Statistical mechanical treatment of the structural hydration of biological macromolecules: Results for *B*-DNA

Gerhard Hummer
Theoretical Biology and Biophysics Group T-10, MS K710, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Dikeos Mario Soumpasis
Biocomputation Group, Max Planck Institute for Biophysical Chemistry, P.O. Box 2841, D-37018 Göttingen, Germany
(Received 13 May 1994)

We constructed an efficient and accurate computational tool based on the potentials-of-mean-force approach for computing the detailed hydrophilic hydration of complex molecular structures in aqueous environments. Using the pair and triplet correlation functions database previously obtained from computer simulations of the simple point charge model of water, we computed the detailed structural organization of water around two *B*-DNA molecules with sequences $d(AATT)_3 \cdot d(AATT)_3$ and $d(CCGG)_3 \cdot d(CCGG)_3$, and canonical structure. [A, T, C, and G denote adenine, thymine, cytosine, and guanine, respectively, and $d(\cdots)$ denotes the deoxyribose in the sugar-phosphate backbone.] The results obtained are in agreement with the experimental observations. A·T base-pair stretches are found to support the marked minor-groove "spines of hydration" observed in x-ray crystal structures. The hydrophilic hydration of the minor groove of the molecule $d(CCGG)_3 \cdot d(CCGG)_3$ exhibits a double ribbon of high water density, which is also in agreement with x-ray crystallography observations of C·G base-pair regions. The major grooves, on the other hand, do not show a comparably strong localization of water molecules. The quantitative results are compared with a computer simulation study of Forester *et al.* [Mol. Phys. **72**, 643 (1991)]. We find good agreement for the hydration of the $-NH_2$ groups, the cylindrically averaged water density distributions, and the overall hydration number. The agreement is less satisfactory for the phosphate groups. However, by refining the treatment of the anionic oxygens on the phosphate groups, almost full quantitative agreement is achieved.

PACS number(s): 87.15.Da, 87.15.By, 68.45.−v

I. INTRODUCTION

In their functional form, biomolecules are surrounded by an aqueous environment. The intrinsic stability of biomolecules and their biochemical reactions depend crucially on the interactions with their environment, i.e., water and ions. In particular, water plays an essential role in mediating molecular association such as antibody-antigen complex formation, protein-DNA interaction, or drug binding to proteins and nucleic acids [1–4]. Structurally localized water can result in strongly enhanced specificity of these reactions. Therefore, proper understanding, modeling, and prediction of the properties and interactions of central biomolecules such as nucleic acids and proteins require an adequate description of their interactions with water.

In the case of the DNA-water interaction, a substantial amount of information has been accumulated during the past 40 years using experimental techniques such as fiber [5,6] and particularly crystal x-ray studies discussed below, infrared [7,8] and Raman spectroscopy [9], buoyant density [10,11], and gravimetry [12], and more recently by volumetry [4], ultrasonics [13,14], Brillouin scattering [15], neutron diffraction [16], and nuclear magnetic resonance (NMR) spectroscopy [17,18]. The experimental results have been widely reviewed [19–24]. These results mainly indicate that the short-range hydrophilic interactions of electronegative DNA atoms with water give rise to a strongly sequence- and conformation-dependent organization of water molecules forming well localized networks on parts of the DNA-solvent interface. The rest of the water in the system can be thought to affect DNA via the poorly understood hydrophobic interaction as well as via its bulk-medium properties, e.g., providing a high dielectric constant, low compressibility fluid background.

Many of the characteristic features of *B*-DNA structural hydration were first observed by Dickerson and collaborators using x-ray crystallography. For example, they first described the "water spine" covering the minor groove of the AATT part in the d(CGCGAATTCG-CG) dodecamer [25,26]. [A, T, C, and G denote adenine, thymine, cytosine, and guanine, respectively, and $d(\cdots)$ denotes the deoxyribose in the sugar-phosphate backbone.] Other research groups subsequently described various patterns of structural hydration in *A*-DNA [27] and *Z*-DNA crystals [28,29], and in *D*-DNA fibers [16]. The presence of a "spine of hydration" in A_nT_n tracts of *B*-DNA in solution was also observed in NMR studies, which showed anomalously long residence times of water molecules in the minor groove [17,18].

On the theoretical modeling side, until recently the

main approach used to study DNA hydration has been the computer simulation technique of placing a DNA oligomer and several hundred to a few thousand water molecules in a simulation box, specifying the interaction potentials, and letting the computer numerically solve Newton's equations for a given period of time (molecular dynamics) or generate particle configurations of the canonical ensemble via the Metropolis Monte Carlo algorithm.

Studies of this kind were first reported 13 years ago by Corongiu and Clementi [30–32]; these studies dealt mainly with energetic aspects of the water-DNA system. Subsequently, a number of Monte Carlo and molecular dynamics studies of various oligomer sequences [33–36] described some features of the experimentally observed DNA structural hydration.

Computer simulations of DNA (and, in general, biomolecular) hydration are subject to serious limitations owing to the complexity of the DNA-water and, even worse, the DNA-water-ion system, and severe methodological difficulties. In conjunction with long relaxation times, the large particle numbers required result in enormous computing times. The treatment of the long-range and strong Coulomb interactions under periodic boundary conditions requires particular care. Another serious technical problem arising in the computation of water densities around biomolecules is the poor density statistics, which is a consequence of the rather open liquid water structure (low bulk-water particle density of about 33.33 nm^{-3}). In order to understand this point consider the fluctuation of particle numbers N in a volume ΔV, which is related to the isothermal compressibility χ_T through $\langle N^2\rangle - \langle N\rangle^2 = \langle N\rangle \rho k_B T \chi_T$, where $\langle N\rangle = \rho \Delta V$. For water with $\chi_T \approx 4.5 \times 10^{-10}$ $\mathrm{m^2\,N^{-1}}$ under normal conditions we obtain $\rho k_B T \chi_T \approx 0.06$. If we assume that in the vicinity of a solvated macromolecule the density varies considerably over distances of, e.g., 0.05 nm (approximately one sixth of the molecular diameter) and calculate average particle numbers in a volume $\Delta V = (0.05\ \mathrm{nm})^3$, we obtain (using bulk-water characteristics) $\langle N\rangle \approx 0.004$ and $(\langle N^2\rangle - \langle N\rangle^2)^{1/2}/\langle N\rangle \approx 1$. Correspondingly, the number of counts in a local density calculation is extremely small and subject to large fluctuations. In view of this brief discussion it becomes clear that new routes are needed in order to allow routine investigations of solvent effects on the hundreds of structures derived from experiments (x-ray, NMR, etc.) or theoretical modeling.

In this work, we present an alternative computational method that is capable of determining the detailed structural organization of water near DNA (and, with trivial adaptations, any other biomolecule and biomolecular assembly) several orders of magnitude faster than computer simulations. It is based on the potentials-of-mean-force (PMF) approach developed in our group [37–41] that aims at the implicit description of the solvent using a database of its correlation functions. This approach already yielded excellent results concerning ionic effects on DNA with a much cruder water description sufficient for diffuse ionic effects but inadequate for hydration studies (for reviews, see Refs. [41] and [42]). A simpler version of the technique described here was used in conjunction with the restricted primitive model of electrolytes and ionic pair correlation functions to compute ionic distributions around DNA conformations [40]. The approach described here was carefully tested in a very recent study of the water density at the ice-water interface [43]. A preliminary account of some of the results obtained concerning DNA hydration was presented in Ref. [44]. From the presentation in the following section it will become clear that because of its generality the approach followed in this work offers many possibilities for applications in various other fields of nonuniform equilibrium fluids involving liquid phases of molecular solvents.

II. THEORY

We start from an atomic representation of a solvated macromolecule with N_α atoms of type α at positions $\mathbf{r}_{i_\alpha}$ and M different types of atoms α. The set of atomic coordinates (obtained from x-ray crystallography, NMR, modeling, etc.) will be denoted by $\{\mathbf{r}_{i_\alpha}\}$. We will assume the molecule to be rigid. For the time being we will consider only equilibrium structures. However, the great computational efficiency of the method allows one to introduce an additional averaging over the set of coordinates of the molecule. This thermal average can, for instance, be performed using the temperature factors obtained from x-ray crystallography or by using the set of different molecular coordinates consistent with the NMR data, interpreting them as a representative ensemble of molecular configurations.

For the sake of notational simplicity, we present the formal development for the case of a one-component solvent. An explicit description of molecular solvents requires the inclusion of additional angular coordinates or, in the case of interaction site models, intramolecular interactions or bond length constraints. However, the formal development follows along the same lines.

The conditional solvent density at position $\mathbf{r}_1$ can be expressed as a configuration space integral. The solvated macromolecule appears by and large as an external field. But for reasons of generality we prefer a notation with a total potential energy U, which may also contain three- and multiparticle interactions. In the case of a canonical ensemble of N solvent particles with a set of coordinates $\{\mathbf{r}_i\}$ we find for the conditional solvent density

$$\rho(\mathbf{r}_1|\{\mathbf{r}_{i_\alpha}\}) = N \frac{\int d\mathbf{r}_2 \cdots d\mathbf{r}_N \exp[-\beta U(\{\mathbf{r}_i\},\{\mathbf{r}_{i_\alpha}\})]}{\int d\{\mathbf{r}_i\} \exp[-\beta U(\{\mathbf{r}_i\},\{\mathbf{r}_{i_\alpha}\})]}\ , \tag{1}$$

where $\beta = 1/k_B T$. The conditional density can then be related to multiparticle correlation functions,

$$\rho(\mathbf{r}|\{\mathbf{r}_{i_\alpha}\}) = \rho\, \frac{g^{(1;\{N_\alpha\})}(\mathbf{r},\{\mathbf{r}_{i_\alpha}\})}{g^{(\{N_\alpha\})}(\{\mathbf{r}_{i_\alpha}\})}\ , \tag{2}$$

where $\rho = N/V$ is the solvent density. The correlation functions are defined in the conventional way as

$$g^{(\{N_\alpha\})}(\{\mathbf{r}_{i_\alpha}\}) = \left(\prod_{\alpha=1}^{M} N_\alpha! \rho_\alpha^{-N_\alpha}\right) \frac{\int d\{\mathbf{r}_i\} \exp[-\beta U(\{\mathbf{r}_i\},\{\mathbf{r}_{i_\alpha}\})]}{\int d\{\mathbf{r}_{i_\alpha}\} d\{\mathbf{r}_i\} \exp[-\beta U(\{\mathbf{r}_i\},\{\mathbf{r}_{i_\alpha}\})]} , \tag{3}$$

$$g^{(1;\{N_\alpha\})}(\mathbf{r}_1,\{\mathbf{r}_{i_\alpha}\}) = N\rho^{-1}\left(\prod_{\alpha=1}^{M} N_\alpha! \rho_\alpha^{-N_\alpha}\right) \frac{\int d\mathbf{r}_2 \cdots d\mathbf{r}_N \exp[-\beta U(\{\mathbf{r}_i\},\{\mathbf{r}_{i_\alpha}\})]}{\int d\{\mathbf{r}_{i_\alpha}\} d\{\mathbf{r}_i\} \exp[-\beta U(\{\mathbf{r}_i\},\{\mathbf{r}_{i_\alpha}\})]} , \tag{4}$$

where $\rho_\alpha = N_\alpha/V$. The higher-order correlation functions $g^{(\{N_\alpha\})}$ and $g^{(1;\{N_\alpha\})}$ are not readily accessible to an explicit calculation. They describe the correlations of sets of atoms $\{N_\alpha\}$ in the limit of infinite dilution in water. However, we can systematically expand them in terms of lower-order correlation functions. This is accomplished using the PMF expansion for $W^{(n)}(\{\mathbf{r}_i\}) = -k_B T \ln g^{(n)}(\{\mathbf{r}_i\})$ [45]. Truncation after pair and triplet correlation levels corresponds to the Kirkwood [46] and Fisher-Kopeliovich [47] superposition approximations for the two- and three-particle correlation functions, respectively. If we retain terms up to the three-particle correlation level in the PMF expansion of Eq. (2), we find for the conditional solvent density

$$\rho(\mathbf{r}|\{\mathbf{r}_{i_\alpha}\}) = \rho\left[\prod_{\alpha=1}^{M}\prod_{i_\alpha=1}^{N_\alpha} g^{(1;\alpha)}(\mathbf{r},\mathbf{r}_{i_\alpha})\right] \times \left[\prod_{\alpha=1}^{M}\prod_{i_\alpha=1}^{N_\alpha}\prod_{\beta=1}^{M}\prod_{i_\beta=1+\delta_{\alpha\beta}i_\alpha}^{N_\beta} \frac{g^{(1;\alpha,\beta)}(\mathbf{r},\mathbf{r}_{i_\alpha},\mathbf{r}_{i_\beta})}{g^{(1;\alpha)}(\mathbf{r},\mathbf{r}_{i_\alpha}) g^{(\alpha,\beta)}(\mathbf{r}_{i_\alpha},\mathbf{r}_{i_\beta}) g^{(\beta;1)}(\mathbf{r}_{i_\beta},\mathbf{r})}\right] \cdots , \tag{5}$$

where $\delta_{\alpha\beta}$ is the Kronecker symbol.

We now simplify Eq. (5) with regard to the number of different atom types used to model the macromolecule, recalling that the main interactions between water and a solvated macromolecule are hydrogen bonds. The highly directional, relatively weak noncovalent hydrogen bonds are also the main reason for the structurally complex nature of dense water phases (liquid, ice). This property of water is well understood in terms of the quantum-mechanical *sp*-hybridization concept. Among other things it explains the highly asymmetric electronic density (lone pairs) on an oxygen atom in water accompanying the formation of H_2O from an oxygen and two hydrogen atoms, to which a hydrogen from another water molecule can couple. The other atom exhibiting the same phenomenon is oxygen's next neighbor in the periodic table, namely, nitrogen, which is even more chemically versatile than oxygen and therefore is incorporated in the heterocyclic base rings of DNA.

With respect to canonically averaged DNA-water interactions we expect the difference in identity between nitrogen and oxygen atoms in DNA and oxygens in water to be only a minor second-order effect. In the case of nucleic acids with no large hydrophobic regions we represent the nonpolar atoms simply as excluded volume. In future refinements, nonpolar atoms (e.g., carbon) can, for example, be modeled by using correlation functions of methane in water. A more intriguing question is the treatment of the charged phosphate groups PO_4^- on nucleic acids. We expect the presence of the negative charge to be well represented by the four oxygens, each carrying a negative partial charge. However, a refined treatment is outlined in Sec. III, where the two anionic oxygens are represented by two "oxygen ions" carrying a fractional negative charge.

In the following we will obtain a good first-order picture by equating all DNA nitrogen and oxygen atoms to water oxygen. This is supported by a large body of the experimental and quantum-mechanical studies of the geometry and the energies of $X\cdots$H— O and X—$H\cdots$O hydrogen bonds (X = N or O). A complete description of this evidence is not within the scope of this article. However, it provides a rather uniform picture of such bonds with X— O distances 0.27–0.32 nm, quasilinearity (X— H— O angles 180° ± 20°), and energies 20 ± 12 kJ/mol.

In addition, it should be noted that the slightly weaker N$\cdots$H— O and N— H$\cdots$O bonds in some molecular complexes are strengthened when nitrogen is part of an aromatic ring because of the delocalized π electrons. Moreover, the number of electronegative DNA atoms is essentially zero compared to the large number of water oxygens. As a consequence, water organization around DNA is dominated by water-water correlations; and apart from trivial steric exclusion owing to the structure and second-order interactions (e.g., dispersion) hydrophilic hydration can be modeled by fixing water oxygens (and if necessary water hydrogens) at the positions of electronegative DNA atoms (or DNA hydrogens).

In view of this brief discussion and the presently feasible crude level of description of real interactions in many-body systems in terms of simple pairwise additive power-law potentials plus large sets of empirical parameters, we think that a more detailed modeling of the difference between nitrogen and oxygen is presently not warranted.

The exact expression Eq. (2) combined with any truncation of the PMF expansion Eq. (5) yields approximate working formulas for computing particle densities anywhere around a defined structure. Accuracy increases with the number of terms retained in Eq. (5). For sim-

ple models with spherically symmetric pair interactions it suffices to work at the pair level, triplet and higher-order correlation functions being only second-order refinements. This is, for instance, the case for restricted-primitive-model descriptions of electrolytes [48] and their application in the computation of ionic densities around DNA [40].

In the case of microscopic water models, it is essential to include triplet correlations because of the highly anisotropic nature of the intermolecular interactions in water (hydrogen bonds, tetrahedral coordination) [43]. In particular, retaining the triplet term in the expansion corrects for the considerable, but wrong, pair contributions of equilateral triplets with edges 0.275 nm. There is only a negligible probability for these triplets to occur in liquid water, which exhibits a preferentially tetrahedral local structure with oxygen distances 0.275, 0.275, and 0.45 nm as a result of the hydrogen bond interactions. Equation (5) truncated at the triplet level yields the working formula for calculating the water-oxygen density around any configuration of electronegative atoms (oxygen and nitrogen for DNA) at positions $\mathbf{r}_1, \ldots, \mathbf{r}_n$,

$$\rho_O(\mathbf{r}|\mathbf{r}_1, \ldots, \mathbf{r}_n) = \rho_0 \prod_{i=1}^{n} g_{OO}^{(2)}(\mathbf{r}, \mathbf{r}_i) \times \prod_{j=1}^{n-1} \prod_{k=j+1}^{n} \frac{g_{OOO}^{(3)}(\mathbf{r}, \mathbf{r}_j, \mathbf{r}_k)}{g_{OO}^{(2)}(\mathbf{r}, \mathbf{r}_j) g_{OO}^{(2)}(\mathbf{r}_j, \mathbf{r}_k) g_{OO}^{(2)}(\mathbf{r}_k, \mathbf{r})} , \quad (6)$$

where ρ_0 is the bulk-water density. The oxygen pair correlation function $g_{OO}^{(2)}$ entering Eq. (6) is known both from experiment and computer simulations employing various water models. The previously unknown triplet correlations $g_{OOO}^{(3)}$ were computed recently in collaboration with Procacci and Corongiu [49], from a molecular dynamics (MD) simulation of the highly refined *ab initio* Niesar-Corongiu-Clementi model of water [50].

In the present study we use the simple point charge (SPC) water model [51], which is simpler but also gives satisfactory results. Many structural, dynamic, and thermodynamic properties of bulk-water phases are reproduced well. The SPC model has a rigid structure with one oxygen and two hydrogen interaction sites. The O— H bond length is 0.1 nm and the H— O— H bond angle is ideally tetrahedral (109.47°). The oxygen and hydrogen sites carry partial charges of $q_O = -0.82|e|$ and $q_H = 0.41|e|$, respectively. These charges are buried inside a Lennard-Jones interaction sphere centered at the oxygen position. Metropolis Monte Carlo computer simulations were used to calculate two- and three-atom correlation functions of SPC water, as described in Ref. [43]. Pair and triplet correlations were obtained at a discretization of 0.005 nm and 0.02 nm, respectively. The particle correlations were sampled up to pair distances of 1.1 and 0.72 nm for pairs and triplets. (Because of the homogeneous and isotropic character of the bulk fluid, the two- and three-atom correlations can be represented as functions of one and three pair distances, respectively.)

Oxygen pair distances below 0.25 nm hardly occur in SPC water under normal conditions; but electronegative oxygen and nitrogen atoms on DNA and other macromolecules are found as close as about 0.2 nm if they are connected through an intermediate carbon. For those configurations only the pair contribution in Eq. (6) can be calculated, since the appropriate three-particle correlations are not known from bulk-water simulations. However, it is possible to calculate the three-particle corrections from computer simulations with two particles constrained to shorter distances than found in bulk simulations. The conditional density at distance r and s from two particles constrained to a distance t is given by

$$\rho(r, s|t) = \rho_0 \frac{g^{(3)}(r, s, t)}{g^{(2)}(t)} . \quad (7)$$

The three-particle correction required in Eq. (6) can then be calculated as $\rho(r, s|t)/[\rho_0 g^{(2)}(r) g^{(2)}(s)]$. Monte Carlo simulations were carried out to calculate $\rho(r, s|t)$ for oxygen distances 0.19, 0.21, 0.23, and 0.25 nm to supplement the $g^{(3)}$ data of the bulk fluid. The same conditions and methods as in the bulk simulations were used [43]. As in the computer simulation calculation of $g^{(3)}$, the calculation of the normalization volumes of $\rho(r, s|t)$ requires some caution with regard to the triangle inequalities [48]. The $\rho(r, s|t)$ data were sampled over 225 000 Monte Carlo passes for each value of t of a system containing 256 SPC water molecules. The oxygen atoms of two water molecules were constrained to a given distance t, with the two molecules otherwise being free to rotate and translate.

III. RESULTS AND DISCUSSION

A. Characteristics of *B*-DNA structural hydration

In this work we will report representative results for DNA duplexes in the *B* family of conformations. A more extensive discussion of the information obtained so far, including the other DNA and RNA conformations [44,52] and also proteins [53] will be presented elsewhere.

Using Eq. (6) along with linear interpolation of the discrete pair and triplet correlation data, water-oxygen densities were computed at vertices of a three-dimensional Cartesian grid covering a volume $3 \times 3 \times 4.2$ nm^3. A grid spacing of 0.03 nm was used. The carbon atoms were surrounded by a sphere of radius 0.3 nm, from which water oxygens were excluded. This radius corresponds approximately to the distance of closest approach of water oxygens to methane. The triplet correction in Eq. (6) was applied to triangles with all edges between 0.19 and 0.71 nm. The atomic coordinates of canonical (Arnott-Hukins) *B*-form structures [54] for d(AATT)$_3$·d(AATT)$_3$ and d(CCGG)$_3$·d(CCGG)$_3$ were used in the computations. These particular sequences were chosen because they contain all possible steps of like base pairs A·T and G·C, i.e., 5′-ApA, 5′-TpA, 5′-ApT, 5′-GpG, 5′-CpG, and 5′-GpC. Here and in the following we adopt the standard

nomenclature for the nucleic acid composition [20]. We use abbreviations for the bases: adenine (A), guanine (G), thymine (T), and cytosine (C). The deoxyribose in the sugar-phosphate backbone is denoted by d(···). The strand polarity is represented by writing the base sequence in the 5′ to 3′ direction. To emphasize the polarity, dinucleotide steps XY are written as 5′-XpY, where p represents the sugar-phosphate linkage. A dot denotes Watson-Crick base pairing (i.e., A·T and G·C).

Unlike the crystallographic structure of B-DNA oligomers, canonical B-DNA does not exhibit variations in the local helical parameters. In particular, the minor-groove width is constant while crystallographic structures show a distinct sequence dependence, with the minor groove in general being narrow at A·T and wide at G·C base pairs. However, the study of these idealized DNA structures with uniform helical parameters allows one to demonstrate the general sequence-dependent characteristics of B-DNA hydration. Similar studies can be done for x-ray crystal structures, for which some results are shown in Ref. [44].

Figure 1 shows a stereographic plot of the main features of the structural hydration obtained for the dodecamer d(AATT)$_3$·d(AATT)$_3$. The positions marked by points correspond to grid points of highest density for water oxygens. A density threshold of five times the bulk value was chosen to illustrate the characteristics of B-DNA hydration. The most pronounced localization is observed in the minor groove of the molecule, with elongated high-density regions in the center of the minor groove between consecutive base pairs. The regions with high water density show the strongest localization at 5′-ApT steps. They are somewhat diffuser at 5′-ApA steps. At 5′-TpA steps, an even larger volume shows water densities above 5. (Here and in the following, water densities will be given in units of the bulk density.) However, the density maxima are higher at 5′-ApA and 5′-ApT steps. The density values at 5′-TpA steps are below 20, whereas at 5′-ApA and 5′-ApT steps few grid points (one to three per base pair for the resolution of 0.03 nm used in these calculations) show densities above 30. (The absolute maximum observed is about 60.)

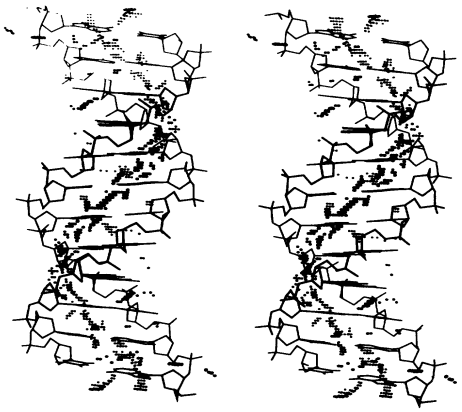

FIG. 1. Hydration of the B-DNA dodecamer with sequence d(AATT)$_3$·d(AATT)$_3$ and canonical structure. The stereographic plot shows grid points with water density larger than five times the bulk-water density.

The pronounced hydration of the minor groove corresponds to the experimental observation of a "spine of hydration" in A·T tracts of B-DNA x-ray crystal structures [25,26]. In addition to the density maxima between the base-pair planes we also observe high water density closer to the base-pair planes, but more distant from the surface of the minor groove. These peaks are split up into two regions above and below the base-pair planes in a distance of about 0.15 nm. For base pairs involved in 5′-TpA steps, they reach density maxima of about 7; for other base pairs, densities of approximately 10 and 20 are observed at the two peaks, respectively. This second layer of density peaks represents water molecules bridging those of the first layer and thus forms the actual "spine of hydration."

The major groove of d(AATT)$_3$·d(AATT)$_3$ shows a less pronounced localization of water molecules compared to the minor groove. Regions of high water density are observed at 5′-ApA steps between consecutive adenines and at the center of the base-pair steps. At 5′-ApT steps, two small high-density regions are found between the adenines and thymines on both sides of the major groove. The major-groove hydration of 5′-TpA steps does not show significant localization.

The sugar-phosphate backbone does not show considerable regions with high water density. However, by reducing the density threshold in Fig. 1 one finds the PO_4^- groups to be surrounded by regions with water density well above 2. The explanation for the comparably low localization of water is a lack of cooperativity. The local density maximum caused by a single polar atom is (within the approximations of our method) limited by the maximum of the oxygen pair correlation function $g^{(2)}$, which is approximately 3 for the SPC water model used in our calculations. Two polar atoms can give a maximum density equivalent to the maximum of $g^{(3)}(\mathbf{r},\mathbf{s},\mathbf{t})/g^{(2)}(\mathbf{s},\mathbf{t})$ with respect to their distance $|\mathbf{t}-\mathbf{s}|$. Only if two, three, or more polar atoms are positioned in a favorable arrangement may one expect a strong cooperative effect resulting in a high water density. In other words, the by and large spherical character of the PO_4^- groups within our approximations and their distance from other polar groups (on B-DNA) make it unlikely to observe similarly high water densities as in the minor groove of B-DNA. This is in agreement with the experimental ordering scheme for water affinities derived from crystallographic data by Kopka *et al.* [26], but disagrees with a scheme supported by infrared spectroscopy measurements [7]. Nevertheless, the phosphate groups are well hydrated, in the sense of being surrounded by regions of water density 2 and higher. The situation is different with structures of other DNA families (e.g., A, Z), where the phosphate groups are in closer proximity [55] resulting in possible cooperative effects. The same is true for the hydration of the sugar-phosphate backbone in crystal structures of DNA molecules. The crystal packing can bring PO_4^- groups in contact with other molecules, resulting in an increased water density.

In the following we study the hydration of the B-DNA

molecule with C·G base pairs. A canonical structure for a *B*-DNA with sequence $d(CCGG)_3 \cdot d(CCGG)_3$ was used. Figure 2 shows a stereographic representation of the structural hydration obtained for this dodecamer. As in the case of the $d(AATT)_3 \cdot d(AATT)_3$ dodecamer, the minor groove shows the most pronounced hydration. However, the qualitative appearance has changed compared to DNA with A·T base pairs. Two side-by-side ribbons of high water density follow the O2 and N2 atoms of the cytosines and guanines in the minor groove of the molecule. These ribbons of high water density in C+G stretches are in agreement with the findings of x-ray crystallography [56].

In addition to the two ribbons, high water densities are found between consecutive base pairs at 5′-CpG and 5′-GpG steps. At 5′-CpG steps, the high-density region is elongated and bridging the two ribbons. At 5′-GpG steps, the high-density peak forms a part of the ribbon on the cytosine strand. The localization of water is found to be less pronounced in the case of C·G base pairs compared to A·T. The high-density regions appear to be less extensive in relation to those shown in Fig. 1. Also the density maxima are approximately 20 and thus significantly lower than those found in the minor groove of $d(AATT)_3 \cdot d(AATT)_3$. (The overall maximum density value on the grid was about 30.)

Similarly to the $d(AATT)_3 \cdot d(AATT)_3$ dodecamer, the major groove does not show a very distinct hydration. We observe density maxima between the base pair planes at 5′-GpC and 5′-GpG steps. The maxima are located between two bases of a strand. In the case of the guanines, the N7 atom appears to be the dominant source of water localization in the major groove. The hydration on the cytosine strand is closer to the center of the base pair and involving mainly the N4 atom of the cytosine. The hydration of the sugar-phosphate backbone again does not show any considerable water localization at an underlying density threshold of 5.

B. Quantitative analysis of *B*-DNA hydration

Various quantities of relevance to the DNA hydration can be easily derived from the computed three-dimensional density distributions. For instance, cylindrical averaging yields the radial water-oxygen densities in the minor and major grooves of the *B*-DNA molecules $d(AATT)_3 \cdot d(AATT)_3$ and $d(CCGG)_3 \cdot d(CCGG)_3$ depicted in Fig. 3. The densities were calculated on a cylindrical grid with spacings $\Delta r = 0.015$ nm in the radial, $\Delta z = 0.015$ nm in the axial, and $\Delta\phi = 2\pi/100$ in the angular direction. The densities were averaged over an axial range of 4×0.338 nm covering the central four base pairs (AATT and CCGG, respectively) to avoid end effects. Global helix axes were obtained from diagonalization of the molecules' tensor of inertia, as discussed in Ref. [57]. The angular coordinates of the C1′ atoms are used to separate minor and major groove. As expected from Figs. 1 and 2, the surface of the minor groove is covered with pronounced high-density regions. In the case of the DNA with A·T base pairs, we observe two peaks reflecting the formation of a "spine of hydration" with water molecules covering the floor of the minor groove that are bridged by a second water layer. The peaks are followed by a minimum between 0.6 and 0.8 nm distance from the helix axis. The water density then gradually increases to its bulk value. (No correction for the excluded volume was applied in the angular averaging of the densities, which explains the density values below 1 despite strong localization.) The radial density distribution in the major groove does not show such a pronounced structuring and approaches its bulk value more continuously.

Integration of the radial densities yields the hydration numbers. Within 1.2 nm of the helix axis we find approximately 7.6 and 7.1 water molecules per base pair in the minor groove of $d(AATT)_3 \cdot d(AATT)_3$ and $d(CCGG)_3 \cdot d(CCGG)_3$, respectively. This reflects a somewhat stronger minor-groove hydration of A·T base pairs. The major groove, on the other hand, shows the

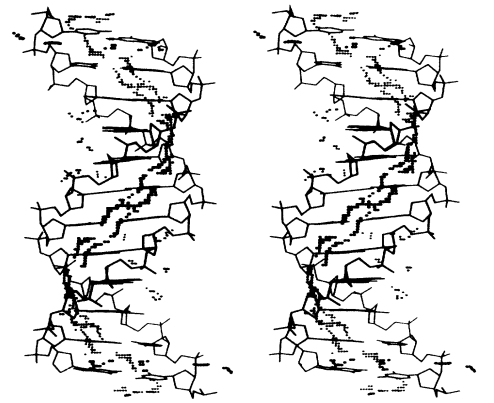

FIG. 2. Hydration of the *B*-DNA dodecamer with sequence $d(CCGG)_3 \cdot d(CCGG)_3$ and canonical structure in stereographic representation. Details as in Fig. 1.

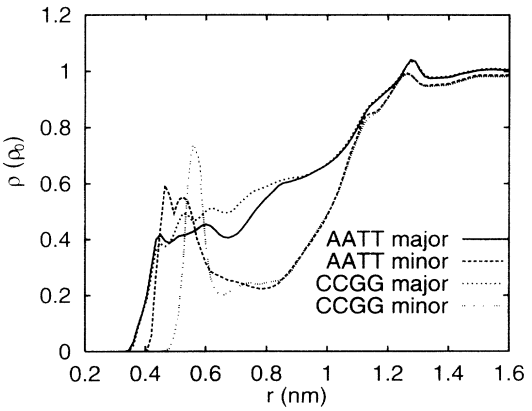

FIG. 3. Radial dependence of the water-oxygen density in the minor and major grooves of the *B*-DNA molecules $d(AATT)_3 \cdot d(AATT)_3$ and $d(CCGG)_3 \cdot d(CCGG)_3$. The densities were averaged in axial and angular directions over the central four base pairs and are plotted versus the distance from the helix axis in units of the bulk-water density ρ_0. (——), major-groove $d(AATT)_3 \cdot d(AATT)_3$; (– – –), minor-groove $d(AATT)_3 \cdot d(AATT)_3$; (- - - -), major-groove $d(CCGG)_3 \cdot d(CCGG)_3$; (·····), minor-groove $d(CCGG)_3 \cdot d(CCGG)_3$.

opposite trend with 18.6 and 19.3 water molecules per base pair within 1.2 nm of the helix axis for the molecules $d(AATT)_3 \cdot d(AATT)_3$ and $d(CCGG)_3 \cdot d(CCGG)_3$, respectively.

We now calculate the number of water molecules per base pair in the first hydration shell. We define two quantities $N(r)$ and $n(r)$. $N(r)$ is the average number of water molecules within a distance r of any polar atom on the DNA molecule; and $n(r)$ is the average water-oxygen density at accessible points ($\rho > 0$) at a distance r of the closest polar atom. $N(r)$ and $n(r)$ can be computed from the density distribution on a Cartesian grid for a volume extending over one symmetry unit in the axial direction (four base pairs for the sequences studied here). $N(r)$ and $n(r)$ are shown in Fig. 4, which also contains the oxygen pair correlation function $g^{(2)}_{OO}(r)$ of SPC water as a reference. The $n(r)$ curves have a strong first peak at about 0.275 nm distance with heights of 2.6 (AATT) and 2.5 (CCGG). They resemble closely the pair correlation function $g^{(2)}_{OO}$ of oxygens in water. However, the ice peak of $g^{(2)}_{OO}$ at $r = 0.45$ nm appears only as a series of small bumps in $n(r)$. For $r > 0.5$ nm, the $n(r)$ and $g^{(2)}_{OO}(r)$ curves are almost identical. The first minimum of $n(r)$ is located at $R = 0.35$ nm. The corresponding numbers of water molecules per base pair within R are 10.8 (AATT) and 10.7 (CCGG).

The experimental values for hydration numbers of B-DNA quoted in the literature vary considerably (see, e.g., Ref. [22]). The reason is that different experimental techniques measure different physical properties of hydration water; and as a consequence the extracted hydration numbers are not directly comparable. The most widely quoted hydration numbers are those of infrared spectroscopy experiments by Falk *et al.* [7,58]. These authors estimate the primary hydration shell of DNA to contain of the order of 10–13 water molecules per nucleotide, i.e., 20–26 per base pair. These numbers are significantly larger than those we obtained from $N(R)$ if R is the first minimum in $n(r)$. We find a corresponding number of about 25 water molecules per base pair if we integrate up to about $R = 0.47$ nm, i.e., if we partly include the water molecules bound to the first layer. But it is clear that there does not exist a simple choice of the distance parameter R that allows a direct comparison between the integrated density distributions $N(R)$ and the experimental data of Falk *et al.*

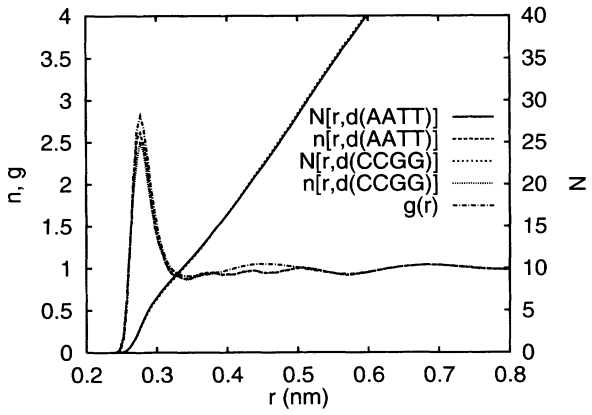

FIG. 4. Hydration number of the B-DNA molecules $d(AATT)_3 \cdot d(AATT)_3$ and $d(CCGG)_3 \cdot d(CCGG)_3$. $N(r)$ is the number of water molecules per base pair within a distance r of the closest polar atom on the DNA; $n(r)$ is the average water density in the accessible volume ($\rho > 0$) at a distance r of the closest polar atom. (——), $N(r, \text{AATT})$; (– – –), $n(r, \text{AATT})$; (- - - - -), $N(r, \text{CCGG})$; (· · · · · ·), $n(r, \text{CCGG})$; (— · —), oxygen pair correlation function $g^{(2)}_{OO}(r)$ of SPC water.

However, the different experimental techniques measuring hydration numbers determine the amount of water having physical properties different from bulk water. Strongly localized water molecules are expected to show the most pronounced deviations from bulk-water behavior. Therefore, we define a quantity $S(\bar{\rho})$,

$$S(\bar{\rho}) = \int_{\rho(\mathbf{r})>\bar{\rho}} d\mathbf{r}\, \rho(\mathbf{r}), \tag{8}$$

which is the number of water molecules in regions with a water density above a density threshold $\bar{\rho}$. $S(\bar{\rho})$ was calculated for a volume extending over the central four base pairs of the B-DNA dodecamers and is shown in Fig. 5 in units of water molecules per base pair. In terms of $S(\bar{\rho})$, the molecule with A·T base pairs shows considerably stronger hydration compared to C·G base-pair DNA, as expected from Figs. 1 and 2. The number of water molecules per base pair in regions with density above 10 are 8.3 (AATT) and 4.4 (CCGG). This observation of a stronger hydration (in terms of strongly localized water molecules) of A+T-rich DNA compared to G+C-rich DNA is in agreement with ultrasonic velocity measurements [14].

C. Comparison with computer simulations

A detailed comparison of our results with the published simulation studies is presently not possible because of the following reasons. First, to our knowledge none of the papers published so far reports a detailed computation

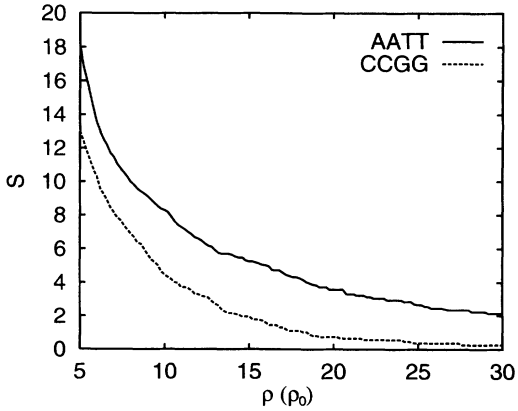

FIG. 5. Number of water molecules $S(\rho)$ per base pair in regions with water density above ρ. (——), $S(\rho, \text{AATT})$; (– – –), $S(\rho, \text{CCGG})$.

of the hydration (e.g., oxygen probabilities) on a lattice comparable to the one presented here. The snapshots of "typical" water configurations usually reported do not contain the information provided by the most quantitative modeling of hydration yielding probability maps like those computed here and also obtained in the x-ray experiments. Second, various groups use different empirical force fields and different amounts of water molecules ranging from 1 (simple energy computations) to about 2000; therefore, the simulations are not easily comparable among themselves. However, by and large the results of the present study agree qualitatively with those of previous studies concerning the most likely formation of "water spines" in the minor groove of A+T-rich structures and the less-marked hydration of G+C-rich segments.

In a recent computer simulation study, Forester and McDonald studied the hydration of a *B*-DNA decamer with sequence $d(CG)_5 \cdot d(CG)_5$ using molecular dynamics simulations of a system comprising one full turn of DNA and 950 SPC water molecules [59]. As in this work, the DNA was assumed to be rigid. The authors reported an average water-oxygen radial distribution function for the four hydrogens of the N4 and N2 atoms on the cytosine and guanine, respectively. In order to allow a quantitative comparison with their data, we calculated the radial water density distribution around each of the hydrogens at a discrete set of radii with intervals of 0.005 nm. For each distance a set of 50 000 points was chosen randomly on a sphere and the density was averaged over those points.

Figure 6 shows the radial water density distributions around each of the hydrogen atoms. The hydrogens H1 point towards the center of the C·G base pair and the hydrogens H2 are oriented towards the grooves (minor and major groove in the case of the N2 and N4 groups, respectively). The average radial distribution is compared with the corresponding curve taken from Fig. 5 of Ref. [59]. The radial distribution data of Forester and McDonald were rescaled with a factor of 0.79, since the authors used the actual water density in the simulation box to normalize their distributions rather than the bulk-water density. The PMF density expansion shows large variations between the radial distributions of the four individual hydrogen atoms. In particular, the H2's lying in the grooves show strong density peaks at a distance of 0.2 nm. The H1's, on the other hand, are involved in hydrogen bonds between the bases and show much weaker peaks at a distance of about 0.35 nm. However, the average distribution function of the PMF expansion agrees very well with the corresponding simulation curve of Forester and McDonald. The positions of the maxima and minima of the curves are reproduced by the PMF expansion. Some quantitative differences are observed for the peak heights. Reflecting the use of a slightly smaller oxygen van der Waals radius for the nitrogen atoms in the PMF expansion, the PMF curve shows nonzero values for distances about 0.02 nm closer than the simulation data. The height of the first minimum is in quantitative agreement with the simulation data; and we also observe a plateau beyond 0.4 nm distance, although at somewhat lower densities.

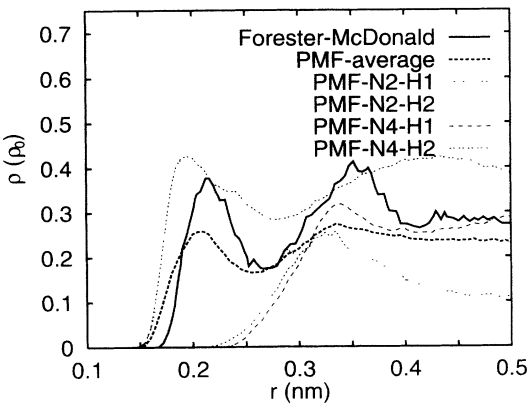

FIG. 6. Radial dependence of the water-oxygen density (in units of the bulk-water density ρ_0) around the hydrogens H1 and H2 of the N4 and N2 atoms of the cytosine and guanine, respectively, for the *B*-DNA molecule with sequence poly[d(CG)]·poly[d(CG)]. H2 atoms are pointing towards the grooves; H1 atoms are involved in the hydrogen bonds of the Watson-Crick base pair. Computer simulation: (——), average over the four hydrogens taken from Fig. 5 of Ref. [59]. PMF expansion results: (– – –), average over the four hydrogens; (- - - -), H1 of the N2 atom (guanine); (· · · · · ·), H2 of the N2 atom (guanine); (— · —), H1 of the N4 atom (cytosine); (- - - -), H2 of the N4 atom (cytosine).

The observation of good agreement for the hydration of the molecular subgroups — NH_2 justifies fundamental approximations made in this work: the substitution of nitrogens by oxygens; and that hydrogens were not explicitly considered. The small differences observed must also be seen in the light of possible statistical and methodological uncertainties in the simulations. As mentioned above, the force-field parameters employed in the simulations are empirical. In addition, the production part of the simulations of Forester and McDonald extended over only 81 ps, a time certainly too short compared to minor-groove residence times of about 1 ns measured in NMR experiments [17].

In Fig. 7, we compare the cylindrically averaged water density for two base pairs in the center of the dodecamer $d(CG)_6 \cdot d(CG)_6$ with the corresponding data of Fig. 6 of Ref. [59]. Unlike Fig. 3, the contributions of minor and major groove were not separated. We find good overall agreement between the computer simulation and PMF expansion curves. However, some differences are observed. In particular, the first peak in the simulation curve appears only as a shoulder and the second peak is too close. An interesting point is that the computer simulation curve does not approach the correct bulk limit of the water density for large r. This is a consequence of the strong inhomogeneity present in the DNA-water system because of the dense packing of DNA molecules under periodic boundary conditions.

Integration of the radial distributions yields the number of water molecules within a given distance from the helix axis. For a distance of 1.5 nm, the PMF expansion gives 54.8 water molecules per base pair. A numerical integration of the simulation curve (Fig. 6 of Ref. [59])

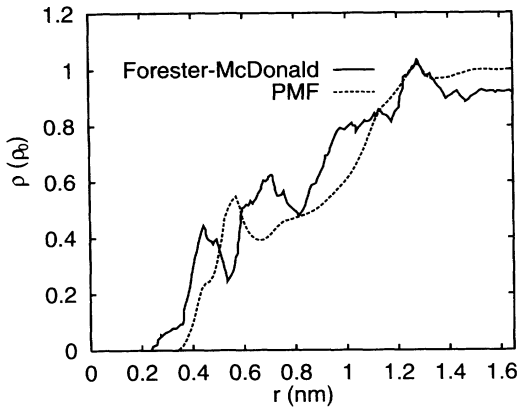

FIG. 7. Radial dependence of the cylindrically averaged water-oxygen density (in units of the bulk-water density ρ_0) for the B-DNA molecule with sequence poly[d(CG)]·poly[d(CG)]. r is the distance from the helix axis. Computer simulation: (——), water-oxygen density distribution taken from Fig. 6 of Ref. [59]. (– – –), PMF expansion result.

yields 56.4 water molecules per base pair, showing excellent agreement of the overall hydration numbers between the computer simulation and the PMF expansion method.

Figure 8 shows the radial water-oxygen distributions for the anionic oxygens O1 and O2 of the PO_4^- groups taken from Fig. 4 of Ref. [59]. As expected, the hydration of the negatively charged phosphate groups shows the most significant discrepancies between computer simulation and PMF expansion. The simulations show a very strong structuring with a high first peak. In the PMF expansion, the angular average with its considerable excluded volume effects gives only a comparably weak first peak in the radial distribution function, though at the right position. Clearly, the modeling of charged atomic

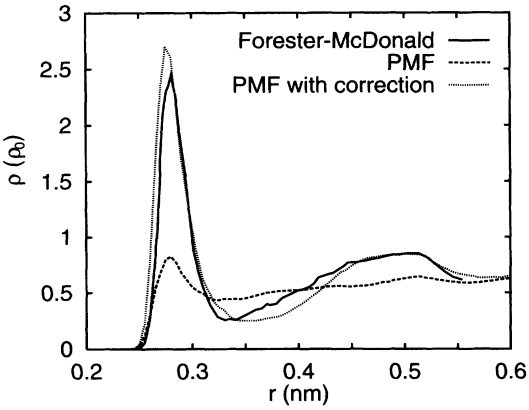

FIG. 8. Radial dependence of the water-oxygen density (in units of the bulk-water density ρ_0) around the oxygens O1 and O2 of the PO_4^- groups, averaged over the O1 and O2 atoms at 5′-CpG and 5′-GpC steps of the B-DNA molecule with sequence poly[d(CG)]·poly[d(CG)]. Computer simulations: (——), average over the four oxygens taken from Fig. 4 of Ref. [59]. PMF expansion results: (– – –), average over the four oxygens; (·····), corrected PMF expansion representing O1 and O2 as negative ions with charge $q = -0.85|e|$.

groups within the expansion formalism requires further refinement. As outlined in Sec. II, this can be accomplished by enlarging the database of correlations used in the calculations.

Figure 8 also shows results for the hydration of the anionic oxygens with a correction applied according to Eq. (5). This correction method illustrates how the PMF expansion method can be systematically extended to study the hydration of molecules with atomic groups carrying a net charge. Here, we studied the water-oxygen distributions around fictitious particles X representing O1 and O2. We approximated these atoms by Lennard-Jones particles ($\sigma = 0.33$ nm, $\epsilon = 0.62802$ kJ/mol [59]) carrying a fractional charge $q = -0.85|e|$ as used in force-field representations of macromolecules [60]. Lorentz-Berthelot mixing rules were applied to the Lennard-Jones interactions with SPC water [61]. The radial water-oxygen distribution function $g_{OX}^{(2)}$ was obtained from a Monte Carlo simulation of a single atom in SPC water. Since the two anionic oxygen atoms are very close on the PO_4^- group, the three-particle correction was calculated using the constraints method described in Sec. II. Two atoms X were kept at a distance of $t = 0.255$ nm and the conditional water-oxygen density $\rho_{O;XX}(r, s|t)$ was averaged as a function of the distances r and s from the two particles (bipolar coordinate system). Otherwise, the conditions and methods were those used in the bulk simulations [43]. In each of the two simulations, 200 000 Monte Carlo passes were used for averaging the distribution functions after extensive equilibration.

Each of the pairs of anionic oxygens on the DNA phosphate groups was then treated as a pair of X particles in the density expansion Eq. (5), resulting in a contribution $g_{O;XX}(r, s|t) = \rho_{O;XX}(r, s|t)/\rho_0$. (For r or s larger than 0.63 nm, the Kirkwood superposition $g_{OX}^{(2)}(r)g_{OX}^{(2)}(s)$ was used instead.) The corresponding result for the hydration of the phosphate groups is shown in Fig. 8. We now observe much closer agreement between the PMF expansion result and the simulation curve of Ref. [59]. In particular, the height of the first maximum is now in quantitative agreement with the simulation data. Also the first minimum and second maximum agree quantitatively. Only for distances r between 0.35 and 0.45 nm is the corrected PMF expansion curve slightly too low compared to the computer simulation result. It is evident that the correction results in a significant improvement of the description of the PO_4^--group hydration. However, a more detailed presentation of the results obtained using corrections to the PMF expansion method using only bulk-water correlations is planned for the near future, in conjunction with studies of the hydration of proteins, RNA, and other DNA families (A, Z).

For a comparison of quantities such as radial distribution functions around atomic sites on a macromolecule, it is important to keep in mind that these functions represent complex angular averages. Unlike simple atomic liquids, the average density on a sphere around an atom shows a strong dependence on the orientation. Therefore, the information content of radial distributions is very limited. In particular, excluded volume effects have a significant influence on radial distributions. This may result in

an exaggeration of certain features or may cause density peaks not to be resolved at all if they are strongly localized in space. It explains the densities of less than 0.3 times the bulk-water density at a distance of 0.5 nm from the — NH_2 hydrogens. However, it also reveals possible problems with these quantities in the case of the PMF density expansion formalism, in which at this stage the nonpolar atoms (carbon) are only represented as spheres, inside which the water density is set to zero. In future extensions it might prove necessary to use, for instance, correlation functions of methane in water. This would better take into account the excluded volume effect, resulting in a density increase near nonpolar atoms irrespective of the weak interactions with water. In particular, quantities involving averages over large volumes or areas might be affected by the presently rather crude description of the nonpolar atoms. However, from the results described above and the physical basis of the method, we expect the PMF expansion formalism to represent well the strongly localized hydrophilic hydration (i.e., the position of the density maxima), which is of major interest in biomolecular applications.

IV. CONCLUDING REMARKS

On the basis of the quality of the results obtained so far and the enormous computational efficiency (calculations like those described above require only about 25 min of CPU time on a workstation), we are convinced that the method proposed here provides the only feasible way presently to describe hydrophilic biomolecular hydration quantitatively of the large number of x-ray, NMR, and modeled structures relevant for molecular biology and biotechnology. Brute force simulations of comparable accuracy and detail require computation times 10^3–10^4 times longer. They are prone to serious statistical problems and involve several rather *ad hoc* assumptions (e.g., periodic boundary conditions, minimum image convention, truncation of potentials, ergodicity, etc.) of uncertain quality in the context of inhomogeneous polar and charged systems of interest to molecular biophysics. We view carefully done simulations as a very useful tool for extracting information necessary for the implementation of effective theoretical frameworks, but not as the panacea for understanding and modeling complex biophysical systems.

At its present state of development, the technique used here involves two central approximations: (i) the neglect of second-order effects owing to the difference between DNA oxygens and nitrogens and water oxygens as well as the assumed weak interactions of nonelectronegative atoms with water, and (ii) the truncation of the water correlations hierarchy at the triplet level. Concerning (i) we have given above physical reasons why this modeling is a good first-order attack on the problem; and, judging from the computational results obtained, we think that our physical intuitions are justified. However, we plan to investigate the matter in a series of computational studies and, if necessary, refine the method as outlined in Sec. III. Concerning (ii), we already performed a benchmarking study [43] for the case of the ice-I*h*–water interface using both the PMF approach and extensive Monte Carlo simulations. The results obtained completely justify both the necessity of including triplet correlations and the negligible role of higher-order correlations.

As mentioned above, hydration analysis of the other DNA conformations [44,52] as well as proteins [53] is under way and will be presented in the near future.

ACKNOWLEDGMENTS

We wish to thank P. Procacci, G. Corongiu, E. Clementi, A. E. García, R. Klement, T. Jovin, and M. Neumann for the collaborations and discussions that greatly contributed to the realization of the project. This work has been funded by the Bundesministerium für Forschung und Technologie, the Max-Planck-Society (both F.R.G), and the Department of Energy (U.S.).

[1] B. F. Luisi and P. B. Sigler, Biochim. Biophys. Acta **1048**, 113 (1990).
[2] F. A. Quiocho, D. K. Wilson, and N. K. Vyas, Nature **340**, 404 (1989).
[3] M. A. Rould, J. J. Perona, D. Söll, and T. A. Steitz, Science **246**, 1135 (1989).
[4] L. A. Marky and D. W. Kupke, Biochem. **28**, 9982 (1989).
[5] R. E. Franklin and R. G. Gosling, Acta Crystallogr. **6**, 673 (1953).
[6] A. G. W. Leslie, S. Arnott, R. Chandrasekaran, and R. L. Ratliff, J. Mol. Biol. **143**, 49 (1980).
[7] M. Falk, K. A. Hartman, Jr., and R. C. Lord, J. Am. Chem. Soc. **85**, 387 (1963).
[8] M. Falk, K. A. Hartman, Jr., and R. C. Lord, J. Am. Chem. Soc. **85**, 391 (1963).
[9] S. C. Erfurth, P. J. Bond, and W. L. Peticolas, Biopolymers **14**, 1245 (1975).
[10] M.-J. B. Tunis and J. E. Hearst, Biopolymers **6**, 1325 (1968).
[11] M.-J. B. Tunis and J. E. Hearst, Biopolymers **6**, 1345 (1968).
[12] M. Falk, K. A. Hartman, Jr., and R. C. Lord, J. Am. Chem. Soc. **84**, 3843 (1962).
[13] V. A. Buckin, B. I. Kankiya, A. P. Sarvazyan, and H. Uedaira, Nucleic Acids Res. **17**, 4189 (1989).
[14] V. A. Buckin, B. I. Kankiya, N. V. Bulichov, A. V. Lebedev, I. Y. Gukovsky, V. P. Chuprina, A. P. Sarvazyan, and A. R. Williams, Nature **340**, 321 (1989).
[15] S. A. Lee, S. M. Lindsay, J. W. Powell, T. Weidlich, N. J. Tao, and G. D. Lewen, Biopolymers **26**, 1637 (1987).
[16] V. T. Forsyth, A. Mahendrasingam, W. J. Pigram, R. J. Greenall, K. Bellamy, W. Fuller, and S. A. Mason, Int. J. Biol. Macromol. **11**, 236 (1989).

[17] E. Liepinsh, G. Otting, and K. Wüthrich, Nucleic Acids Res. **20**, 6549 (1992).
[18] M. G. Kubinec and D. E. Wemmer, J. Am. Chem. Soc. **114**, 8739 (1992).
[19] J. Texter, Prog. Biophys. Mol. Biol. **33**, 83 (1978).
[20] W. Saenger, *Principles of Nucleic Acid Structure* (Springer, Berlin, 1984).
[21] W. Saenger, Annu. Rev. Biophys. Biophys. Chem. **16**, 93 (1987).
[22] V. A. Buckin, Mol. Biol. **21**, 615 (1987) [Mol. Biol. (USSR) **21**, 512 (1987)].
[23] E. Westhof and D. L. Beveridge, in *Water Science Reviews*, edited by F. Franks (Cambridge University Press, Cambridge, England, 1990), Vol. 5, pp. 24–136.
[24] H. M. Berman, Curr. Opinion Struct. Biol. **1**, 423 (1991).
[25] H. R. Drew and R. E. Dickerson, J. Mol. Biol. **151**, 535 (1981).
[26] M. L. Kopka, A. V. Fratini, H. R. Drew, and R. E. Dickerson, J. Mol. Biol. **163**, 129 (1983).
[27] O. Kennard, W. B. T. Cruse, J. Nachman, T. Prange, Z. Shakked, and D. Rabinovich, J. Biomol. Struct. Dyn. **3**, 623 (1986).
[28] A. H.-J. Wang, T. Hakoshima, G. van der Marel, J. H. van Boom, and A. Rich, Cell **37**, 321 (1984).
[29] B. Chevrier, A. C. Dock, B. Hartmann, M. Leng, D. Moras, M. T. Thuong, and E. Westhof, J. Mol. Biol. **188**, 707 (1986).
[30] G. Corongiu and E. Clementi, Biopolymers **20**, 551 (1981).
[31] G. Corongiu and E. Clementi, Biopolymers **20**, 2427 (1981).
[32] E. Clementi and G. Corongiu, Biopolymers **21**, 763 (1982).
[33] P. S. Subramanian and D. L. Beveridge, J. Biomol. Struct. Dyn. **6**, 1093 (1989).
[34] F. Eisenhaber, V. G. Tumanyan, F. Eisenmenger, and W. Gunia, Biopolymers **28**, 741 (1989).
[35] V. P. Chuprina, U. Heinemann, A. A. Nurislamov, P. Zielenkiewicz, R. E. Dickerson, and W. Saenger, Proc. Nat. Acad. Sci. U.S.A. **88**, 593 (1991).
[36] A. V. Teplukhin, V. I. Poltev, and V. P. Chuprina, Biopol. **32**, 1445 (1992).
[37] D. M. Soumpasis, Proc. Nat. Acad. Sci. U.S.A. **81**, 5116 (1984).
[38] D. M. Soumpasis, J. Wiechen, and T. M. Jovin, J. Biomol. Struct. Dyn. **4**, 535 (1987).
[39] A. E. García and D. M. Soumpasis, Proc. Nat. Acad. Sci. U.S.A. **86**, 3160 (1989).
[40] R. Klement, D. M. Soumpasis, and T. M. Jovin, Proc. Nat. Acad. Sci. U.S.A. **88**, 4631 (1991).
[41] D. M. Soumpasis, A. Garcia, R. Klement, and T. Jovin, in *Theoretical Biochemistry and Molecular Biophysics*, edited by D. L. Beveridge and R. Lavery (Adenine Press, Schenectady, NY, 1990), Vol. 1, pp. 343–360.
[42] D. M. Soumpasis, in *Computation of Biomolecular Structures*, edited by D. M. Soumpasis and T. M. Jovin (Springer, Berlin, 1993), pp. 223–239.
[43] G. Hummer and D. M. Soumpasis, Phys. Rev. E **49**, 591 (1994).
[44] G. Hummer and D. M. Soumpasis, in *Structural Biology: The State of the Art*, edited by R. H. Sarma and M. H. Sarma (Adenine Press, Schenectady, NY, 1994), Vol. 2, pp. 273–278.
[45] A. Münster, *Statistical Thermodynamics* (Springer, Berlin, 1969), Vol. 1, p. 338.
[46] J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).
[47] I. Z. Fisher and B. L. Kopeliovich, Dokl. Akad. Nauk. SSSR **133**, 81 (1960) [Sov. Phys. Dokl. **5**, 761 (1960)].
[48] G. Hummer and D. M. Soumpasis, J. Chem. Phys. **98**, 581 (1993).
[49] D. M. Soumpasis, P. Procacci, and G. Corongiu, IBM DSD report, 1991 (unpublished).
[50] U. Niesar, G. Corongiu, E. Clementi, G. R. Kneller, and D. K. Bhattacharya, J. Phys. Chem. **94**, 7949 (1990).
[51] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, in *Intermolecular Forces: Proceedings of the 14th Jerusalem Symposium on Quantum Chemistry and Biochemistry*, edited by B. Pullman (Reidel, Dordrecht, 1981), pp. 331–342.
[52] G. Hummer, D. M. Soumpasis, and A. E. García, Biophys. J. **66**, A25 (1994).
[53] A. E. García, G. Hummer, and D. M. Soumpasis, Biophys. J. **66**, A130 (1994).
[54] S. Arnott and D. W. L. Hukins, Biochem. Biophys. Res. Commun. **47**, 1504 (1972).
[55] W. Saenger, W. N. Hunter, and O. Kennard, Nature **324**, 385 (1986).
[56] G. G. Privé, K. Yanagi, and R. E. Dickerson, J. Mol. Biol. **217**, 177 (1991).
[57] D. M. Soumpasis, C.-S. Tung, and A. E. García, J. Biomol. Struct. Dyn. **8**, 867 (1991).
[58] M. Falk, A. G. Poole, and C. G. Goymour, Can. J. Chem. **48**, 1536 (1970).
[59] T. R. Forester and I. R. McDonald, Mol. Phys. **72**, 643 (1991).
[60] S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, Jr., and P. Weiner, J. Am. Chem. Soc. **106**, 765 (1984).
[61] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 2nd ed. (Academic Press, London, 1986).