J|A|C|S
ARTICLES
Published on Web 11/14/2002

DNA Dynamics in a Water Drop without Counterions

Alexey K. Mazur*

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

Received December 14, 2001. Revised Manuscript Received September 23, 2002

Abstract: Because of its polyionic character, the DNA double helix is stable and biologically active only in salty aqueous media where its charge is compensated by solvent counterions. Monovalent metal ions are ubiquitous in DNA environment, and they are usually considered as the possible driving force of sequence-dependent modulations of DNA structure that make it recognizable by proteins. In an effort to directly examine this hypothesis, MD simulations of DNA in a water drop surrounded by vacuum were carried out, which relieves the requirement of charge neutrality. Surprisingly, with zero concentration of counterions, a dodecamer DNA duplex appears metastable, and its structure remains similar to that observed in experiment, including the minor groove narrowing in the dodecamer d(CGCGAATTCGCG)$_2$ often considered as the most evident cation effect. It is suggested that the same computational approach will allow one to simulate dynamics of long DNA chains more efficiently than with periodical boundary conditions.

Introduction

Environment effects upon the structure and dynamics of nucleic acids is of fundamental importance for their biological function. It has been long recognized that, because of the polyionic nature of DNA, solvent counterions are required for its stability and that important structural changes in DNA can be provoked by changing the concentration, the charge, or the type of counterions.[1−3] More recently, it has been proposed that direct contacts with free ions can cause significant DNA deformations.[4−8] Most controversial is the role of the common monovalent cations Na^+ and K^+. They are ubiquitous in the DNA environment and can be readily available for any purpose. Until recently, they remained invisible in experimental DNA structures because it is difficult to detect them in water, and it has been suggested that they are perhaps responsible for the most widespread deformations of the double helix, narrowing of the minor groove and bending. It is assumed that counterions are sequestered in the minor groove of some sequences, which breaks the symmetry of the repulsive electrostatic forces and provokes deformations. This model is general, and it easily explains other puzzling effects in DNA structure.

The foregoing hypothesis is supported by many observations. Penetration of monovalent cations into the minor DNA groove has been confirmed by X-ray diffraction,[9−17] NMR spectroscopy,[18,19] and MD simulations.[5,20−26] It appears difficult, however, to find a discriminating setup for testing the cause and consequence relationship between solvent ions and the fine DNA structure. All available experimental and computational evidences have more than one interpretation, making this problem highly controversial.[22,27,28] For example, correlations between ion positions and DNA structure observed in MD

* To whom correspondence should be addressed. E-mail: alexey@ibpc.fr. Fax: +33[0]1 58 41 50 26.

(1) Anderson, P.; Bauer, W. *Biochemistry* **1978**, *17*, 594−601.
(2) Nishimura, Y.; Torigoe, C.; Tsuboi, M. *Nucleic Acids Res.* **1986**, *14*, 2737−2749.
(3) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
(4) Jayaram, B.; Beveridge, D. L. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 367−394.
(5) Young, M. A.; Jayaram, B.; Beveridge, D. L. *J. Am. Chem. Soc.* **1997**, *119*, 59−69.
(6) Hud, N. V.; Feigon, J. *J. Am. Chem. Soc.* **1997**, *119*, 5756−5757.
(7) Rouzina, I.; Bloomfield, V. A. *Biophys. J.* **1998**, *74*, 3152−3164.
(8) Williams, L. D.; Maher, L. J., III. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 497−521.
(9) Rosenberg, J. M.; Seeman, N. C.; Kim, J. J. P.; Suddath, F. L.; Nicholas, H. B.; Rich, A. *Nature* **1973**, *243*, 150−154.
(10) Bartenev, V. N.; Golovanov, E. I.; Kapitonova, K. A.; Mokulskii, M. A.; Volkova, L. I.; Skuratovskii, I. Y. *J. Mol. Biol.* **1983**, *169*, 217−234.
(11) Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. *Biochemistry* **1998**, *37*, 8341−8355.
(12) Shui, X.; Sines, C. C.; McFail-Isom, L.; VanDerveer, D.; Williams, L. D. *Biochemistry* **1998**, *37*, 16877−16887.
(13) Tereshko, V.; Minasov, G.; Egli, M. *J. Am. Chem. Soc.* **1999**, *121*, 470−471.
(14) Tereshko, V.; Minasov, G.; Egli, M. *J. Am. Chem. Soc.* **1999**, *121*, 3590−3595.
(15) Woods, K. K.; McFail-Isom, L.; Sines, C. C.; Howerton, S. B.; Stephens, R. K.; Williams, L. D. *J. Am. Chem. Soc.* **2000**, *122*, 1546−1547.
(16) Sines, C. C.; McFail-Isom, L.; Howerton, S. B.; VanDerveer, D.; Williams, L. D. *J. Am. Chem. Soc.* **2000**, *122*, 11048−11056.
(17) Howerton, S. B.; Sines, C. C.; VanDerveer, D.; Williams, L. D. *Biochemistry* **2001**, *40*, 10023−10031.
(18) Hud, N. V.; Sklenár, V.; Feigon, J. *J. Mol. Biol.* **1999**, *286*, 651−660.
(19) Denisov, V. P.; Halle, B. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 629−633.
(20) Young, M. A.; Ravishanker, G.; Beveridge, D. L. *Biophys. J.* **1997**, *73*, 2313−2336.
(21) Feig, M.; Pettitt, B. M. *Biophys. J.* **1999**, *77*, 1769−1781.
(22) McConnell, K. J.; Beverdige, D. L. *J. Mol. Biol.* **2000**, *304*, 803−820.
(23) Štrahs, D.; Schlick, T. *J. Mol. Biol.* **2000**, *301*, 643−663.
(24) Štefl, R.; Koča, J. *J. Am. Chem. Soc.* **2000**, *122*, 5025−5033.
(25) Hamelberg, D.; McFail-Isom, L.; Williams, L. D.; Wilson, W. D. *J. Am. Chem. Soc.* **2000**, *122*, 10513−10520.
(26) Hamelberg, D.; Williams, L. D.; Wilson, W. D. *J. Am. Chem. Soc.* **2001**, *123*, 7745−7755.
(27) McFail-Isom, L.; Sines, C. C.; Williams, L. D. *Curr. Opin. Struct. Biol.* **1999**, *9*, 298−304.
(28) Chiu, T. K.; Zaczor-Grzeskowiak, M.; Dickerson, R. E. *J. Mol. Biol.* **1999**, *292*, 589−608.

simulations cannot answer whether the ions perturb DNA or just bind "opportunistically" in the sites of low potential near the already deformed double helix.[25,26] To clarify the issue of cause and effect, one would have to remove solvent ions and check if the supposed counterion effects disappear with them. Unfortunately, the most reliable computational procedures presently employed require that the simulation cell that holds DNA carries zero net charge; therefore, the counterion effects cannot be completely eliminated.

The requirement of charge neutrality can be relieved in computations without cutoffs if periodical boundary conditions are effectively abolished. At present, there is no way to do this for continuous media. In contrast, for an isolated system surrounded by vacuum, charge neutrality is not required, and it is technically possible to compute energy and forces without cutoff. One can try, therefore, to treat the DNA molecule covered by several water layers as a large isolated cluster. These conditions can also present more general interest for future simulations of long DNA chains. Long DNA cannot be simulated in an elongated periodical unit cell because it can easily break periodical boundaries due to rotation and bending.[29] Consequently, the cell must grow in all directions, which makes calculations prohibitively costly. Alternative approaches propose to treat the solvent implicitly[30] or semiimplicitly.[31] The cluster conditions may present another appealing possibility if it turns out that the necessary number of water molecules need not grow as the cube of the DNA length.

To implement the foregoing plan, I have adapted the particle mesh Ewald algorithm[32,33] for modeling dynamics of DNA in a water drop surrounded by vacuum. Similar calculations are long known in physics. Hockney was apparently the first to show how the Poisson's problem for isolated clusters can be treated with Fourier-based mesh-particle methods.[34,35] He showed that one can compute the *exact* value of the electrostatic potential within a unit cell surrounded by vacuum by (i) replicating the system periodically in space so that neighboring images are separated by a distance equal to the box length (L), and (ii) truncating the Green function at $2L$ while continuing it periodically. A detailed analysis of his original ideas and some improvements have been reported by Eastwood and Brownrigg,[36] as well as by other groups.[37,38] This approach was employed in cosmic physics and sometimes mentioned in other domains.[39,40] Martyna and Tuckerman have recently developed a similar idea in the context of most popular chemical physics applications.[41] To my knowledge, however, such methods were never applied to DNA or other biological systems.

In this paper, the original schema by Hockney is slightly modified to make it more suitable for DNA, and the method is applied for MD simulations of a dodecamer B-DNA duplex in water, with free vacuum boundaries and unperturbed Coulomb electrostatics. A detailed comparison is presented of DNA dynamics in different conditions including zero counterion concentrations as well as conventional simulations with periodical boundary conditions. The results obtained in all of these simulations appear very similar regarding the fine DNA structure. This confirms earlier conclusions that the possible effects of periodical DNA images in simulations with Fourier-based mesh-particle algorithms usually are not very significant.[29,42] Surprisingly, it appears to be the case for the effects of monovalent solvent counterions as well. With their concentration equal to zero, DNA dynamics virtually do not change, and the fine DNA structure remains similar to that observed in other calculations and experiments.

Methods

We will compute electrostatic energy and forces for a large isolated system by taking advantage of fast Fourier-based lattice summation approaches. Similar to earlier such methods,[34,35,41] the system in question is surrounded by a large enough empty zone and next continued in space periodically. Calculations employ the Fourier transform techniques, which means that they are carried out simultaneously for all space. At the same time, the interaction potential is modified so that the contributions of neighboring images cancel out, while all interactions within a single image remain exact. Considered below is an easy way to achieve that by shifting Coulomb's potential.

Ewald Splitting for Shifted Coulomb Potential. First consider a single charge placed at the origin of coordinates. The shifted Coulomb potential is

$$\varphi(\mathbf{r}) = qg(\mathbf{r})$$

where

$$g(\mathbf{r}) = \begin{cases} \dfrac{1}{r} - \dfrac{1}{R_0} & r \le R_0 \\ 0 & r > R_0 \end{cases} \tag{1}$$

Here and below we distinguish between vectors and their modules by boldface and normal symbols, respectively, and always assume dot product when two vectors are multiplied. The standard Ewald splitting is applied as

$$\varphi = \varphi_{\mathrm{d}} + \varphi_{\mathrm{r}} \tag{2}$$

where indices stand for "direct" and "reciprocal", with

$$\varphi_{\mathrm{d}}(\mathbf{r}) = q\int [\delta(\mathbf{t}) - G(\beta,\mathbf{t})]g(\mathbf{r}-\mathbf{t})\,\mathrm{d}\mathbf{t} \tag{3}$$

and

$$\varphi_{\mathrm{r}}(\mathbf{r}) = q\int G(\beta,\mathbf{t})g(\mathbf{r}-\mathbf{t})\,\mathrm{d}\mathbf{t}$$

while $G(.,.)$ denotes the Gaussian

$$G(\beta,\mathbf{t}) = \left(\frac{\beta^2}{\pi}\right)^{3/2} \mathrm{e}^{-\beta^2 t^2}$$

(29) Cheatham, T. E., III; Kollman, P. A. In *Interactions and Expression of Biological Macromolecules. Proceedings of the 10th Conversation, State University of New York, Albany, N. Y. 1998*; Sarma, R. H., Sarma, M. H., Eds.; Adenine Press: New York, 1998; pp 99−116.
(30) Tsui, V.; Case, D. A. *J. Am. Chem. Soc.* **2000**, *122*, 2489−2498.
(31) Mazur, A. K. *J. Am. Chem. Soc.* **1998**, *120*, 10928−10937.
(32) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(33) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(34) Hockney, R. W. *Methods Comput. Phys.* **1970**, *9*, 135−211.
(35) Hockney, R. W.; Eastwood, J. W. *Computer Simulation Using Particles*; McGraw-Hill: New York, 1981.
(36) Eastwood, J. W.; Brownrigg, D. R. K. *J. Comput. Phys.* **1979**, *32*, 24−38.
(37) Maruhn, J. A.; Welton, T. A.; Wong, C. Y. *J. Comput. Phys.* **1976**, *20*, 326−335.
(38) James, R. A. *J. Comput. Phys.* **1977**, *25*, 71−93.
(39) Pollock, E. L.; Glosli, J. *Comput. Phys. Commun.* **1996**, *95*, 93−110.
(40) Sagui, C.; Darden, T. A. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155−179.
(41) Martyna, G. J.; Tuckerman, M. E. *J. Chem. Phys.* **1999**, *110*, 2810−2821.
(42) Norberto de Souza, O.; Ornstein, R. L. *Biophys. J.* **1997**, *72*, 2395−2397.

Substitution of eq 1 into eq 3 gives

$$\varphi_d(\mathbf{r}) = q\int_{|\mathbf{t}-\mathbf{r}|<R_0}[\delta(\mathbf{t}) - G(\beta,\mathbf{t})]\frac{d\mathbf{t}}{|\mathbf{r}-\mathbf{t}|} - q\int_{|\mathbf{t}-\mathbf{r}|<R_0}[\delta(\mathbf{t}) - G(\beta,\mathbf{t})]\frac{d\mathbf{t}}{R_0} \quad (4)$$

We need $\varphi_d(\mathbf{r})$ only for $r < r_{cut} \approx 10$ Å, with parameter β chosen so that $\exp(-\beta^2 r_{cut}^2) \ll 1$. Therefore, with $R_0 > 2r_{cut}$, we can safely extend integration in eq 4 to infinity, which zeroes the second integral. The result becomes identical to that in the Ewald method:

$$\varphi_d(\mathbf{r}) = \frac{q}{r}erfc(\beta r) \quad (5)$$

Now consider N charged particles in a rectangular unit cell replicated periodically in space. The total potential $\Phi(\mathbf{r})$ is split similarly to eq 2

$$\Phi = \Phi_d + \Phi_r \quad (6)$$

Term Φ_d is obtained by summing eq 5 over all charge pairs separated by less than r_{cut}, which does not differ from conventional Ewald calculations. For Φ_r, we have

$$\Phi_r(\mathbf{r}) = \int_t \sum_{n=0}^{\infty}\sum_{j=1}^{N} q_j G(\beta,\mathbf{t} - \mathbf{r}_j^n) g(\mathbf{r}-\mathbf{t})\, d\mathbf{t} \quad (7)$$

where n is the cell number, and integration involves the whole space. Note that because of the shifting in eq 1, all lattice sums involved in eqs 6 and 7 are absolutely convergent regardless of charges, and the result does not depend on the order of summation. Therefore, the straightforward calculation below is also perfectly rigorous. The system is invariant with respect to periodical translations. Because the solution is unique, we can look for a periodical potential $\Phi_r(\mathbf{r})$, with its Fourier coefficients computed as

$$\hat{\Phi}_r(\mathbf{k}) = \int_{\mathbf{r}}\int_t \sum_{n=0}^{\infty}\sum_{j=1}^{N} q_j G(\beta,\mathbf{t} - \mathbf{r}_j^n) g(\mathbf{r}-\mathbf{t}) e^{-i\mathbf{kr}}\, d\mathbf{t}\, d\mathbf{r}$$

where integration over $\mathbf{r}$ involves only the central cell. Now summation over n can be dropped with integration over $\mathbf{r}$ extended to infinity. This standard manipulation only requires that $g(\mathbf{r})$ does not depend on the vector direction. Integration over $\mathbf{t}$ and $\mathbf{r} - \mathbf{t}$ gives Fourier transforms of the Gaussian and $g(\mathbf{r})$, respectively, and results in

$$\hat{\Phi}_r(\mathbf{k}) = \sum_{j=1}^{N} q_j e^{-i\mathbf{kr}_j} e^{-\mathbf{k}^2/4\beta^2}\hat{g}(\mathbf{k})$$

with

$$\hat{g}(\mathbf{k}) = \begin{cases} \dfrac{4\pi}{k^2}\left(1 - \dfrac{\sin(kR_0)}{kR_0}\right) & k \neq 0 \\ 2\pi R_0^2/3 & k = 0 \end{cases}$$

The corresponding energy term is

$$E_r = \frac{1}{2}\sum_{j=1}^{N} q_j \Phi_r(\mathbf{r}_j) = \frac{1}{2}\sum_{j=1}^{N} q_j \sum_{m=-\infty}^{\infty}\frac{1}{V}\hat{\Phi}(\mathbf{k}_m)e^{i\mathbf{k}_m\mathbf{r}_j} = \frac{1}{2V}\sum_{m=-\infty}^{\infty} e^{-\mathbf{k}_m^2/4\beta^2} S(\mathbf{k}_m)S(-\mathbf{k}_m)\hat{g}(\mathbf{k}_m) \quad (8)$$

with the structure factor $S(\mathbf{k})$ defined as

$$S(\mathbf{k}) = \sum_{j=1}^{N} q_j e^{i\mathbf{kr}_j}$$

Equation 8 is identical to the corresponding relation in ref 32 with the Fourier transform of $1/r$ substituted by $\hat{g}(\mathbf{k})$. All other calculations are carried out as in the original SPME method.[32] The resultant total electrostatic energy is shifted from the exact value by

$$E_{corr} = \frac{1}{2R_0}\left(\sum_{j=1}^{N} q_j\right)^2$$

which is taken into account in accuracy checks.

In MD simulations, the DNA molecule is placed in a roughly spherical water drop, and a rectangular unit cell is constructed around the drop with the minimal separation of $R_0/2$ between the water molecules and the cell sides. The cell is replicated periodically, which gives an infinite lattice of water drops with at least R_0 spacing between the closest neighbors. With the shifted Coulomb's law, eq 1, all interactions between periodical images are eliminated. Because the shifting does not affect the forces, the system behaves in dynamics as if surrounded by an infinite vacuum. Within the drop, the electrostatics are effectively evaluated with R_0 as the long distance cutoff. For dodecamer DNA duplexes commonly used in benchmark tests, its value can be made larger than the system size.

Regarding mathematics, the shifting of the Coulomb potential is not very different from the original Hockney method,[34] but it provides some practical advantages. It allows one to use the same size of the empty buffer zone in all dimensions. This size need not be changed in time according to the variable unit cell dimensions, which is useful for dynamics with vacuum boundaries. In addition, large systems extended in one direction can be treated more efficiently, which is potentially important for long DNA. The shifting approach is also intuitively simple, and, in fact, it does not even need a proof. In contrast, the original Hockney's argumentation raised some doubts in the literature.[35−37]

Simulation Protocols. Several different MD simulations have been performed with the Dickerson−Drew dodecamer (CGCGAATTCGCG,[43]) surrounded by TIP3P water molecules.[44] The canonical B-DNA conformation[45] was used as the starting point. In water drop simulations, the DNA molecule was first immersed in a large rectangular water box, and next external solvent molecules were removed by using a spherical distance cutoff from DNA atoms. The cutoff radius was adjusted to keep 4000 water molecules remaining. This procedure results in a roughly spherical drop with a diameter around 50 Å. In some cases, the drop was neutralized by randomly adding 22 Na^+ ions at least 5 Å away from DNA.

As explained above, the drop was placed to the center of a large rectangular unit cell so that any cell side was $R_0/2$ from the closest atom. The shape of the drop fluctuates in dynamics, and the size of the unit cell is adjusted accordingly at every time step. Evaporated water molecules are detected and excluded from calculations to prevent explosion in the cell size. After every 50 ps, the calculation is stopped, and water molecules that have left the drop are reintroduced with zero velocities by scattering them randomly near the surface of the drop. The rate of evaporation was around 84 mol/ns; that is, on average four molecules had to be reintroduced at each stop.

Simulations in continuous water media with periodical boundaries were carried out with a rectangular unit cell of 45 × 45 × 65 Å, 4200 TIP3P water molecules, and 22 Na^+ ions. In the starting state, the DNA placed to the center of the unit cell was at least 12 Å from any cell side. Although significant rotations were observed in dynamics, the

(43) Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Nature* **1980**, *287*, 755−758.
(44) Jorgensen, W. L.; Chandreskhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(45) Arnott, S.; Hukins, D. W. L. *Biochem. Biophys. Res. Commun.* **1972**, *47*, 1504−1509.

DNA molecule always remained within the borders of the unit cell. These simulations were carried out in NVT ensemble conditions. The bulk water density estimated for the zone between periodical images was 1.000 ± 0.007.

Every system was energy minimized first with the solute held rigid and then with all degrees of freedom. Dynamics were initiated with the Maxwell distribution of generalized momenta at 250 K and equilibrated at this temperature during several picoseconds. The second phase of equilibration included 0.5 ns of dynamics in production conditions, that is, with the temperature bound to 300 K by the Berendsen algorithm[46] with a relaxation time of 10 ps. All calculations were carried out with AMBER98 parameters[47,48] by using the ICMD method[49−51] with increased inertia of water and planar sugar angles[31] and the time step of 0.01 ps. The van der Waals and SPME direct sum interactions were truncated at 9 Å with the value of Ewald parameter $\beta \approx 0.35$. These values are currently most common for DNA simulations. They ensure the residual level of 10^{-5} for the Gaussian charge density at the cutoff distance and give the rms relative error of 10^{-4} in the direct space forces.[32]

The data presented in the sections below have been obtained in three production trajectories of 5 ns each. The first two trajectories (Tj1 and Tj2) were obtained for DNA in a water drop with and without neutralizing Na^+ ions, respectively. Parameter R_0 equaled 50 Å in both cases. The third trajectory (Tj3) was computed for DNA in continuous water with periodical boundary conditions. The SPME spline order of 6 with a 1.5 Å mesh was used in all three trajectories. These parameters correspond to a reasonable compromise between accuracy and the speed of calculations in water drop simulations (see below). The conformations were saved with a 2.5 ps interval, and the last 2.5 ns were used for statistical analysis. Programs Curves,[52] XmMol,[53] and Mathematica by Wolfram Research Inc. were also employed in the data processing.

Several additional trajectories for the same dodecamer have been computed to check the effect of the specific choice of parameters above. Notably, water drop calculations in conditions of Tj1 and Tj2 have been repeated with $R_0 = 70$ Å to verify that the 50 Å long cutoff has no effect. Calculations with periodic boundaries have been carried out with the spline interpolation order of 4 and the mesh size 1 Å, which corresponds to the most common literature conditions, and also with a smaller number of water molecules in the simulation cell (3900). Durations of these trajectories were between 2 and 5 ns, and the results obtained were similar to those in the production runs.

Results and Discussion

The Accuracy of Calculations. The correctness of the present implementation of the SPME method was carefully checked by parallel comparison with the AMBER suite of programs[54] for identical sets of atom coordinates with periodical boundary conditions. With vacuum boundaries, the results can be compared to straightforward pair summation without cutoff. Because extensive accuracy checks of the original method have been reported by the authors,[32] our discussion here is limited to the issues related directly to the main subject of the paper. Tables 1 and 2 show the results of accuracy checks for a pure water drop and that with DNA, respectively. The total electrostatic energy and gradients were computed as described above with $R_0 = 50$ Å and $\beta \approx 0.35$, corresponding to the rms error of 4×10^{-4} for the direct space forces.[32] The corresponding reference values were obtained by direct summation of atom−atom Coulomb interactions without cutoff.

It has been shown[32] that, with the mesh size around 1 Å, the relative rms errors in the reciprocal space energies and forces are below 10^{-5}. Therefore, it is expected that the rms errors in the *total* electrostatic energy and forces in Tables 1 and 2 converge to those of the direct space. It is seen that this is the case for the pure water drop, but not for DNA where the rms

Table 1. Relative Errors in the Electrostatic Energy and Gradients (Generalized Forces) for a Spherical Drop of 2038 TIP3P Water Molecules[a]

spline order	mesh (Å)	ΔE $\times 10^5$	ΔF $\times 10^4$
4	0.75	5.64	4.07
4	1.0	2.61	14.2
4	1.25	15.9	28.5
4	1.50	30.8	55.5
4	1.75	63.0	85.3
4	2.00	139.	114.
6	0.75	7.53	1.58
6	1.0	6.82	1.83
6	1.25	4.27	3.30
6	1.50	3.83	10.0
6	1.75	24.9	22.4
6	2.00	74.4	39.9
8	0.75	7.58	1.58
8	1.0	7.49	1.59
8	1.25	6.72	1.71
8	1.50	2.14	3.67
8	1.75	14.4	11.1
8	2.00	52.8	24.6

[a] Here $\Delta E = |(E - E_0)/E_0|$, and $(\Delta F)^2 = \sum(F^i - F_0^i)^2/\sum(F_0^i)^2$, where E_0 and F_0^i are electrostatic energies and forces computed by straightforward summation of pair Coulomb interactions without cutoff.

Table 2. Relative Errors and the Necessary Computer Time for Calculation of the Electrostatic Energy and Gradients for Dodecamer DNA in a Drop of 4000 TIP3P Water Molecules[a]

order	mesh (Å)	ΔE $\times 10^3$	ΔF $\times 10^3$	time (s)
4	0.75	1.54	2.20	
4	1.0	1.51	2.53	4.30
4	1.25	1.45	3.05	2.57
4	1.50	1.38	6.86	1.51
4	1.75	1.28	8.18	1.04
4	2.00	0.90	11.8	0.90
6	0.75	1.55	2.24	
6	1.0	1.55	2.24	4.59
6	1.25	1.54	2.25	2.89
6	1.50	1.48	2.55	1.83
6	1.75	1.36	3.00	1.30
6	2.00	1.10	3.96	1.15
8	0.75	1.55	2.24	
8	1.0	1.55	2.24	5.19
8	1.25	1.55	2.24	3.43
8	1.50	1.52	2.24	2.35
8	1.75	1.41	2.50	1.80
8	2.00	1.19	3.03	1.76

[a] The errors are computed as in Table 1.

(46) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(47) Cheatham, T. E., III; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845−862.
(48) 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.
(49) Mazur, A. K. *J. Comput. Chem.* **1997**, *18*, 1354−1364.
(50) Mazur, A. K. *J. Chem. Phys.* **1999**, *111*, 1407−1414.
(51) Mazur, A. K. In *Computational Biochemistry and Biophysics*; Becker, O. M., MacKerell, A. D., Jr., Roux, B., Watanabe, M., Eds.; Marcel Dekker: New York, 2001; pp 115−131.
(52) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63−91.
(53) Tuffery, P. *J. Mol. Graphics* **1995**, *13*, 67−72.
(54) Case, D. A.; Pearlman, D. A.; Caldwell, J. C.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 6*; University of California: San Francisco, CA, 1999.

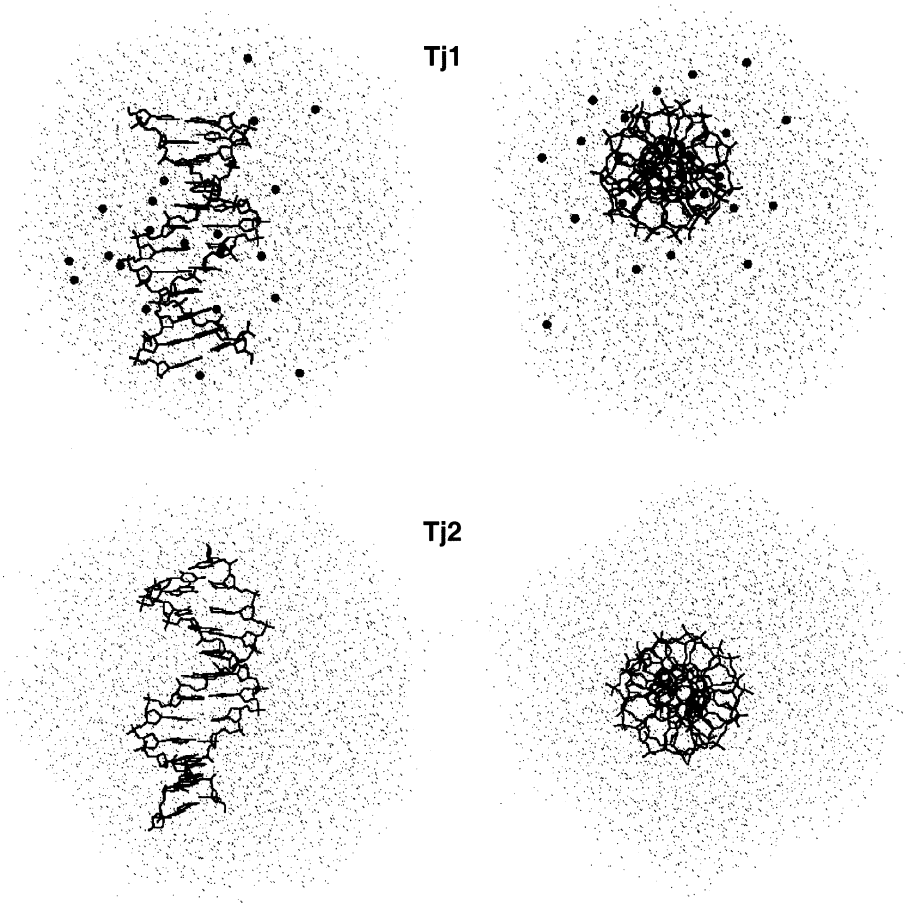

Figure 1. Two perpendicular views of the snapshots from the last nanosecond of Tj1 and Tj2, with water oxygen and Na^+ ion positions shown by small and large dots, respectively.

error values converge to somewhat higher values. This difference can be attributed to the correction of the 1−3 and 1−4 interactions in DNA (masked interactions in terms of ref 32) which makes an additional contribution in Table 2 because in Table 1 tests water molecules were considered rigid. The overall error is still acceptable, and one can see that, with higher order spline interpolation, it exhibits rather small variations with the mesh size. On the other hand, the timing comparison in Table 2 shows that, with a large number of additional mesh nodes in the empty zone around the drop, one can somewhat improve the speed without losing the accuracy by using a slightly wider mesh and a higher spline order. The mesh size of 1.5 Å and the order of 6 used in our simulations represent a reasonable compromise between the speed and the accuracy.

Computed DNA Structures. Figure 1 shows two snapshots form the last nanoseconds of the two water drop simulations with and without counterions. In both trajectories, the DNA duplex looked stable and exhibited no signs of significant deformations or denaturation that could have been caused by the unusual environment conditions. It was always well within the drop and separated from vacuum by several water layers, but did not stay exactly in the middle. The water media remained continuous without internal bubbles.

The surface of the charged drop was covered by spiky "protuberances", and this is the main apparent difference between the two snapshots in Figure 1. Noteworthy is the shape of the charged drop viewed along the DNA axis which resembles a regular polygon rather than a circle. These repetitive features are not yet explained. The spikes were somewhat larger during the early phases of dynamics. Additional control simulations confirmed that they are caused by the high electric field around the charged drop. With the total number of water molecules reduced, the spikes increase and eventually become comparable with the drop size. In this case, the DNA molecule sometimes appears partially stripped of water and may collapse. In contrast, with increased drop size, the spikes become smaller. If the drop is neutralized by counterions, the spikes do not appear regardless of its size.

Table 3. Some Structural Parameters of Standard and Computed DNA Conformations[a]

	rise	twist	RMSD-A[b]	RMSD-B[b]
A-DNA	2.6	32.7	0.0	6.2
B-DNA	3.4	36.0	6.2	0.0
Tj1	3.2 ± 0.085	34.1 ± 0.80	4.57 ± 0.47	2.18 ± 0.38
Tj2	3.3 ± 0.089	33.8 ± 0.75	4.52 ± 0.55	2.36 ± 0.45
Tj3	3.3 ± 0.097	33.6 ± 1.00	4.46 ± 0.45	2.41 ± 0.32

[a] Sequence averaged helical parameters were computed with the program Curves[52] for conformations averaged over the last 2.5 ns of dynamics. Standard deviations are the time variances for the corresponding parameter during the same period. All distances are in angströms, and angles are in degrees. [b] Heavy atom root mean square deviation from the corresponding canonical DNA form.

The DNA structures averaged over the last 2.5 ns are characterized in Table 3. These are B-DNA conformations with the helical parameters similar in all three cases. It is understood that, with zero counterion concentration, the system should have exploded if the duplex length was increased beyond a certain

value because otherwise the electrostatic energy would eventually become infinite. However, no clear trend is seen in Table 3, which indicates that critical lengths of catastrophic deformations are much larger, whereas the dodecamer B-form is metastable in water even if it is charged.

The scales of differences between the DNA conformations in the three trajectories are well characterized by the data in Table 3, and they were always far from statistically significant. The corresponding experimental data are available only for the twist; its value drops by 0.1° when the NaCl concentration is reduced from 0.3 to 0.05 M.[1] A shift by 0.1° in the average twist is too small to be detected in a 5 ns simulation, because, for this relatively small molecule, its fluctuations between consecutive 1 ns averaged structures can reach 1.0°. The absolute twist value is lower than that in experiment, which is a known general feature earlier discussed in the literature.[47]

The Possible Influence of Long Cutoff. The principal goal of these simulations is to see what happens if one gets rid of the artificial conditions in standard PME calculations, that is, periodical boundaries and charge neutralization. This would have little sense if we were bound to introduce new potential sources of artifacts in place of the old ones. There are two essential new elements in the present calculations that may be important. First, the drop boundary can affect many solvent properties such as pressure, mobility, orientation isotropy, and so forth. We should certainly care about these effects, but not for eliminating them because they are due to physical factors that would also work in an experiment with a similar setup. With the current water models, these factors perhaps are not well reproduced, but the situation can be improved. The second new element is the long distance cutoff R_0. This is an artificial limitation, and one should make sure that it does not affect the results and conclusions.

The value $R_0 = 50$ Å approximately equals the initial size of the drop, and it is larger than any atom−atom distance in DNA during dynamics. For water and counterions, this is not the case, however, because the shape of the drop fluctuates. One could imagine, for example, that the Na^+ ions would prefer mutual distances beyond cutoff, and this could have perturbed DNA dynamics. Radial distribution functions present a simple and sensitive test for the possible accumulation of like-charge ion pairs at the cutoff distance.[55] The radial distribution functions for water and counterions in Tj1 and Tj2 are shown in Figure 2. Note that these are nonnormalized distributions; therefore, the areas cut by the vertical line at 50 Å are proportional to the fractions of the corresponding atom pairs affected by the cutoff. Their number is very small for ions and somewhat larger for water molecules. In all cases, however, the plots show no perturbations at the cutoff distance even though no smoothing was applied. The water distribution for the charged drop is slightly shifted toward longer distances probably due to the specific fluctuations of its shape described above.

The data shown in Figure 2 suggest that the finite cutoff value $R_0 = 50$ Å used in our production runs had a negligible effect upon the results. This conclusion was additionally confirmed by shorter simulations with $R_0 = 70$ Å in otherwise similar conditions.

DNA Conformational Mobility. The issue of conformational mobility is important in the context of the possible future application of such methods for DNA simulations. The surface tension at the water−vacuum boundary is likely to increase the internal pressure and can enhance local water ordering around DNA, reducing conformational fluctuations that are the main focus of MD studies. This issue is sometimes invoked as a possible general drawback of simulations with finite water shells.[56] Data in Table 3 indicate, however, that the time fluctuations of the DNA structure were not significantly different in all three trajectories. The same conclusion may be drawn from Figure 3, where detailed patterns of atom fluctuations are presented. The data in Figure 3 can be compared to similar patterns published earlier for a shorter PME simulation with periodical boundaries[57] and ICMD calculations with semiimplicit electrostatics.[31] The results appear to be qualitatively similar in all of these studies, with only minor quantitative differences. Notably, the mobility of bases in Figure 3 is slightly higher and in better agreement with experimental data,[43] which is probably due to a 5 times longer sampling period as compared to ref 57. As compared to the earlier "minimal DNA" ICMD calculations,[31] the fluctuations in Table 3 and Figure 3 are larger by approximately a factor of 1.5.

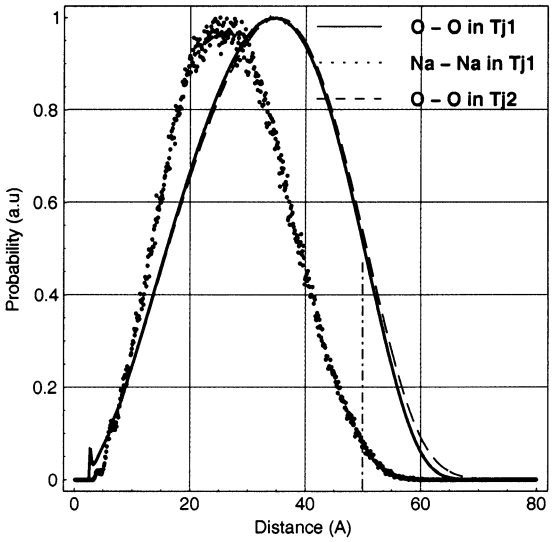

Figure 2. Calculated distribution functions for the distance between water oxygens and that between Na^+ ions in Tj1 and Tj2. The distributions are not volume normalized; that is, each point gives the relative number of atom pairs at a given distance. Data were counted with a window of 0.1 Å over all system configurations saved during dynamics. The plots are scaled to have the same maximum equal one. No smoothing was applied. The long cutoff distance of 50 Å is marked by a dash−dot line.

Water and Counterion Distribution around DNA. In Figure 4, computed distribution functions of Na^+ ions around DNA are compared for simulations with periodical and vacuum boundaries. The distributions are rather similar, with small differences that should be attributed to poorly sampled long time fluctuations of the DNA structure. In both distributions, three peaks can be distinguished within the 10 Å DNA radius, at approximately 3.5, 5, and 8 Å. The first of them is just a shoulder of the larger second peak, but it is reproduced in both simulations. All of these peaks correspond to penetration of counterions into the DNA grooves. Beyond the DNA radius,

(55) Auffinger, P.; Beveridge, D. L. *Chem. Phys. Lett.* **1995**, *234*, 413−415.
(56) Cheatham, T. E., III; Young, M. A. *Biopolymers* **2000**−**2001**, *56*, 232−256.
(57) Duan, Y.; Wlkosz, P.; Crowley, M.; Rosenberg, J. M. *J. Mol. Biol.* **1997**, *272*, 552−572.

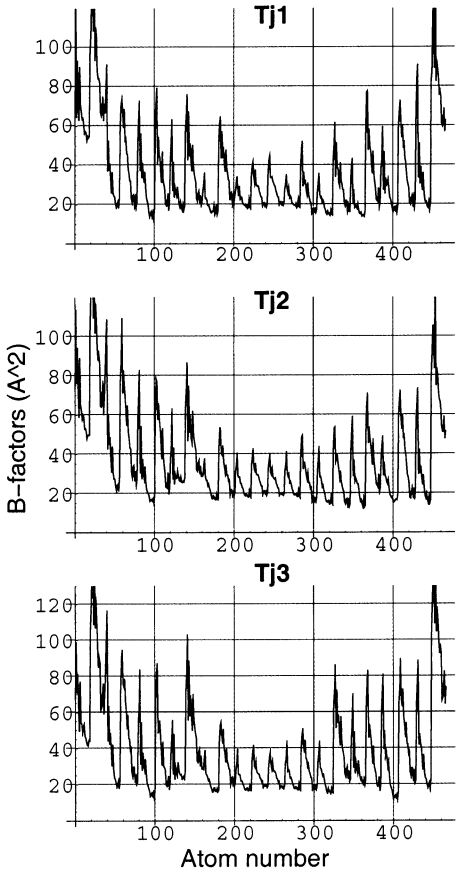

Figure 3. Comparison of calculated *B* factors computed from time fluctuations of atom coordinates in the three trajectories. The rms *B* factors were estimated as $B = (8\pi^2/3)\langle\Delta r^2\rangle$ from atom position fluctuations in superimposed duplex conformations during the last 2.5 ns of dynamics. Atoms in nucleotides are ordered according to expected *B* factors; that is, the phosphate group is the first, and the base atoms are the last. As a result, each nucleotide produces one distinct peak in the figure. The nucleotides are grouped according to base pairing, with the four middle peaks corresponding to the central AT step of the duplex.

there is one broad peak at 13 Å which seems to be split in two at its top. Both the positions and the heights of the main peaks in Figure 4 are in good agreement with analogous data reported earlier.[20]

The distributions obtained in the same way for water oxygens are compared in Figure 5. These plots reveal some unexpected features. Had DNA been replaced by a hard cylinder, one should have distinguished for $r > 10$ Å a few maxima corresponding to the first water layers with relaxation to a uniform distribution. Here two such maxima are seen, but the water density seems to reduce with the distance. In principle, both for Tj1 and for Tj3 this can be an artifact. In Tj1, fluctuations of the shape of the drop can produce such an effect. In Tj3, this may result from the counting procedure when DNA is inclined with respect to the faces of the unit cell (see caption to Figure 4). The trend, however, is observed at fairly short distances, and it is very similar in the two cases. It seems possible, therefore, that water density is really increased near DNA, and this effect deserves verification.

The three peaks in Figure 5 distinguishable at shorter distances can be related to the DNA structure. The first peak at 4 Å corresponds to the first water shell in the major groove. A cylindrical surface with this radius goes too close to bases in the minor groove. The next peak results from the second major

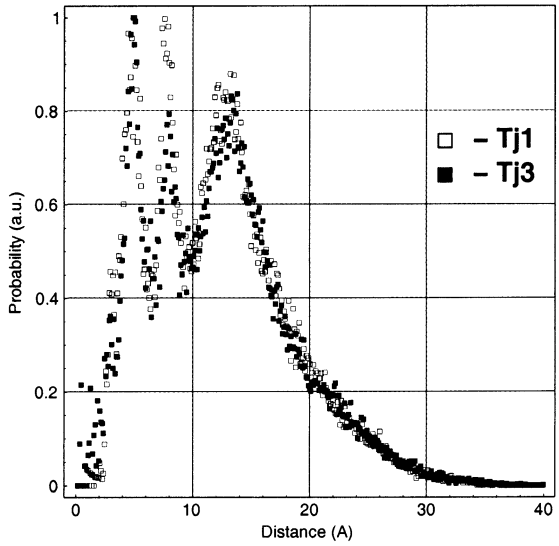

Figure 4. Cylindrical radial distribution functions for Na^+ ions around DNA in Tj1 and Tj3. DNA structures saved in dynamics together with surrounding counterions were superimposed with a canonical B-DNA structure with the global coordinate OZ direction as its helical axis. The Na^+ ions were counted in co-axial 0.1 Å thick cylinders, with their length approximately equal to that of the canonical B form to exclude ions beyond the termini. However, when the instantaneous DNA conformation is curved, the OZ axis can be partially exposed to solvent, which gives trace quantities of ions detected at distances below 3 Å. The distributions are volume normalized, that is, scaled with a factor of $1/r$. For Tj3, only one periodical unit cell was considered with DNA in its center; therefore, only distances less than ~20 Å are meaningful. No smoothing was applied. The plots are scaled to have the same maximum equal one.

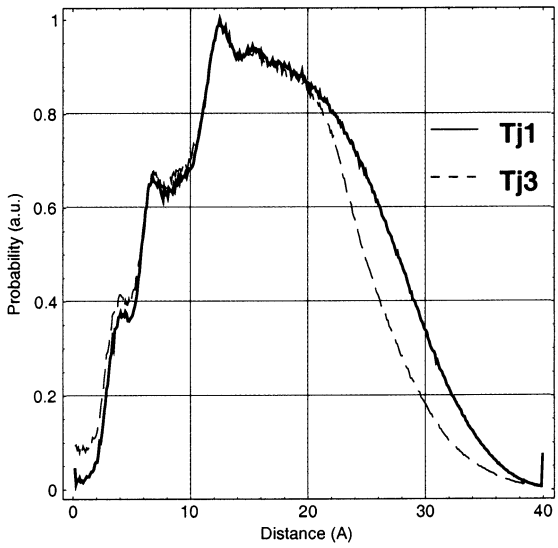

Figure 5. Cylindrical radial distribution functions for water oxygens around DNA in Tj1 and Tj3. The distributions were computed in the same way as for Na^+ ions in Figure 4. A nonzero probability for distances beyond 40 Å is due to evaporated water molecules from Tj1.

groove water layer plus the first one in the minor groove. It is larger than the first peak because the relative accessible volume is increased. In contrast, the third peak at ~10 Å is similar to the previous because it corresponds to the next water layers in both grooves. Between 10 and 13 Å, the relative accessibility is again increased because cylindrical surfaces go beyond DNA; therefore, the next peak is again significantly higher.

One may note, finally, that there is some correspondence between the distributions shown in these two figures regarding

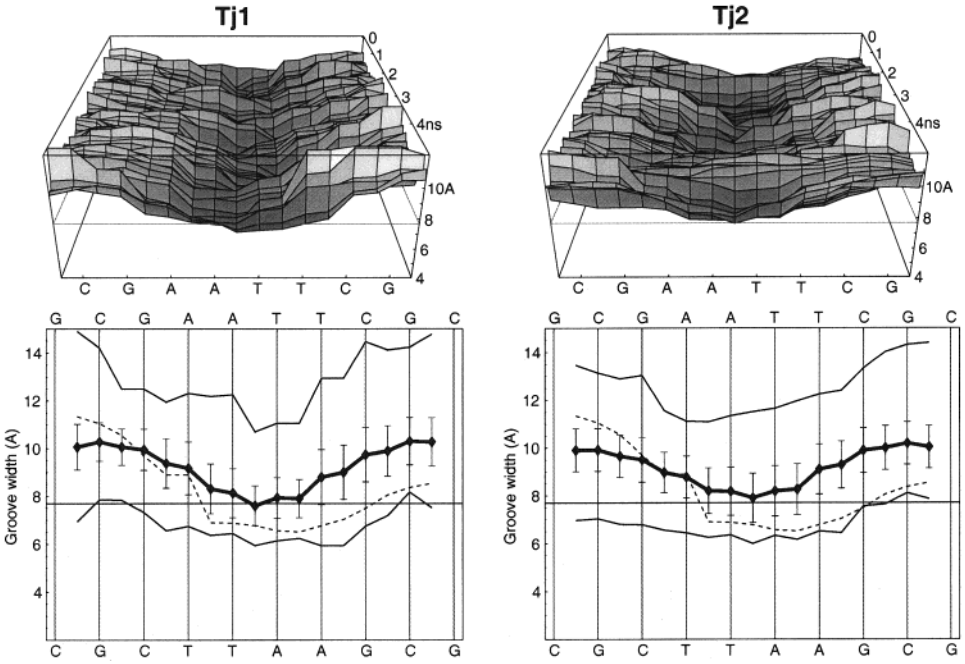

Figure 6. The time evolution of the minor groove and its profile averaged over the last 2.5 ns in Tj1 and Tj2. The surface plots are formed by time-averaged successive minor groove profiles, with that on the front face corresponding to the final DNA conformation. In lower plates, the central traces represent the average groove width with rms fluctuations shown as error bars. The upper and lower solid traces show the maximal and minimal values, respectively. The dotted traces exhibit the experimental profile.[43] The groove width is evaluated by using space traces of C5′ atoms.[67] Its value is given in angströms, and the corresponding canonical B-DNA level of 7.7 Å is marked in all plates by the thin straight lines. Note that the groove width can be measured only starting from the third base pair from both termini. Despite the narrowing, the groove width remains larger than the canonical value,[67] which corresponds to the lower average twist.

the number and the positions of the peaks. This observation indicates that the counterion distributions in Figure 4 result from DNA−Na^+ interactions as well as the space accessibility around the double helix. Comparison of the relative weights of the peaks shows that the ions manifestly keep close to DNA and prefer to enter the DNA grooves rather than to stay outside. Within the DNA radius, however, they are more or less evenly distributed over the accessible space, suggesting that, on average, the radial component of the electric field is closer to zero.

Minor Groove Modulations. The most famous feature of this DNA molecule is the middle AATT fragment. It is long known from experiments that the minor DNA groove always narrows in this and some similar sequences, called A-tracts, and widens outside of them. Figure 6 exhibits dynamics and the average minor groove profiles in Tj1 and Tj2. It has a characteristic waving shape with a narrowing in the middle. The amplitude of this modulation is similar to that in the experimental X-ray structure. The minimal width is 1.5 Å larger than the experimental value, which is probably linked mechanically to the lower average twist. Similar results were obtained for Tj3 and other trajectories, and they are close to earlier reported simulation studies carried out with nonzero counterion concentrations.[20]

The sequence-dependent groove-width modulations in DNA are well established experimentally, and, in recent years, they have been proposed to result from interactions with bound monovalent metal ions commonly undetectable in X-ray crystal maps.[4,8,26] The present results evidence that it is not the case, supporting recent conclusions of different groups.[19,22,28] They explain also why groove modulations and intrinsic DNA bending could be reproduced in MD simulations with simplified treatment of electrostatic interactions that ignored sequence specific counterion effects.[31,58−60]

Concluding Discussion

According to the counterion condensation theory, DNA in aqueous environment should be always covered by a shell of counterions, and its charge should be compensated by around 75% regardless of the bulk ion concentrations.[61] The results presented here do not contradict this theory, but they are somewhat at odds with an implicit assumption that the counterion cloud should always strongly affect the DNA conformation. Such effects certainly exist, but the example considered here shows that not all experimental DNA deformations can be attributed to invisible free ions. The correlations observed in the recent MD simulations of the Dickerson−Drew dodecamer[26] apparently were due to the binding of counterions in sites of low potential near an already narrowed minor groove. Therefore, these interactions are structure-specific rather than sequence-specific, and they cannot be the driving force of the A-tract minor groove narrowing. At the same time, the structure-specific interactions are not less important because they can equally well highlight conformations that otherwise would not be populated. Strong counterion effects upon DNA can probably result from their collective behavior even without direct binding. These mechanisms are more intricate than the simple "association-deformation" paradigm, and they remain important issues for future studies.

(58) Mazur, A. K. *J. Am. Chem. Soc.* **2000**, *122*, 12778−12785.
(59) Mazur, A. K. *J. Comput. Chem.* **2001**, *22*, 457−467.
(60) Mazur, A. K.; Kamashev, D. E. *Phys. Rev. E* **2002**, *66*, 011917(1−13).
(61) Manning, G. S. *Q. Rev. Biophys.* **1978**, *2*, 179−246.

The small difference in DNA dynamics simulated with and without added salt is not completely unexpected. MacKerell earlier reported about a no-salt DNA simulation with periodical boundaries and a 13 Å cutoff.[62] Cheatham and Kollman carried out PME calculations with DNA charge neutralized by smearing the opposite charge equally over all atoms (neutralizing plasm).[29] In both cases, the authors reported about the absence of dramatic changes in no-salt conditions. However, specific experimental phenomena attributed to counterions were not modeled, and these results were not considered as a counter argument because one could not claim that the counterion contribution was completely eliminated. Note, for instance, that cutoff truncation of electrostatic interactions is equivalent to placing neutralizing charges on the cutoff sphere of every atom.[63] In any case, the electric field around DNA, which in real conditions is quenched by ions, was also quenched. The same argument applies to simulations with partial hydration and implicit treatment of counterions where minor groove modulations were observed.[31,59]

The simulations described here are more convincing because they take into account all interactions with no artificial damping of electrostatics. Calculations with vacuum boundaries are based upon only two assumptions: (i) Coulomb's law is correct, and (ii) the force field is sufficiently accurate. Unlike periodical boundaries, the water drop conditions, in principle, can be exactly reproduced in experiment. These calculations demonstrate that, even in the absence of counterions, the strong electrostatic field around one turn of the double helix is suppressed by water to a low level that cannot significantly affect the structure. It is rather counter intuitive because the DNA molecule is known to be easily deformable by even small perturbations. This result also qualitatively disagrees with earlier energy calculations with implicit solvent models.[64,65]

As noted above, theoretically, the conditions of the charged drop can be reproduced in experiment. Assuming that it is done and the result of such an experiment confirmed calculations, one still cannot claim that DNA in such conditions corresponds to that in infinite water with low ion concentration. As seen in Figure 1, the Na^+ ions keep close to DNA, while one-half of the drop is free from ions. It is quite possible that the ions will not notice if we start adding extra water layers around the drop, which would agree with the counterion condensation idea.[61] However, other theoretical predictions indicate that the counterions should leave DNA when the Debye length becomes much larger than that of DNA.[66] Our calculations can neither confirm nor disprove the last assertion. With zero ion concentration, the Debye length is formally infinite, and, as we see, it is possible for ions to leave DNA. This does not mean, however, that they would do it spontaneously. To check this, long simulations with intermediate Debye lengths are necessary. One Na^+ ion per 4000 water molecules already gives concentration corresponding to a Debye length around 40 Å. To increase it 2 times, we need a 4 times larger water drop. Therefore, calculations with neutral drops and larger Debye lengths are too computationally demanding.

The approach to simulations of DNA dynamics applied here continues earlier attempts to find an efficient way to model long double helices without the computational burden of continuous media with periodical boundary conditions.[30,31,59] It is much more faithful than calculations with semiimplicit hydration with a "minimal DNA" model[31,59] because no force field modifications are involved. At the same time, the physical effects related to the limited solvent shell certainly remain. It is hoped that the result exhibited in Figure 2, a negligible effect of the long cutoff, will hold for larger systems as well. In this case, one can imagine simulations of dynamics of long DNA covered by a shell of water and counterions approximately as thick as in the present water drop calculations, which would allow one to study rather long DNA molecules soon.

JA012706E

(62) MacKerell, A. D., Jr. *J. Phys. Chem. B* **1997**, *101*, 646−650.
(63) Wolf, D.; Keblinski, P.; Phillpot, S. R.; Eggebrecht, J. *J. Chem. Phys.* **1999**, *110*, 8254−8282.
(64) Zhurkin, V. B.; Lysov, Y. P.; Florentiev, V. L.; Ivanov, V. I. *Nucleic Acids Res.* **1982**, *10*, 1811−1830.
(65) Flatters, D.; Zakrzewska, K.; Lavery, R. *J. Comput. Chem.* **1997**, *18*, 1043−1055.
(66) Woodbury, C. P., Jr.; Ramanathan, G. V. *Macromolecules* **1982**, *15*, 82−86.
(67) Mazur, A. K. *J. Mol. Biol.* **1999**, *290*, 373−377.