.

Comparison with Earlier Studies

Many attempts to reproduce the structure of the same duplex in MD have been reported,[1,16−24] including several recent long-time simulations[25−29] with AMBER94 parameters and PME technique. These last studies gave rather consistent results, with stable B-DNA conformations relatively close to experimental data. The final MD states obtained here can be compared with the earlier results in terms of average conformational and helicoidal parameters, as well as in terms of rmsd from reference

(50) Cheatham, T. E., III; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 4805−4825.

Table 6. Nonhydrogen Atom rmsd from the Crystal Form Computed for Various Fragments of d(CGCGAATTCGCG)$_2$

segment[a]	X-ray[b]	TJA	TJB	TJC	Duan[c]	Young[d]	Ini[e]
1−12	1.77	1.45	1.38	1.51	2.4	2.5	2.30
1−3	0.62	1.06	1.06	1.41	1.53		1.17
2−4	0.73	1.28	1.03	1.30	1.52		1.05
3−5	0.57	1.09	1.01	1.04	1.43		0.97
4−6	0.53	0.81	0.85	0.60	1.01		0.56
5−7	0.27	0.59	0.71	0.43	0.81		0.50
6−8	0.27	0.42	0.61	0.41	0.74		0.54
7−9	0.53	0.60	0.83	0.69	1.00		1.08
8−10	0.57	0.83	1.08	0.88	1.29		1.44
9−11	0.73	0.95	1.24	0.95	1.51		1.09
10−12	0.62	0.97	0.95	0.93	1.52		0.81
1−4	0.74	1.23	1.10	1.41	1.71		1.27
9−12	0.74	1.02	1.22	0.99	1.70		1.19
5−8	0.27	0.62	0.73	0.46	0.82		0.71
4−9	0.61	0.77	0.87	0.72	1.13	1.46	1.08
3−10	0.80	0.96	1.02	1.04			1.57
2−11	1.41	1.28	1.25	1.28			1.96
1−6	0.99	1.45	1.11	1.28		1.26	1.22
7−12	0.99	0.96	1.14	1.15		1.87	1.85

[a] Duplex segments are labeled by residue numbers in the first strand of the pdb file of the crystal structure. [b] The crystal conformation[14] is compared with itself by exchanging the assignment of the two palindromic strands. [c] Data from ref 25 for the average conformation from the last nanosecond of a 2 ns trajectory. [d] Data from ref 26 for the average conformation from a 5 ns trajectory. A 5 ns averaging for TJA and TJB gives for all residues 1.62 and 1.45 Å, respectively. [e] The starting conformation of TJA.

structures. A close look at these data reveals significant discrepancies.

All published free MD simulations with the PME method and AMBER94 parameters[25,26,28] converged to average conformations with rmsd around 2.5 Å from canonical and crystal B forms, which is significantly larger than the corresponding values in Table 2. Much smaller values reported in ref 22 are not appropriate for such comparison because they were obtained in the crystal cell environment which is very restrictive. For a more detailed and direct comparison we show in Table 6 rmsd's computed for shorter duplex fragments, which were used by other authors.[25,26]

This table shows that, when computed for shorter fragments, the rmsd's of the final TJA, TJB, and TJC states from the crystal conformation are not very different from earlier results. Generally, the shorter the fragment and the closer it is to the duplex center, the smaller the difference. A part of this effect is due to local deformations of chemical groups in our averaged MD conformations which have not been corrected. Nevertheless, this comparison clearly shows that the lower rmsd values for the whole molecule result from better relative orientations of distant duplex fragments or, in other words, from a better overall shape. This improvement has been achieved in the course of dynamics. To show this we have included in this table the results of a similar comparison for the initial conformation from TJA, that is before equilibration. It has a 2.3-Å rmsd from the crystal form but, for some fragments, shows even better local fitting than the final TJA state.

Having arrived at this conclusion one reasonably asks how this improvement reflects on helical parameters? Comparison of Tables 3−5 with the literature data shows that they are generally similar with one notable exception, namely, the average twist. In earlier free MD simulations this value was always smaller, less than even in the canonical A-DNA.[25,26,28] The deviations of atom coordinates caused by the average twist grow with the helix lengths, which explains well the data in Table 6. A 3° difference, for example, would result in an approximately 33° smaller total winding angle of the duplex, which already gives deviations around 3 Å for terminal base pairs in superimposed structures.

Interestingly, the early PME simulations made in the crystal environment[22] gave an rmsd around 1.2 Å for the cell averaged structure and a much better average twist. This large difference seems strange because since then the atom parameters for nucleic acids have been significantly improved.[5] The detailed DNA conformations observed in crystals and in solution can certainly differ because of numerous physical factors, but the average helical parameters of B-DNA, notably twist and rise, are rather similar in different environments.[51,48,47] Note, however, that with a head-to-tail junction of DNA fragments in a crystal cell, periodic boundary conditions effectively impose restraints upon the total winding angle of the duplex. Unwinding of a single double helix requires opposite rotations at the two ends, which, with the specific crystal packing of this dodecamer,[14,15] would result in atom−atom clashes. To reduce the average twist the whole chain of linked duplexes must rotate concertedly, but this is not compatible with periodical boundaries. It is noteworthy, however, that at the same time the average rise quickly dropped from the initial 3.4 to 3.2 Å, which corresponds to a 2-Å shortening of the duplex.[22]

These observations may have alternative interpretations, but they all are explained once we assume that the crystal conformation of d(CGCGAATTCGCG)$_2$ in PME calculations suffers from a small internal strain which tends to relax one way or another. It should be noted also that, for other B-DNA duplexes, similar calculations give even lower average twist values.[50,52] This discrepancy has been noticed and discussed by different authors.[26,28,52] It is much larger than the experimental error in the corresponding X-ray values, and it can hardly be attributed to dynamic fluctuations because the computed values fall into a rather narrow interval regardless of the sequence. Our results may help to understand the origin of this discrepancy because they give an example where the effect disappears.

Concluding Discussion

The results described above can be summarized in three statements. *First*, we have presented ICMD as a new powerful method in simulation nucleic acids. The fact that one of the production runs was made with Cartesian coordinate MD should not mislead. This was the last trajectory we computed, and it took approximately 5 times more time than the other two. If we were bound to repeat similarly all the trial calculations involved in this study, the time spent for this project would increase very significantly.

Second, we have shown that large water boxes, explicit counterions, and PME calculations are not absolutely necessary in MD simulations of B-DNA duplexes. Surprisingly accurate and stable structures can be obtained in inexpensive free MD simulations with AMBER94[5] parameters and semiimplicit solvent representation. This is encouraging since it opens the way to MD simulations of much longer DNA molecules and other important biological systems where more rigorous models would be prohibitively expensive.

Third, we have found, unexpectedly, that MD trajectories of a partially hydrated duplex with a very approximate treatment of long-range electrostatic effects converge to conformations which are significantly closer to experimental data than in PME calculations. This certainly does not mean that the simplified

(51) Dickerson, R. E. *Methods Enzymol.* **1992**, *211*, 67−111.
(52) Cheatham, T. E., III; Kollman, P. A. *J. Mol. Biol.* **1996**, *259*, 434−444.

treatment used here is in fact the most accurate. Rather, our results indicate that the converged MD conformations are determined by local interactions in DNA, in minor groove water and at their interface, and therefore, the high precision of the resulting structures should be totally attributed to the high quality of AMBER94[5] parameters. The long-range electrostatic interactions are also important but less specific, and probably any technique that scales them in a generally similar way should work equally well.

We believe that modeling of nucleic acids with partial hydration and implicit approximate treatment of long-range electrostatics should not be considered obsolete. On the contrary, this approximation is reasonable and it can significantly expand the scope of application of the modern computational techniques. For many applications it is only important that the model would meet two requirements. Namely, the experimental conformations must be within the allowed area in the conformational space, close to the free energy minimum. In addition, the model must allow MD trajectories to sample from the surrounding conformational space. Our results show that both of these requirements can be satisfied. The sampling is possibly reduced with respect to the full hydration, but it is still sufficient to approach the minimum energy conformations from rather distant initial states. Our results give no indication that, because of the reduced fluctuations, the system may be frozen near the starting point. Figures 2−4 and 8 demonstrate that the drift of different structural parameters significantly exceeds the noise level and actually is similar or larger than in the PME calculations starting from the same initial conformations.[25,26]

The difference from the PME results is intriguing, and hopefully, it should help to further improve our understanding of internal DNA mechanics and the specific role of the electrostatic effects. The simplified computations used here certainly cannot be considered as an alternative to the PME method and similar techniques which are potentially able to treat the effects of water and counterion environment upon DNA with all details. Our results also do not mean that there are some inherent difficulties in the PME method. A better agreement with the experimental data can result from fortunate mutually compensating artifacts. One of the possible such interpretations is as follows.

In our calculations, a large water cloud remained always "docked" in the minor DNA groove and caused its narrowing with respect to the vacuum minimum energy conformation. This effect may be interpreted as a nonspecific negative pressure or a capillary force. It has been recently suggested that, in reality, the minor groove narrowing is caused by specific binding of small counterions.[27] The recent high-resolution crystallographic studies[53] showed that four water sites in the minor groove spine may be partially occupied by sodium ions. In our model, water is held near the solute by long-range charge−dipole interactions; it is pulled into the minor groove, where the field is particularly low[54,55] and helps to relax the local "electric strain", thus playing in a vacuum the same role as counterions play in water. With Na^+ concentration around 0.3 M used in simulations,[27] the duration of trajectories accessible with PME method are just enough to simulate the condensation if it is controlled by diffusion. If an entropic or enthalpic barrier is involved, or if the process goes via several stages, the necessary time can well be a couple of orders longer. One may note in this connection that the trajectories described here qualitatively differ from the published PME simulations by the slow drift of structures observed in the nanosecond time scale after the first rapid relaxation phase. Until now, only a single relaxation phase was always observed in PME simulations, which may suggest that the later phases are much slower.

This explanation, however, admits that a very approximate treatment of electrostatics added by equally approximate and nonspecific modeling of counterion condensation in the minor groove is able to produce high-precision B-DNA conformations. The important question to answer in this connection is whether the water effect in our model is sequence specific or it just nonspecifically narrows the minor groove. In any case, the level of agreement with experimental data achieved here sets up a new benchmark for future comparisons. Whatever the cause of a good agreement it is always useful since it allows one to ask sensible questions. The models and methods described here should help to rapidly answer some of them.

Acknowledgment. I thank R. Lavery and B. Hartmann for useful discussions and critical comments to the first version of the manuscript.

JA981498W

(53) Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. *Biochemistry* **1998**, *37*, 8341−8355.
(54) Klein, B. J.; Pack, G. R. *Biopolymers* **1983**, *22*, 2331−2352.
(55) Jayaram, B.; Sharp, K. A.; Honig, B. *Biopolymers* **1989**, *28*, 975−993.