JCTC Journal of Chemical Theory and Computation

Another Coarse Grain Model for Aqueous Solvation: WAT FOUR?

Leonardo Darré,[†] Matías R. Machado,[†] Pablo D. Dans,[†] Fernando E. Herrera,[†,‡] and Sergio Pantano*,[†]

Institut Pasteur de Montevideo, Calle Mataojo 2020, CP 11400, Montevideo, Uruguay, and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Avda. Rivadavia 1917 - CP C1033AAJ - Cdad. de Buenos Aires, Argentina

Received July 7, 2010

Abstract: Biological processes occur on space and time scales that are often unreachable for fully atomistic simulations. Therefore, simplified or coarse grain (CG) models for the theoretical study of these systems are frequently used. In this context, the accurate description of solvation properties remains an important and challenging field. In the present work, we report a new CG model based on the transient tetrahedral structures observed in pure water. Our representation lumps approximately 11 WATer molecules into FOUR tetrahedrally interconnected beads, hence the name WAT FOUR (WT4). Each bead carries a partial charge allowing the model to explicitly consider long-range electrostatics, generating its own dielectric permittivity and obviating the shortcomings of a uniform dielectric constant. We obtained a good representation of the aqueous environment for most biologically relevant temperature conditions in the range from 278 to 328 K. The model is applied to solvate simple CG electrolytes developed in this work (Na^+, K^+, and Cl^-) and a recently published simplified representation of nucleic acids. In both cases, we obtained a good resemblance of experimental data and atomistic simulations. In particular, the solvation structure around DNA, partial charge neutralization by counterions, preference for sodium over potassium, and ion mediated minor groove narrowing as reported from X-ray crystallography are well reproduced by the present scheme. The set of parameters presented here opens the possibility of reaching the multimicroseconds time scale, including explicit solvation, ionic specificity, and long-range electrostatics, keeping nearly atomistic resolution with significantly reduced computational cost.

Introduction

Computer simulation of biological systems is continuously experiencing a tremendous expansion urged by the ever-growing computer power that allows for the treatment of always more complex systems and for time scales that continuously approach biological relevancy.[1] Parallel to this, the greediness to achieve structural and dynamical descriptions of yet longer and bigger sized systems has prompted the scientific community to develop simplified models of molecular assemblies that mimic arbitrarily intricate molecular systems with a lower degree of complexity. These simplified or coarse grain (CG) representations reduce significantly the computational demands but still capture the physical essence of the phenomena under examination.[2,3] Starting from the pioneering simplified models used to describe protein folding,[4,5] a huge number of successful applications covering a wide range of biomolecular and nanotechnologically relevant applications have been presented.[6−18] For an exhaustive review of this area, the book *Coarse-Graining of Condensed Phase and Biomolecular Systems*[19] is recommended. In this context, the accurate treatment of solvent effects is still a challenging issue. In fact, many CG approaches use a uniform dielectric constant, which may produce an incorrect partition of

* Corresponding author. Tel.: +598-2522 0910. Fax: +598-2522 0910. E-mail: spantano@pasteur.edu.uy.

[†] Institut Pasteur de Montevideo.

[‡] CONICET.

10.1021/ct100379f
Published on Web 11/04/2010

hydrophilic molecules in a hydrophobic medium. Recently, elaborated and/or systematic developments of CG models for simulating water, Hbond (hydrogen bond) bound, and/or ionic liquids with high accuracy have been presented.[20−23] Here, we present a new and simple CG model for water derived from elementary physicochemical concepts and fitting the interaction parameters to reproduce some characteristic features of liquid water. The main advantage of our model is that all of the interactions are described by a typical Hamiltonian for classical simulations, explicitly including long-range electrostatics. This model is composed of four interconnected beads arranged in a tetrahedral conformation (Figure 1). Each bead carries an explicit partial charge. In this way, the liquid generates its own dielectric permittivity, avoiding the use of a constant dielectric medium. The model achieves a reasonable reproduction of some common properties of liquid water in the range of temperatures relevant for most biological applications.

As examples of the potentiality of the model, we study first the solvation of CG monovalent electrolytes developed in this work (Na^+, K^+, and Cl^-). Then, we present molecular dynamics (MD) simulations of a recently published CG model for DNA.[24] This model was shown to provide nearly atomistic resolution information of the structure and dynamics of double-stranded DNA under the generalized Born model approach for implicit solvation. In this contribution, we present an extension of that model for explicit solvation.

We show that this CG scheme is able to reproduce solvation spines, electrolyte specificity, and cation-driven narrowing of the minor groove. These examples illustrate the usefulness of the model in incorporating electrostatic effects in a physiological medium, keeping the chemical details of the different ionic species within CG simulations and overcoming the drawbacks of implicit solvation.

Methods

Description of the Model. The underlying idea of the model is that, due to its molecular characteristics, pure water behaves as a structured liquid forming (among other structural arrangements) transient tetrahedral clusters.[25] These clusters are composed of a central water coordinated by four identical molecules that form an elementary tetrahedral arrangement (Figure 1A). In this arrangement, the central molecule is buried and unable to interact with any other particle outside of the cluster. Our working hypothesis is that, owing to the replication of this structure in the bulk, the central molecule of any tetrahedron can be taken into account implicitly passing from a highly packed (atomistic) to a more granular (CG) liquid (Figure 1B). Aimed at reproducing the structural organization of the liquid, we generated a molecular topology in which four "covalently bound" beads are placed on the geometric positions of four oxygen atoms at the corners of an ideal tetrahedron (hence, the name WAT-FOUR or WT4 for short). The proposed topology implies that within an elementary cluster, the Hbond interactions that hold together the atomistic liquid water are represented by spring constants linking four beads (Figure 1B). The interactions between elementary clusters are taken into account by normal vdW and electrostatic terms in the classical Hamiltonian (Table S1, Supporting Information). These forces reproduce the overall tetrahedral ordering of water, allowing the elementary clusters to diffuse freely.

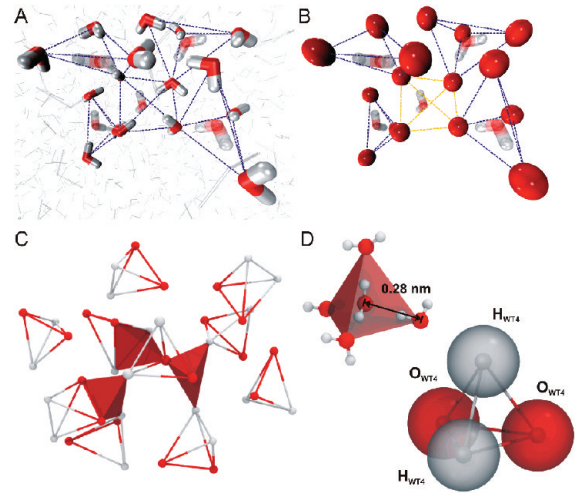

Figure 1. From atomistic to CG water. (A) Snapshot taken from a MD simulation showing the typical ordering of bulk water molecules. Gray molecules represent the liquid bulk. The structural organization is illustrated with a few opaque, thick water molecules which occupy the corners of irregular tetrahedrons. They saturate the Hbond capacity of a (semitransparent) molecule located in the center of each tetrahedron. Hbonds are indicated with dashed lines. (B) The positions of each of the oxygen atoms at the corners of the tetrahedra in A are now indicated with red beads. The concept behind the WAT FOUR (WT4) model is that those elementary tetrahedral clusters can be represented by four harmonically linked beads. The covalent bonds included in the WT4 model are represented by dark dashed lines, while intercluster interactions (vdW and electrostatics) are indicated with light dashed lines. The model implies that a number of water molecules are taken into account implicitly (represented as semitransparent molecules). Notice that water molecules can be implicitly represented even between noncovalently bound beads (take, for example, the central water molecule in the picture). The positions of all of the elements in A and B are identical. (C) Structural organization of WT4 in the bulk solution taken from a MD snapshot. The model reproduces higher-granularity tetrahedral organization in the space through noncovalent interactions. Red planes evidence the presence of rough tetrahedrons formed between different WT4 molecules comprising an implicit water molecule. (D) The ideal organization of a tetrahedral water cluster leads to the geometry of WT4. The separation of 0.28 nm between the water oxygen located at the center of the tetrahedron and its corners corresponds to the oxygen−oxygen (first neighbor) distance. This elementary cluster can be mapped to a WT4 molecule (bottom) composed by four harmonically bonded beads. White and red beads (hydrogen-like, H_{WT4}, and oxygen-like, O_{WT4}) carry positive and negative partial charges of 0.41*e*, respectively.

In analogy with the nearly tetrahedral water molecule that promotes a tetrahedral ordering in the surrounding space, a WT4 molecule recreates a roughly similar arrangement with a higher granularity (Figure 1C). Indeed, the structure of a WT4 molecule is replicated in its neighborhood, leaving holes that can be regarded as atomistic waters implicitly taken

Table 1. Interaction Parameters of the CG Models for Water and Ions[a]

					bond parameters	
	mass (au)	charge (e)	σ[b] (nm)	ε (kJ mol^{-1})	d_{eq} (nm)	K_{bond}[c] (kJ mol^{-1} nm^{-2})
SPC[27]	Ow:16 Hw:1	Ow:−0.82 Hw:+0.41	0.3166	0.650	0.1[d]	172500
TIP3P[28]	Ow:16 Hw:1	Ow:−0.8340 Hw:+0.4170	0.315061	0.6364	0.09572[d]	251208
WT4	O_{WT4}:50 H_{WT4}:50	O_{WT4}:−0.41 H_{WT4}:+0.41	0.42	0.55	0.45[e]	2092
NaW^+	130.99	1	0.58	0.55		
KW^+	147.1	1	0.645	0.55		
ClW^-	143.45	−1	0.68	0.55		

[a] The parameters of two common atomistic water models (SPC and TIP3P) are included for comparison. [b] Distance from the atomic center to the minimum of the vdW function. [c] Corresponds to a harmonic approximation of the form $E_{bond} = K_{bond}(d - d_{eq})^2$. [d] Hydrogen−oxygen distance. [e] Interbead distance.

into account by the CG scheme. These implicit waters can be present not only within the four bonded beads but also between tetrahedrons formed by beads belonging to a different molecule (Figure 1B and C). This suggests that the WT4 molecules in the bulk solution have the capacity to form interactions alike to Hbond networks.

The distance between the central oxygen of a tetrahedral water cluster (Figure 1D) and any other oxygen is ∼0.28 nm, as determined from diffraction experiments.[26] Taking this into account and the geometry of a perfect tetrahedron, the equilibrium distance between beads was set to 0.45 nm. The bond stretching force constant was set to mimic the interaction energy involved in typical hydrogen bonds. We tried harmonic constants within a range from 837 kJ/mol nm^2 to 4184 kJ/mol nm^2 (2 kcal/mol Å^2 to 10 kcal/mol Å^2). A value of 2092 kJ/mol nm^2 (5 kcal/mol Å^2) was chosen, as it results in a better fit of different water properties. This weak link confers the molecule a certain degree of structural plasticity, resulting in small deviations from a perfect tetrahedron upon temperature effects. These deformations could be identified with the nonperfect tetragonal ordering present in liquid water at room temperature. Given the tetrahedral symmetry, only these two bonded parameters for intramolecular interactions are needed (Table 1 and Figure 1D).

Intermolecular nonbonded interactions are ruled by normal van der Waals and electrostatic parameters, listed in Table 1. Partial charges were assigned considering that the central water molecule in a given atomistic tetrahedral cluster neutralizes the atomic charges of the waters in the corners by Hbond formation. If the water atomic charges are q for the hydrogen and $-2q$ for the oxygen, this yields two positive corners with charge q (alike to Hbond acceptors) and two negative corners with charge $-q$ (alike to Hbond donors, Figure 1D). The assignment of partial charges is a largely unsolved issue in classical force fields. In the particular case of water, this task has been addressed in many different ways, from adjusting parameters to reproduce experimental quantities in the liquid or gas phase to ab initio potentials derived from calculations using small clusters of molecules. However, no available model is capable of reproducing all of the water properties with good accuracy. Given the roughness of our model, we just sought to keep the electrostatic interactions engaged by CG beads comparable to atomistic Hbonds. Therefore, we simply tried the same atomic charge values used in common three-point water models (Table 1). Among several atomistic three-point water models tried, the charge distribution that better fit the experimental values was that of the SPC model.[27] The van der Waals radii and well deepness were used as free parameters. Intramolecular nonbonded interactions were excluded.

The mass of each bead was assigned to fit the density of liquid water. To this task, we used a computational box containing 497 WT4 molecules simulated at 300 K and 1 bar. The mass per bead necessary to match a density close to 1 g/mL resulted in 50 au. Taking into account that the mass of each atomistic water molecule is 18 au, it is implied that each WT4 bead represents ∼2.8 water molecules (50 au/18 au). This corresponds on average to about 11 real waters per WT4 molecule. Namely, we assume that each WT4 molecule represents 11 real water molecules in the CG scheme. Therefore, whenever we compare with physicochemical properties, a renormalization factor of 11 is taken into account (see below).

The packing factor of the WT4 spheres calculated as the volume of the cubic box that contains the WT4 molecules divided into the excluded volume of beads is ∼0.47, close to the 0.5 calculated for the SPC model. These values are significantly lower than the ideal 0.74 expected for the hexagonal closest packing (the maximum compaction for rigid spheres). This suggests that the bulk structure of WT4 leaves a number of interstitial cavities in a slightly higher proportion than in the SPC model.

CG Model for the Ions. Three ionic species were developed to represent, at the CG level, the hydrated states of Na^+, K^+, and Cl^- (hereafter called NaW^+, KW^+ and ClW^-, respectively).

Ions were developed considering that six water molecules are always attached to them[29] (i.e., roughly considering an implicit first solvation shell). Therefore, their masses were set as the sum of the ionic mass plus that of six water molecules (Table 1). Partial charges were set to unitary values. The van der Waals radii were adopted to match the first minima of the radial distribution function (RDF, also known as $g(r)$) of hydrated ions as obtained from neutron diffraction experiments.[30] The deepness of the well was set to the same values as the WT4 beads. This was done to ensure compatibility since when a WT4 molecule contacts a CG ion it interacts with its first solvation shell,

Table 2. Description of the Simulated Systems

system		water model	number of molecules	ionic species (number of ions)[a]	solute	temperature (K)	simulation time (ns)	ionic pair concentration (M)
AA[b]	S_1^{AA}	SPC	2483[c]			278−323	45	
AA	S_2^{AA}	SPC	5368[c]	Na^+(1) Cl^-(1)		300	20	0.01
AA	S_3^{AA}	SPC	5368[c]	K^+(1) Cl^-(1)		300	20	0.01
AA	S_4^{AA}	TIP3P	2483[c]			278−323	45	
AA	S_5^{AA}	TIP3P	7612[c]	Na^+(34) Cl^-(34)		300	15	0.5
CG[d]	S_1^{CG}	WT4	497[e]			300	100	
CG	S_2^{CG}	WT4	497[e]			278−328	200	
CG	S_3^{CG}	WT4	268[e]			300	3	
CG	S_4^{CG}	WT4	268[e]			300	3	
CG	S_5^{CG}	WT4	473[e]	NaW^+(1) ClW^-(1)		300	100	0.01
CG	S_6^{CG}	WT4	473[e]	KW^+(1) ClW^-(1)		300	100	0.01
CG	S_7^{CG}	WT4	456[e]	NaW^+(44) ClW^-(44)		300	30	0.5
CG	S_8^{CG}	WT4	456[e]	KW^+(44) ClW^-(44)		300	30	0.5
CG	S_9^{CG}	WT4	174[e]	NaW^+(7) ClW^-(7)		300	200	0.2
CG	S_{10}^{CG}	WT4	174[e]	KW^+(7) ClW^-(7)		300	200	0.2
CG	S_{11}^{CG}	WT4	170[e]	NaW^+(11) ClW^-(11)		300	200	0.3
CG	S_{12}^{CG}	WT4	170[e]	KW^+(11) ClW^-(11)		300	200	0.3
CG	S_{13}^{CG}	WT4	655[e]	NaW^+(34) ClW^-(34)		300	100	0.5
CG	S_{14}^{CG}	WT4	655[e]	KW^+(34) ClW^-(34)		300	100	0.5
CG	S_{15}^{CG}	WT4	506[e]	NaW^+(19) KW^+(19) ClW^-(16)	CG-DNA	300	4000	0.15[f]

[a] Parameters from Berendsen et al.[44] and van Gunsteren et al.[45] In system S_5^{AA}, the CHARMM PARAM27 parameters[46] were used. [b] AA: all atoms. [c] Atomistic water molecules. [d] CG: coarse grain. [e] WT4 molecules. [f] Not considering 22 neutralizing counterions.

which is implicitly considered. A list of nonbonded interaction parameters for the CG monovalent ions is detailed in Table 1.

CG Model for DNA. The CG system used for DNA was essentially the same as that previously presented by us.[24] This CG model reduces the complexity of the atomistic picture to six beads per nucleobase (see Supporting Information Figure S1 for the coarse graining scheme). This mapping keeps the "chemical sense" of specific Watson−Crick recognition allowing the 5′−3′ polarity. Similarly to the approach taken here for water and ions, molecular interactions are evaluated using a classical Hamiltonian. The beads used in this representation carry partial charges, which permits the use of explicit electrostatics

Minor changes have been introduced to the interaction parameters to improve the stability of the double strand using a time step of 20 fs. Back mapping of the atomic coordinates during the trajectory permitted an evaluation of the overall structural quality of the DNA dodecamer in terms of helical parameters (Supporting Information Figure S2). This new parameter set reproduces equally well the structural features of the double-stranded helix.

The complete set of new parameters for DNA is listed in Supporting Information Table S1.

A similarity index between the present implementation and that using the GB model for implicit solvation was calculated from the covariance matrices obtained from the trajectories performed in the present work and that performed in Dans et al.[24] for the Drew−Dickerson dodecamer.

Molecular Dynamics. MD simulations were performed using Gromacs 4.0.5[31−34] in the NPT ensemble unless otherwise stated. The temperature was coupled using the Nose−Hoover thermostat,[35,36] while pressure was kept at 1 bar by means of a Parrinello−Rahman[37,38] barostat, with coupling times of 1 and 5 ps, respectively. A cutoff for nonbonded interactions of 1.2 nm was used, while long-range electrostatics were evaluated using the Particle Mesh Ewald approach.[39,40] A time step of 2 fs was used in all-atom (AA) simulations, while in the CG simulations the time step was set to 20 fs. In order to ensure that the use of such a relatively long integration step does not introduce energy conservation problems, we performed a series of simulations at constant energy (NVE ensemble) using such a time step and varying the cutoff. For an acceptable accuracy in the integration of the equations of motion, one should expect the fluctuations of the total energy to be lower than one-fifth (20%) of the kinetic or potential energy components of the system.[41] According to our results, this criterion is well fulfilled with total energy fluctuations representing 5% of potential or kinetic energy fluctuations, using cutoff values of 1.0, 1.2, and 1.5 nm (Supporting Information Table S2). It was decided to use a cutoff of 1.2 nm, which besides ensuring energy conservation also includes direct nonbonded interactions up to the second neighboring WT4 molecule in solution. Additionally, NVT simulations were performed for some systems in order to compute the WT4 surface tension and the ionic osmotic pressure as detailed below.

All of the interactions (i.e., WT4−WT4, WT4−ion, ion−ion, ion−DNA, WT4−DNA, and DNA−DNA) were straightforwardly calculated within the pairwise Hamiltonian of Gromacs 4.0.5, which is common to many popular MD packages. The van der Waals cross interactions were calculated using the Lorentz−Berthelot combination rules.

Five atomistic (S_{1-5}^{AA}) and 15 CG systems (S_{1-15}^{CG}) were constructed to evaluate different properties of interest (see Table 2). Atomistic simulations were used to obtain reference properties to be compared with the CG models for water and ions. Systems S_1^{AA} and S_4^{AA} were used to compute density and diffusion coefficient profiles in a relevant range of temperatures (see Table 2). The temperature scan was carried out raising the reference temperature by 5° in steps of 5 ns.

Both radial distribution functions (ion−Ow) and electrostatic potential (on the line connecting both ions) were

calculated from systems S_2^{AA} and S_3^{AA}, where the cation−anion distance was kept fixed at 3.6 nm during the whole simulation. This last property was also calculated for system S_1^{AA} at room temperature in order to use it as a reference state for pure water. System S_5^{AA} was used to validate the methodology for measuring the osmotic pressure (described in the Supporting Information).

Regarding the CG simulations, bulk water properties under room conditions were obtained from system S_1^{CG}. The behavior of the model in the range of temperatures from 278 to 328 K was assessed using system S_2^{CG}. The temperature scan was carried out as in the corresponding atomistic simulations (S_1^{AA} and S_4^{AA}) but using time windows of 20 ns instead of 5 ns.

Surface tension and isothermal compressibility at the CG level were computed from systems S_3^{CG} and S_4^{CG}, respectively, according to the following steps, as proposed elsewhere.[42] First, an initial configuration at 300 K and 1 bar (generated by a short NPT equilibration of a simulation box containing 268 WT4 molecules) underwent a 0.1 ns equilibration in the NVT ensemble. The resulting configuration was used, on one hand, to construct system S_3^{CG} by adding vacuum slabs above and below the water bulk, so the box length in the z direction was tripled. A 3 ns production NVT simulation was conducted in such a system at 300 K, from which the surface tension was computed from the pressure tensors:

$$\gamma = \frac{L_z}{2}\left\langle P_{zz} - \left(\frac{P_{xx} + P_{yy}}{2}\right)\right\rangle \quad (1)$$

On the other hand, the NVT equilibrated configuration was also used as the starting structure (system S_4^{CG}) for a 3 ns NPT simulation at 300 K and 1 bar, from which the isothermal compressibility was computed according to[43]

$$\kappa = \frac{\langle V^2\rangle - \langle V\rangle^2}{\langle V\rangle k_B T} \quad (2)$$

Radial distribution functions (CG ion−WT4) and an electrostatic potential profile (obtained in the same way as in the atomistic system) were calculated for systems S_5^{CG} and S_6^{CG} and compared with systems S_2^{AA} and S_3^{AA}, respectively, in order to assess the ability of the CG model to reproduce atomistic results.

Systems S_7^{CG} and S_8^{CG} were used to compute radial distribution functions (CG ion-WT4) using an ionic concentration of roughly 0.5 M, in order to compare them with experimental data.[30]

Bjerrum (λ_B) and Debye (κ^{-1}) lengths were calculated as

$$\lambda_B(\rho) = \frac{(1.0 \times 10^9)\beta\, e^2}{4\pi\varepsilon_0\, \varepsilon_r(\rho)} \quad (3)$$

$$\kappa^{-1}(\rho) = \left(\frac{2(1.0 \times 10^{-15})F^2\rho}{RT\varepsilon_0\, \varepsilon_r(\rho)}\right)^{-1/2} \quad (4)$$

where β is the thermal energy, $\varepsilon_0 = 8.85 \times 10^{-12}$ C^2 J^{-1}m^{-1}, $F = 96485.3399$ C mol^{-1}, and $R = 8.314472$ J mol^{-1} K^{-1}.

The dielectric constants of the electrolyte aqueous solutions (Na^+Cl^- and K^+Cl^-) at different concentrations (0.2 and 0.3 M) were obtained from simulations of systems S_{9-12}^{CG} (see Table 2).

The osmotic pressure measurement was based on the methodology presented by Roux and Luo[47] (systems $S_{13,14}^{CG}$). The idea behind it is to simulate an aqueous solution where the ions are restrained to stay only in one-half of the simulation box and from the force exerted by the restraints, calculate the osmotic pressure. To accomplish this, we used a restraining strategy previously developed in our group called BRIM[48] (see the Supporting Information for a more exhaustive explanation).

Finally, we performed a 4 μs unconstrained simulation (S_{15}^{CG}) of a CG version[24] of a double-stranded DNA using the Drew−Dickerson sequence 5′-d(CGCGAATTCGCG)-3′ in an octahedron box filled with WT4 and CG ions (see Table 2). Global DNA structural behavior, DNA hydration, and specific DNA−water and DNA−ion interactions were evaluated. Helical parameters for DNA were computed using the Curves+ software.[49] The cation-induced narrowing of the minor groove was studied. Such structural changes were estimated from the average interphosphate distance between opposite strands measured for the following pairs: {(*5*, 24), (*6*, 23), (*7*, 22), (*8*, 21), (9, *20*), (10, *19*), (11, *18*), (12, *17*)} (italics indicate the residue numbers at the AT track). Cations were considered to be bound to the minor groove if their distance to the phosphate groups of both opposite strands was below 0.5 nm.

Results

In the following paragraphs, we describe the performance of the WT4 model to reproduce some common parameters of pure water. Comparisons are made, whenever possible, against experimental data. However, some of the calculated properties are also confronted with the results obtained from popular atomistic water models just to provide a reference frame for our results against well established AA models used by the broad scientific community. Subsequently, we analyze the solvation structure of simple electrolyte representations. Finally, to provide an example of application to a biologically relevant system, we briefly present a simulation of a CG DNA double helix in the presence of explicit solvent and mixed salts. A more detailed study on different properties of DNA (flexibility, breathing, DNA−solvent interactions on the multi-microsecond time scale, etc.) will be published elsewhere.

WT4 in the Bulk. A characteristic feature of water is its intrinsic ordering. A good reproduction of the oxygen−oxygen radial distribution function is a common goal for most water models in atomistic detail. The shape of the radial distribution function (RDF) at points far from the first spheres of hydration may furnish an idea of the liquid character of the substance under study. While for a liquid the RDF is expected to converge to a unitary value after a certain point, repetitive behavior is indicative of a crystalline state.

Although the RDF obtained with our model retains some characteristic features of liquid water, comparison of the RDF

obtained for WT4 with other atomistic models reveals some dissimilarities. The most evident difference with respect to the RDF calculated for SPC or TIP3P simulations (systems S_1^{AA} and S_4^{AA}) is the complete lack of the first solvation peak. Owing to the size and topology of the beads, WT4 presents a void space from the center of each bead up to the distance corresponding to the second solvation shell of real water. In this sense, the WT4 representation can be considered a second shell solvation model. In fact, the position of the first maximum in WT4 corresponds to the second peak of atomic water[50] (Figure 2A). It is important to notice that the normalization to the bulk value and the more granular character of the CG model generates a difference in the relative heights of the probability distribution of WT4 with respect to real water. Furthermore, the harmonic bonds existing within the tetrahedron translate into an overestimation of the probability of finding the first neighbor in the WT4 solution. After this global maximum, the relatively large size of WT4 generates some residual ordering that extends up to ~1.2 nm. The radial distribution function converges to one (bulk density) beyond 1.3 nm.

An important property for models of liquid water is their capability to reproduce the correct water diffusion. Clearly, the diffusivity of the WT4 molecules is much lower than that of atomistic water. At 300 K, we obtained a value of 2.03×10^{-6} cm^2 s^{-1}. However, the displacement of a WT4 molecule implicitly represents the movement of the center of mass of ~11 water molecules. Taking into account that the average mean squared displacement of the center of mass of N molecules is N times slower than the average mean squared displacement of N molecules diffusing separately,[53,54] we can conclude that the self-diffusion coefficient for the water molecules represented by the CG model at room temperature is 2.23×10^{-5} cm^2 s^{-1}, which is in good agreement with the experimental value (Table 3).

The WT4 model includes the explicit treatment of the electrostatic interactions as each bead carries a point charge (Table 1). This gives rise to a dielectric permittivity without imposing a continuum dielectric medium. The dielectric permittivity simulated by WT4 is 110.[55] Although this value is nearly 30% higher than that of real water, it must be noticed that this has been a problematic point even for more sophisticated atomistic models of water, and values ranging from 53[56] to 116[57] have been reported.

An important issue regards the long-range ordering of the WT4 molecules. In fact, some CG models for water present a freezing point very close to room temperature.[41] Therefore, we sought to perform a temperature scan over a range from 278 to 328 K. This range of temperatures covers most of the potential and biologically relevant applications of the model. Calculation of the RDF along the studied temperature range suggests that WT4 retains its liquid character, as no significant changes are found between 278 and 328 K (Figure 2A, inset).

The density of the WT4 model was set to match the value of pure water at 300 K. However, a reasonably good reproduction of the variations of the density versus temperature is also desirable. From the qualitative point of view, we obtained the expected reduction of the density with the

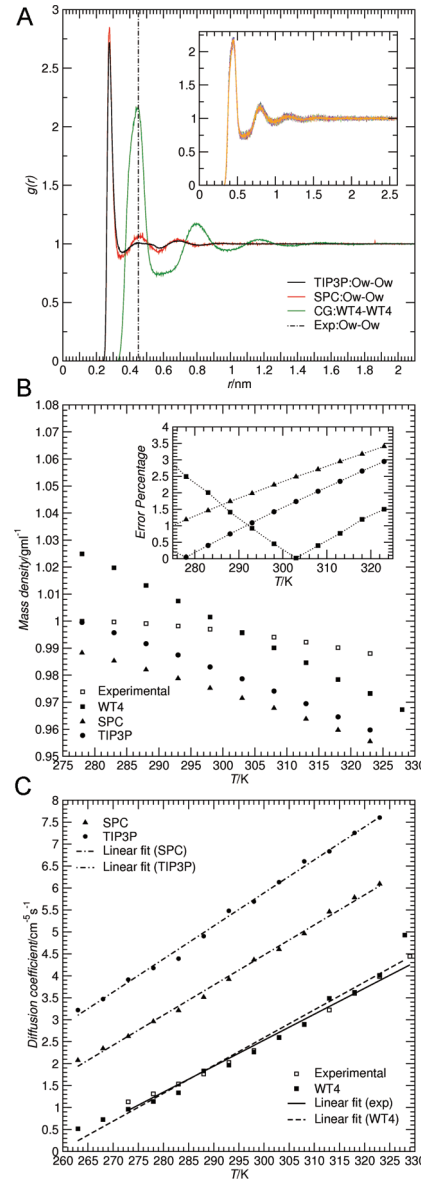

Figure 2. Bulk properties of WT4. (A) RDF calculated over all of the WT4 beads at room temperature from system S_1^{CG} (green line). Comparison is made with the oxygen−oxygen RDF calculated from TIP3P and SPC atomistic simulations as obtained from systems S_1^{AA} and S_4^{AA} at 298 K (black line and red line, respectively). The position of the second solvation peak obtained from experiments[50] is also shown (vertical, dot-dashed line). The inset shows the behavior of the RDF upon temperature variations (system S_2^{CG}) in the range from 278 to 328 K. No significant changes are observed. (B) The variation of the CG water mass density with the temperature (filled squares) calculated from system S_2^{CG} as compared with experimental data (empty squares)[51] and simulations of SPC (triangles) and TIP3P (circles) systems (S_1^{AA} and S_4^{AA}, respectively). The inset shows the relative error of the WT4, SPC, and TIP3P models compared to the experimental data. (C) The dependence of the diffusion coefficient on temperature is compared between the WT4, SPC, and TIP3P models (S_2^{CG} (filled squares), S_1^{AA} (triangles), and S_4^{AA} (circles), respectively) and experimental data[52] (empty squares). All four profiles present an almost linear trend, as revealed by the corresponding linear fits.

Table 3. Bulk Water Properties at Room Conditions for Atomistic Water Three-Point Models (SPC and TIP3P), WT4, and Experimental Data

	dielectric constant	diffusion coefficient (10^{-5} cm^2 s^{-1})	expansion coefficient (10^{-4} K^{-1})	mass density (g mL^{-1})	number density[a] ($\times$ 10^{22} mL^{-1})	surface tension (mN m^{-1})	isothermal compressibility (GPa^{-1})
WT4	110	2.23	11.6	1.0001	0.3	17	2.43
SPC	65[58]	3.85[59]	7.3[60]	0.9705[61]	3.2	53.4[62]	0.53[63]
TIP3P	82[56]	5.19[59]	9.2[64]	1.002[64]	3.4	49.5[62]	0.58[63]
Exp.	78.4[65]	2.27[66]	2.53[51]	0.9970[51]	3.3	71.2[67]	0.46[68]

[a] Calculated from the corresponding mass density, considering the molar mass of water (18 g mol^{-1}) and WT4 (200 g mol^{-1}). Accordingly, the number density for the atomistic models and real water corresponds to the number of water molecules per milliliter, while for WT4 it corresponds to the number of WT4 molecules (~11 water molecules) per milliliter.

temperature and an almost perfectly linear behavior of the system's density against temperature in the explored range (Figure 2B). Although the functional dependence of real water against temperature is certainly not linear, it is a good approximation within the temperature range chosen. In fact, the relative error of the WT4 density with respect to that of the real water in this temperature window remains always below 3%, with the higher deviations near the critical point of real water (Figure 2B). This behavior is comparable with those of the SPC and TIP3P atomistic models (Figure 2B).

Following the volume changes upon thermal variations at constant pressure allows also for the calculation of the isobaric expansion coefficient of our model. We obtain an overestimation of this quantity at 298 K (Table 3). The expansion coefficient of WT4 gives a value of 11.6×10^{-4} K^{-1} as compared with the experimental value of 2.53×10^{-4} K^{-1}.[51] Although overestimated, it is comparable with the values reported for widely used three-point water models (Table 3).

Another frequently calculated property for CG models is the surface tension. In our case, we obtained a value of 17 mN m^{-1}, which is nearly 4 times smaller than the experimental value. Similarly, we found a 5 fold higher isothermal compressibility as compared with the experimental value (Table 3). These discrepancies are very frequently found in CG models that lump a number of water molecules into one single entity.[42] The origin of this effect may be the loss of fully atomic interactions that decrease the cohesive forces and increase the granularity of the system.

A more stringent test for our representation comes from the calculation of the diffusion coefficient. Clearly, a rise in the diffusion must occur upon heating. Experimental data indicates that pure water experiences a nearly linear increase in the diffusion coefficient between 278 and 328 K. The model shows the correct dependence of the diffusion coefficient on temperature. Indeed, it shows good agreement with the experimental behavior within the explored range (Figure 2C).

Taking into account the above results, the range of validity of the model may be delimited by the following considerations: the lower limit should not go below 278 K. Applications at lower temperatures are strongly discouraged since ice formation implies quantum effects that can, obviously, not be achieved by simplified models. On the upper limit, the relative error in the renormalized diffusion coefficient arrives at ~11% at 328 K, suggesting that simulations at higher temperatures could require some reparameterization to keep the accuracy at acceptable levels.

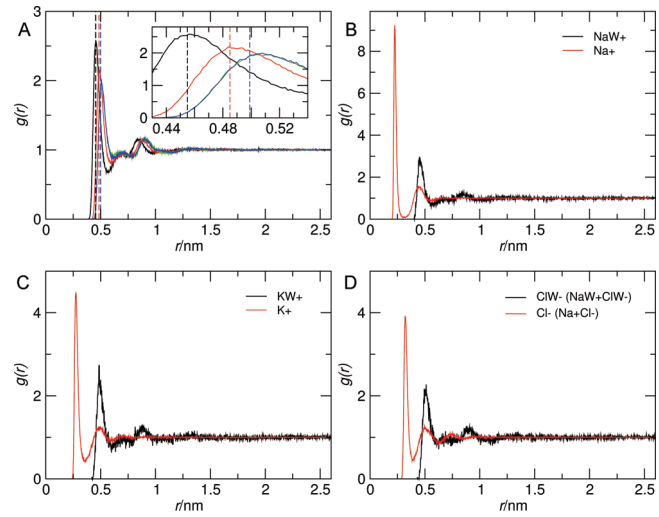

Figure 3. Ionic solvation. (A) The RDF of WT4 around CG electrolytes computed for systems S_7^{CG} and S_8^{CG} (NaW, black; KW, red; ClW, blue for NaW^+ClW^- and green for KW^+ClW^-). Vertical dashed lines indicate the position of the second solvation peak as determined from neutron scattering experiments.[30] The inset shows a closeup on the region between 0.43 and 0.54 nm allowing for a more precise comparison. (B, C, and D) Comparison between RDFs obtained from atomistic and CG simulations (systems S_2^{AA}, S_3^{AA}, S_5^{CG}, and S_6^{CG}). The plot corresponding to the solvation structure around chlorine ions in the presence of potassium is similar to that shown for the case of sodium. It is omitted for brevity.

Ionic Solvation. The characteristics of the WT4 model open the possibility to study the solvation properties of systems in which electrostatics are dominant. In this context, we developed the CG parameters of three simple electrolytes: Na^+, K^+, and Cl^-. Since we can imagine WT4 as a second solvation shell model, we represent ions together with their first sphere of hydration. Aimed at exploring the solvation structure generated by the WT4 model on the CG ions, simulations were conducted at roughly similar ionic concentrations to those reported in neutron diffraction experiments.[30] As depicted in Figure 3A, there is good correspondence, especially for the cations, between the first solvation maximum found for WT4 and the second hydration shell estimated from the experimental data.[30] A second solvation peak is found at nearly 0.9 nm, which has to be considered mainly as an artifact of the geometry of the model since the beads located in the last peak are harmonically linked to those

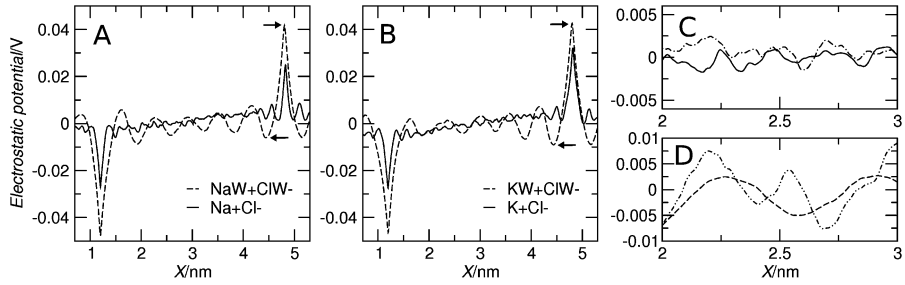

Figure 4. Profiles of electrostatic potential. (A) Electrostatic potential calculated along the line connecting the ionic pairs Na^+Cl^- (filled line, S_2^{AA}) and NaW^+ClW^- (dashed line, S_5^{CG}). Arrows indicate the points where the differences in the electrostatic potential were calculated. (B) Same as A for systems K^+Cl^- (S_3^{AA}) and KW^+ClW^- (S_6^{CG}). (C) Comparison of the electrostatic potential between the central portion of panel A (filled line, S_2^{AA}) against the analogous quantity calculated along a box containing pure SPC water (dot-dashed line, S_1^{AA}). (D) Same as C but for the CG systems (dashed line, S_5^{CG}, and double-dot-dashed line, S_1^{CG}).

of the first. After that point, the RDF converges to the bulk value in all cases.

Unfortunately, experimental data for ionic solvation is only available at high electrolyte concentrations. To explore lower (and more physiological) concentrations for which no experimental data are available, we tried a comparison with atomistic simulations confronting systems S_2^{AA} and S_3^{AA} with systems S_5^{CG} and S_6^{CG}, respectively, both having an ionic concentration of 0.01 M (Figures 3B, C, and D).

In close analogy with the case of pure WT4, the RDF of WT4 around CG ions shows a complete lack of the first solvation shell. A good reproduction of the position of the second solvation peak is observed, confirming the behavior of WT4 as a second solvation shell solvent. As expected, WT4 is not able to reproduce the third solvation shell. The relevance of this inaccuracy is, however, uncertain and could only be relevant in the case of chlorine ions, where such a shell is slightly more pronounced.[30]

Electrostatic Potential. Having analyzed the hydration structure of simple electrolytes, we turned our attention to the profiles of electrostatic potential and the screening properties. This was done by comparing the results of systems S_2^{AA} and S_3^{AA} against those of S_5^{CG} and S_6^{CG}, respectively. These systems consist of an ionic pair of Na^+Cl^- (or K^+Cl^-) kept at a fixed position during the simulation. The separation between both ions was 3.6 nm. Atomistic ionic pairs were immersed in a computational box containing an equivalent number of water molecules. This setup allowed us to compare under similar conditions atomistic and CG simulations as well as the behavior of the different ionic species. In order to assign the proper weight to the perturbations introduced by the electrolytes, we also made comparisons with the fluctuations produced by pure solvent (atomistic and CG) in the profiles of the electrostatic potential. In this way, it is possible to separate the observed features into two components: intrinsic bulk fluctuations and ionic perturbations. Furthermore, this approach gives an idea about the relaxation of the ionic potential at increasing distances from the ion and compares it with pure water and electrolyte solution.

A comparative view of the atomistic versus CG simulation can be acquired from Figure 4A. The first noticeable difference regards the height of the peaks centered on the positions of the ions. Owing to its smaller size, the SPC waters can get closer to the atomistic ion generating a more pronounced electrostatic screening. In the CG counterpart, the corresponding first solvation shell, which is implicit in the NaW^+ and ClW^- ions, only serves to create a void space without screening properties. This translates to a higher electrostatic potential induced by the CG ion. The implicit consideration of the first solvation shell in the CG ions implies that the first minimum observed in the atomistic system is absent in the CG system (Figure 4A). Furthermore, the position of the first minimum observed in the CG simulation (second solvation shell) roughly corresponds to the position of the second minimum around the ion in the atomistic system. Clearly, this effect derives from the solvation structure around the electrolytes; i.e., the first and second minima around the position of the ions (both, Na^+ and Cl^-) shown in Figure 4A correspond to the position of the oxygen atom in the first and second solvation shells shown in Figure 3B and D. Similar features are observed for the cases of K^+Cl^- and KW^+ClW^- ionic pairs (Figure 4B).

The distinctive characteristics of both cations evidenced by the solvent organization around NaW^+ and KW^+ (Figure 3A) can also be obtained from the calculation of the difference in electrostatic potential measured at the position of the cation with respect to that of its first minima (Figure 4A,B). This difference was about 10% higher for the case of KW^+ with respect to NaW^+, in qualitatively good agreement with the ~25% obtained from the atomistic case. This behavior may reflect the fact that water around potassium is bound in a more disorderly fashion than around sodium,[29] probably generating a less marked electrostatic screening in the case of potassium.

As seen from Figure 4A and B, the CG scheme presents higher fluctuations in the potential than the atomistic system. Aimed at excluding the possibility of a spurious ordering of WT4 molecules around the electrolytes, we compared the perturbations in the electrostatic potential introduced by the ions against those observed for pure solvent (both, atomistic and CG). This was assessed by computing the electrostatic potential along an arbitrary axis in two simulation boxes containing pure SPC and WT4 (systems S_1^{AA} and S_1^{CG},

Table 4. Thermodynamic Properties of Electrolyte Solutions

		Bjerrum length (nm)		Debye length (nm)		osmotic pressure[a] (bar)
ρ (molarity)		0.2 M	0.3 M	0.2 M	0.3 M	0.5 M
NaW^+ClW^-/WT4		0.57	0.61	0.76	0.6	35 (s.d. 15)
KW^+ClW^-/WT4		0.55	0.61	0.78	0.6	33 (s.d. 16)
Exp.[b]	NaCl	0.75	0.77	0.66	0.53	~25 (taken from Roux and Luo[47])
	KCl	0.74	0.76	0.67	0.54	

[a] The value obtained in the atomistic simulations using CHARMM PARAM 27 was 37 (s.d. 9) bar. [b] The function $\varepsilon_r(\rho) = \varepsilon(0)/(1 + A\rho)$ (NaCl, $A = 0.27$; KCl, $A = 0.24$), which results from fitting to experimentally obtained dielectric constants,[69] was used to estimate $\varepsilon_r(\rho)$ at the desired concentration, which is necessary for the computation of both Bjerrum and Debye lengths.

respectively). Superposition of both profiles (Figure 4C and D) suggests that both pure water systems show important fluctuations in the electrostatic potential of nearly the same magnitude as those observed in the region between the ions in the ionic solution. This indicates that the perturbations observed in those regions are not an effect induced by the ions but correspond to variations in the electrostatic potential, which are intrinsic to the pure solution. According to this, the difference in the amplitude of the fluctuations observed between the atomistic and CG models (Figure 4A and B) are explained by the augmented granularity of the CG model. An estimation of such a difference is obtained from the approximate amplitudes observed in both atomistic (~0.002 V) and CG (~0.018 V) simulations. This indicates that the oscillations in the CG system have amplitudes nearly 1 order of magnitude higher than the ones in the atomistic system.

Bulk Electrolytic Properties. The vast majority of empirical parametrizations for single ions are typically developed to fit single ion properties, such as those examined in the previous sections. In order to complement the structural description of the CG aqueous solutions we studied some thermodynamic properties regarding ion−ion interactions: in particular, the Bjerrum and Debye lengths. The first represents the separation between two elementary charges at which the electrostatic interaction is comparable in magnitude to the thermal energy, and the second provides information regarding the distance at which the electrostatic potential of one ion is screened by the ionic strength of the surrounding medium. From the qualitative point of view, we retrieved the correct tendency in Bjerrum and Debye lengths upon changes in the ionic concentration (Table 4). Calculation of the Bjerrum and Debye lengths at 0.2 and 0.3 M gave values within a maximum error of 13% with respect to experimental values (Table 4). We obtained an underestimation of the Bjerrum length and, correspondingly, an overestimation of the Debye length, which is indicative of a slightly higher global electrostatic screening in the bulk solution, independent of the salt considered in the simulation (i.e., NaW^+ClW^- or KW^+ClW^-).

A direct measurement of the strength of the effective solvent-mediated interaction between ions is also very relevant, and it can be obtained from the osmotic pressure. For the case of NaW^+ClW^- at an ionic concentration of 0.5 M, we obtained a value of 35 bar (33 bar for KW^+ClW^-), which is essentially identical to that obtained by atomistic simulations using the CHARMM force field. Despite the large standard deviations, these values are in agreement with experimental reports (Table 4), suggesting a satisfactory balance in ion−ion and ion−WT4 interactions.

CG Solvation of Double-Stranded DNA. As a final example of application, we analyzed the explicit solvation of a dodecameric segment of double-stranded DNA. For this task, we used the already published CG scheme for simulating nucleic acids within the framework of the generalized Born model for implicit solvation.[24] In this contribution, the same system was simulated in the presence of explicit solvent and added salts. Both approaches furnish a similar picture of the structural and dynamical behavior of the double-helical segment of DNA with a maximum pairwise RMSD between both average structures of 0.25 nm. This is in agreement with the good reproduction of helical parameters obtained upon backmapping from the DNA simulation in explicit CG solvent (Figure S2, Supporting Information). Furthermore, the superposition of the covariance matrices calculated along the MD trajectories of the Drew−Dickerson dodecamer performed using implicit and explicit solvation gives an identity of 84%. This strongly suggests that both approaches sample nearly equivalent conformational spaces.

During the dynamics in the presence of explicit CG solvent, the global structure of the DNA dodecamer was fairly well conserved with an average RMSD of 0.25 nm from the starting (canonical) conformer. This can be inferred from the good superposition of snapshots taken at different times of the simulation (Figure 5A). Moreover, a good agreement is also obtained at the atomistic level upon backmapping. The all atoms RMSD of those shapshots compared with the X-ray structure 1BNA resulted in values of 0.35 nm (blue), 0.39 nm (green), and 0.34 nm (orange) (Figure S3, Supporting Information).

The WT4 molecules and cations closely interact with the CG nucleobases. It can be observed that the ordering of the WT4 molecules around the DNA qualitatively resembles the hydration features encountered in atomistic systems at both experimental and theoretical levels.[70−75] Conical arrangements of WT4 beads form around the phosphate groups (Figure 5B). The molecules of WT4 acquire an orientation guided by the electrostatic attraction between the positive (hydrogen-like) beads and the negatively charged phosphate superatoms. This results in the formation of structures alike to hydration cones (Figure 5B). In this kind of solvent arrangement around the backbone, WT4 molecules can be replaced by cations from the solution (Figure 5B) as observed experimentally.[76] Furthermore, ions can also remain transiently bound to the DNA visiting different positions within the minor groove. Extended hydration of the major groove and the formation of hydration spines in the minor groove are also observed, as illustrated in Figure 5C. A comprehensive picture of the hydration structure can be obtained from the WT4 beads' occupancy density map projected in the plane perpendicular to the DNA axis placed at the AT step (Figure 5D).

A more quantitative view of the solute/solvent interaction can be obtained from the cumulative RDF of the different species around the phosphate superatoms (Figure 5E). The

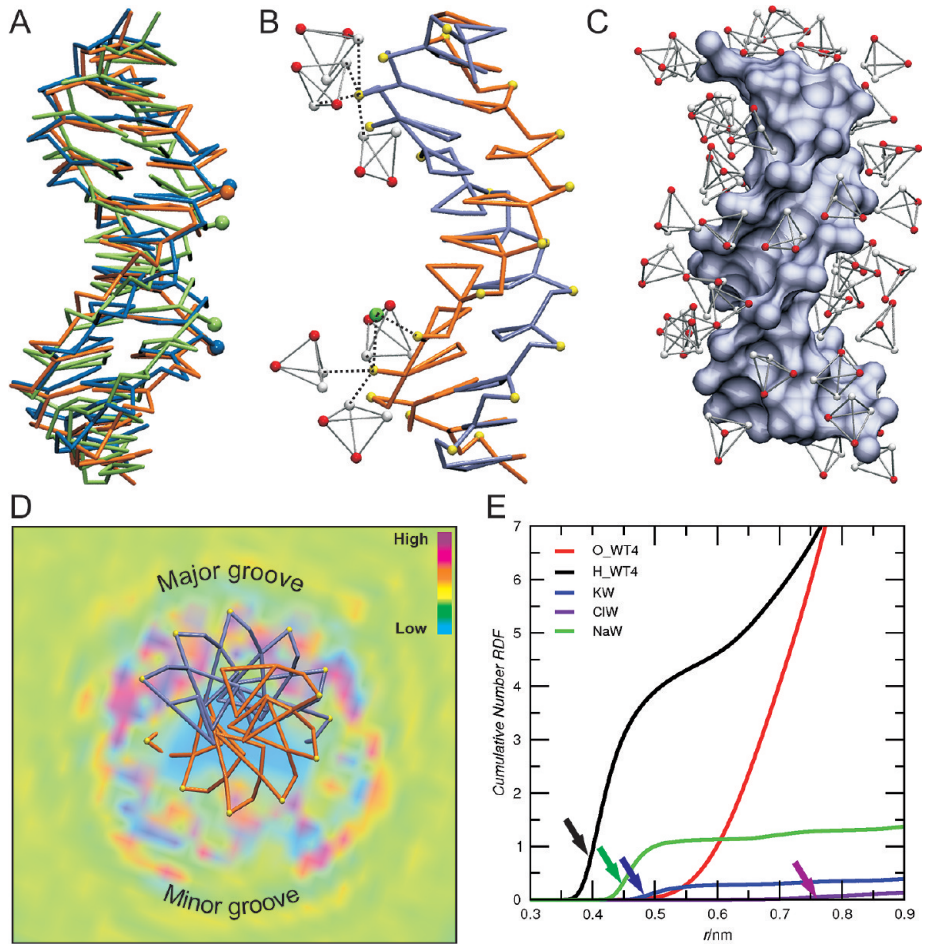

Figure 5. DNA and solvation structure. (A) Superposition of the DNA conformers taken from the first (blue), middle (green), and last (orange) frame of the simulation of system S_{15}^{CG}. Spheres indicate a pair of phosphate superatoms from opposite strands, which are highlighted in order to show the minor groove narrowing. An atomistic view of this superposition obtained from backmapped CG coordinates can be seen in Figure S6, Supporting Information. (B) WT4 and NaW specific interaction with the phosphate groups taken from a representative MD snapshot. The dashed lines highlight the conical arrangement of WT4 beads around phosphates (top) and the competition for the phosphate groups by WT4 and NaW (bottom). (C) WT4 solvation in the major and minor grooves from a random frame. The extensive hydration of the major groove and spines of hydration within the minor groove are evident. (D) WT4 occupancy density map projected onto a plane orthogonal to the DNA axis, located in the central AT step. The color scale represents the occupancy level, with a color range from cyan (low occupancy) to purple (high occupancy). Differences in major and minor groove are plain. Notice also the more punctuated location of WT4 within the minor groove indicative of solvation spines. (E) Cumulative number (integral of the RDFs) of negative and positive WT4 beads (red and black, respectively), NaW^+ (green), KW^+ (blue), and ClW^- (violet), with respect to the phosphate groups. Arrows indicate inflection points, which correspond to the first maxima of each RDF.

directionality in the WT4− phosphate interaction is evident from the right shift observed in the integral of the RDF corresponding to the oxygen-like beads' position with respect to that of the hydrogen-like beads (compare red and black lines in Figure 5E). The position of the first WT4 solvation shell forming conical structures lies at 0.4 nm from the phosphate superatom. This distance is in good agreement with the 0.38 nm found in atomistic simulations[24] and is comparable to the minimum distance of 0.32 observed in X-ray structures.[77]

Our model can also take into account the specificity in the DNA−cation interactions. As expected, sodium ions are more prone than potassium to interact with the solute. As mentioned above, sodium is frequently found in the close neighborhood of the phosphate moieties and even within the minor groove.[76,78,79] The closest sodium shell is localized at 0.45 nm from the phosphate, as compared with the 0.5 nm found for the bulkier potassium. In contrast, the radial distribution of the chlorine ions is much more shifted to the right with a first peak at 0.76 nm (Figure 5E).

The fraction of DNA charge neutralized within a cylinder of 0.9 nm from the exterior surface of the double-stranded helix is 0.75. This is in good agreement with the fraction of condensed counterions calculated within the condensation volume using Manning's counterion condensation theory for polyelectrolytes.[80] Moreover, this number is comparable with a fraction of 0.76 obtained by previous atomistic simulations using the same DNA sequence.[81] Among the fraction of condensed counterions, 76% corresponds to sodium and 24% to potassium; this is in qualitative agreement with a series of experimental and theoretical studies (see Savelyev and Papoian[82] and references therein).

While the global distribution of cations around the DNA contributes to the stability of the double helix, the specific

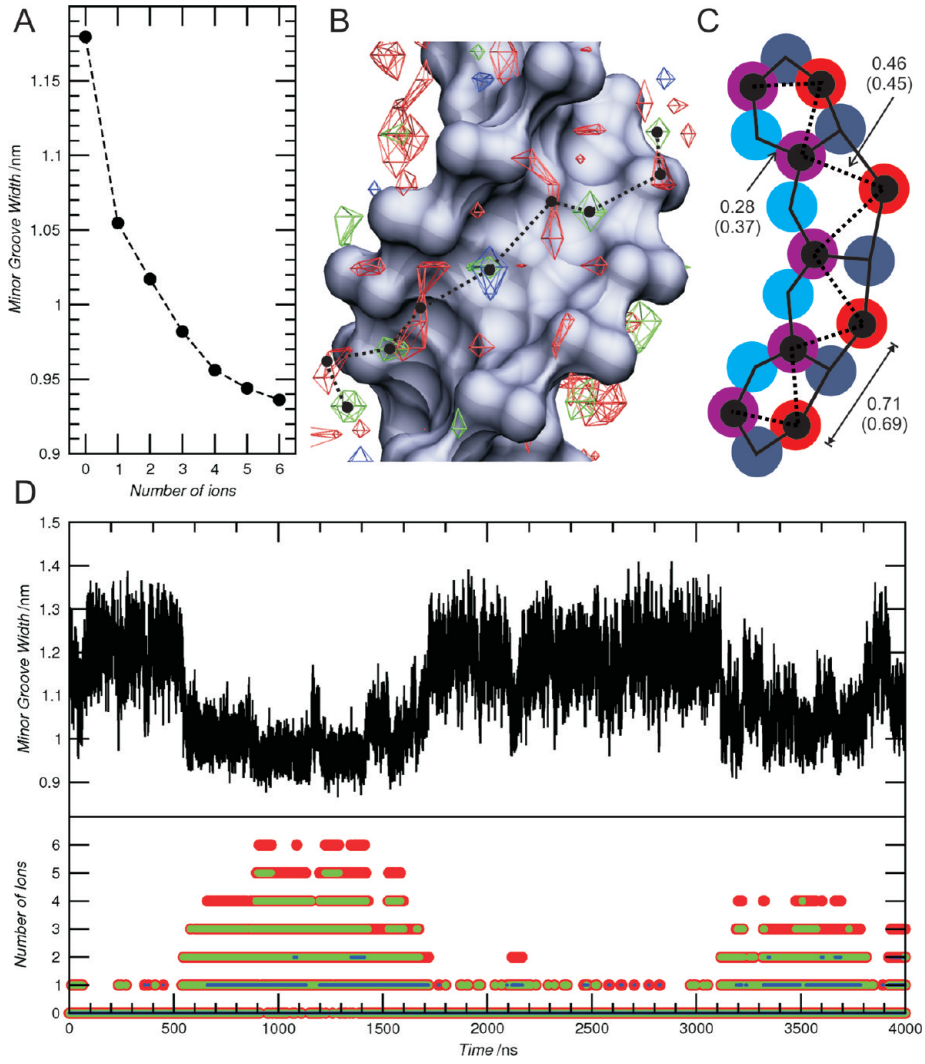

Figure 6. Binding of cations within the minor groove. (A) The minor groove width averaged over all frames with an equal number of bound cations is plotted against the number of bound ions (according to the criteria explained in the Methods section). (B) WT4 (red), NaW^+ (green), and KW^+ (blue) occupancy isosurfaces located in the minor groove of the AT track. Dashed path connecting black points indicate the zig-zag motif formed by the cations and WT4 beads in the minor groove. (C) Scheme showing the superimposition of the zig-zag motif (black circles connected by dashed line) observed in the CG simulation over the fused hexagon motif (continuous line) formed by the solvent sites (cyan, violet, gray, and orange circles) experimentally observed. These sites can by occupied by both water or cations.[77] The distances between corresponding solvation sites in the fused hexagon motif are shown and compared to the corresponding ones in the zig-zag motif (parentheses). (D) Minor groove width (top) and number of bound cations plotted against time (bottom). The number of cations is shown as the number of NaW^+ (green), number of KW^+ (blue), and total number of cations (sum of the number of NaW^+ and number of KW^+, in red).

interaction of cations with DNA has been related to local structural distortions. In particular, the binding of sodium ions within the minor groove has been proposed to mediate a narrowing in the minor groove.[83] In agreement with this proposal, we observed a clear correlation between the width of the minor groove and the binding of cations. Moreover, there seems to be a cumulative effect between these two events; i.e, a higher number of bound ions induces a more pronounced narrowing. This is clear from a measure of the average width of the minor groove with respect to the total number of bound ions (Figure 6A). The binding of one single ion is enough to induce a sensible change in the minor groove. Upon the successive incorporation of ions, the narrowing becomes more marked, reaching a minimum when six ions are concomitantly bound. Experimental studies on the same dodecamer also reveal a high occupancy of cations in the minor groove, leading to its narrowing.[78] Furthermore, a highly ordered structure is formed when cations and water interact with the AT track of the DNA. Such a structure is organized in four layers of solvent sites and resembles a series of fused hexagonal motifs.[78] Figure 6B shows the 3D occupancy map of WT4 and cations around the minor groove of the AT track. This map reveals sites highly occupied by WT4 (red wire mesh), NaW^+ (green wire mesh), and KW^+ (blue wire mesh) that resembles a zig-zag structure (dashed path connecting black points in Figure 6B). When such a zig-zag structure is superimposed onto the experimentally observed fused hexagon motif, good agreement is obtained for the second and fourth solvent-site layers, as confirmed by the inter solvent-site distances (Figure 6C).

The minor groove narrowing process appears to take place on two different time scales. The first is related to the binding

of one or two ions for up to a few dozen nanoseconds, while the second corresponds to the simultaneous binding of three to six ions for a period of nearly 1 μs (Figure 6D). This last induces a more marked and persistent but always reversible structural distortion with an average minor groove width of 0.98 nm (three bound cations) to 0.94 nm (six bound cations). The magnitude of this DNA distortion is in very good agreement with the average value of 0.96 nm obtained experimentally.[79]

It is worth noticing that temporal scales for sodium binding are coincident with the faster events found in this study have been reported for MD simulations at the atomistic level.[83−85] Unfortunately, the longest atomistic simulation reported in this system was carried out for 1.2 μs.[83] Although only nanosecond binding events were reported in that work, the agreement of the position of the binding sites and DNA distortion with X-ray data[78,79] may allow for speculation that a lack of longer binding events in the atomistic simulation could be related to insufficient sampling. Clearly, longer simulation times that go beyond the introductory scope of this paper would be needed to properly sample these long lasting events. This issue will be addressed in a forthcoming publication.

There is a marked selectivity for sodium against potassium. Indeed, while the simultaneous binding of more than two sodium ions is very frequent, only two potassium ions were present within the minor groove simultaneously, and this rather rare event was detected only five times in the 4 μs trajectory (Figure 6D and Figure S4, Supporting Information).

Finally, to complete the picture regarding the ionic structure around DNA, we analyzed the ionic distribution at longer distances from the double helix. This was done by calculating the number density of the three types of ions present in the system at increasing distances from DNA. In good agreement with the prediction from Poisson−Boltzmann theory,[86] the amount of electrolytes along a direction perpendicular to the DNA principal axis follows an exponential decay (Figure S5, Supporting Information).

Discussion and Conclusions

In this work, we have presented a model for simulating water at a coarse grain level. The WT4 model presented here is based on the transient tetrahedral structure adopted by water molecules in solution, preserving the molecular characteristics of the atomistic liquid. Due to the large number and heterogeneity of the CG models proposed in the literature, it is difficult to establish a fair comparison in terms of a computational speedup obtained with WT4. However, a comparison is more straightforward if we restrict it to the simplest models that condense three or four water molecules into a single bead.[53,54,87] This implies a coarse graining factor from 9 to 12, as compared to the value of ∼8 obtained for WT4. In addition to a similar coarse graining factor, our model offers some advantages, like the capacity to interact via explicit short- and long-range electrostatic interactions, and a dielectric permittivity. This grants the model the ability to reproduce some of the characteristic properties of water and electrolytic solutions.

The bead's masses were assigned to fit the water density at 300 K. Although this may raise some concerns about the suitability of the model at different temperatures, the relative error for the WT4 density with respect to the experimental determination of pure water remains below 3% in the range of 278 to 328 K (Figure 2B).

A strong assumption of the model is the fact that the existences of these five-member water clusters are supposed to be permanent, while their average lifetime in real water is on the order of picoseconds. This defect is partially compensated by setting a loose harmonic constraint between the beads of our representation. This allows for bond stretching variations of about 10% in the bond lengths, conferring a large plasticity to the WT4 molecules and the possibility of adapting their conformation according to its molecular environment.

The use of the WT4 model to solvate simple ions reproduces their hydration structure and some thermodynamic properties such as osmotic pressure, which is often considered a quality gauge of the parametrization.

We notice that important properties such as the isothermal compressibility and surface tension are poorly described by WT4. This may be of particular relevance in the study of self-assembly phenomena, and for such treatment special caution is advised. Despite this deviation from ideal behavior, the description of the double-stranded DNA segment does not seem to be compromised. This suggests that the long−medium range screening properties of the solvation model are suitable for overcoming the strong electrostatic repulsion generated by the negatively charged phosphate groups of DNA. In fact, the addition of explicit solvation and different ionic species highly enhanced the description of the DNA dynamics, allowing, for instance, the reproduction of the cation-mediated narrowing of the minor groove that could not be studied within the implicit solvation approach. While the implicit solvation approach can provide a good and faster description of sequence dependent effects on the structural and dynamical stability, inclusion of the explicit solvent can allow for the study, for instance, of the influence of intrinsic versus extrinsic sources of DNA flexibility, solvent mediated effects, ionic specificity, etc. Furthermore, the use of periodic boundary conditions and explicit electrostatics permits a more realistic consideration of long-range effects.

WT4 together with the CG electrolyte model represent correctly the gross solvation structure around DNA, as noted by the percentage of DNA charge neutralized at 0.9 nm that closely resembles that of atomistic simulations and that predicted by counterion condensation theory. Moreover, DNA hydration features like the extensive major groove hydration, minor groove hydration spines, and conical arrangement around phosphate groups that resembles the hydration cones observed in atomistic simulations and experimental data are well reproduced. It is important to note in this context that the development of interaction parameters has always been carried out within the philosophy of fitting structural properties of water, ionic solvation, and DNA. In this respect, we first developed the representation for WT4 in the bulk and then added the description of simple CG

electrolytes. Finally, the existing DNA parameters for implicit solvation were slightly modified to further refine its structural description when embedded in explicit solvent. In this sense, good agreement with experimental determinations can be considered emerging properties of the model because no specific fittings of cross interaction potentials have been performed.

The simulation scheme presented here allows for runing at a rate of ~1 μs per day (S_{15}^{CG}) on a dual quad core PC (Intel Xeon 2.66 GHz). This performance along with the nearly atomic resolution achievable upon backmapping of the coordinates in our DNA model[24] make the millisecond time scale reachable. This would effectively bridge the gap between the time scales feasible to MD and those that are biologically relevant.

Finally, we would like to stress the point that the model presented here computes all of the interactions using a typical Hamiltonian function, avoiding *ad hoc* code modifications/recompilations. Topologies, interaction parameters, and coordinate files for GROMACS implementation are available from the authors upon request.

Acknowledgment. This work was supported by ANII—Agencia Nacional de Investigación e Innovación, Programa de Apoyo Sectorial a la Estrategia Nacional de Innovación—INNOVA URUGUAY (Agreement n8 DCI - ALA/2007/19.040 between Uruguay and the European Commission) and Grant FCE_60-2007. L.D. and M.R.M. are beneficiaries of the National Fellowship System of ANII.

Supporting Information Available: Mapping scheme between atomistic and CG model for DNA. Helical parameters of the backmapped trajectory. Table containing interaction parameters sets for DNA. Energy fluctuations analysis. Cations in the minor groove frequency. Supperposition of backmapped DNA structures. Details on osmotic pressure calculation. Long range ionic structure. This information is available free of charge via the Internet at http://pubs.acs.org/.

References

(1) Karplus, M.; McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646–652.

(2) Klein, M. L.; Shinoda, W. Large-scale molecular dynamics simulations of self-assembling systems. *Science* **2008**, *321*, 798–800.

(3) Murtola, T.; Bunker, A.; Vattulainen, I.; Deserno, M.; Karttunen, M. Multiscale modeling of emergent materials: biological and soft matter. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1869–1892.

(4) Levitt, M.; Warshel, A. Computer simulation of protein folding. *Nature* **1975**, *253*, 694–698.

(5) Tanaka, S.; Scheraga, H. A. Statistical Mechanical Treatment of Protein Conformation. I. Conformational Properties of Amino Acids in Proteins. *Macromolecules* **1976**, *9*, 142–159.

(6) Yin, Y.; Arkhipov, A.; Schulten, K. Simulations of membrane tubulation by lattices of amphiphysin N-BAR domains. *Structure* **2009**, *17*, 882–892.

(7) Arkhipov, A.; Yin, Y.; Schulten, K. Four-scale description of membrane sculpting by BAR domains. *Biophys. J.* **2008**, *95*, 2806–2821.

(8) Wee, C. L.; Gavaghan, D.; Sansom, M. S. P. Interactions between a voltage sensor and a toxin via multiscale simulations. *Biophys. J.* **2010**, *98*, 1558–1565.

(9) Ayton, G. S.; Voth, G. A. Hybrid coarse-graining approach for lipid bilayers at large length and time scales. *J. Phys. Chem. B* **2009**, *113*, 4413–4424.

(10) Treptow, W.; Marrink, S.; Tarek, M. Gating motions in voltage-gated potassium channels revealed by coarse-grained molecular dynamics simulations. *J. Phys. Chem. B* **2008**, *112*, 3277–3282.

(11) Yefimov, S.; van der Giessen, E.; Onck, P. R.; Marrink, S. J. Mechanosensitive membrane channels in action. *Biophys. J.* **2008**, *94*, 2994–3002.

(12) Periole, X.; Huber, T.; Marrink, S.; Sakmar, T. P. G protein-coupled receptors self-assemble in dynamics simulations of model bilayers. *J. Am. Chem. Soc.* **2007**, *129*, 10126–10132.

(13) Durrieu, M.; Bond, P. J.; Sansom, M. S. P.; Lavery, R.; Baaden, M. Coarse-grain simulations of the r-snare fusion protein in its membrane environment detect long-lived conformational sub-states. *Chem. Phys. Chem.* **2009**, *10*, 1548–1552.

(14) Srinivas, G.; Discher, D.; Klein, M. Self-assembly and properties of diblock copolymers by coarse-grain molecular dynamics. *Nature* **2004**, *3*, 638–644.

(15) Nielsen, S.; Lopez, C.; Srinivas, G.; Klein, M. Coarse grain models and the computer simulation of soft materials. *J. Phys.: Condens. Matter* **2004**, *16*, R481–R512.

(16) Arkhipov, A.; Freddolino, P. L.; Schulten, K. Stability and dynamics of virus capsids described by coarse-grained modeling. *Structure* **2006**, *14*, 1767–1777.

(17) Srinivas, G.; Klein, M. Molecular dynamics simulations of self-assembly and nanotube formation by amphiphilic molecules in aqueous solution: a coarse-grain approach. *Nanotechnology* **2007**, 18.

(18) Zhou, J.; Thorpe, I. F.; Izvekov, S.; Voth, G. A. Coarse-grained peptide modeling using a systematic multiscale approach. *Biophys. J.* **2007**, *92*, 4289–4303.

(19) Voth, G. A. *Coarse-graining of condensed phase and biomolecular systems*, 1st ed.; Taylor & Francis Group: New York, 2009; pp 1−455.

(20) DeMille, R. C.; Molinero, V. Coarse-grained ions without charges: reproducing the solvation structure of NaCl in water using short-ranged potentials. *J. Chem. Phys.* **2009**, *131*, 034107.

(21) Molinero, V.; Moore, E. B. Water modeled as an intermediate element between carbon and silicon. *J. Phys. Chem. B* **2009**, *113*, 4008–4016.

(22) Savelyev, A.; Papoian, G. A. Molecular renormalization group coarse-graining of electrolyte solutions: application to aqueous NaCl and KCl. *J. Phys. Chem. B* **2009**, *113*, 7785–7793.

(23) Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J. Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput. Biol.* **2010**, *6*, e1000810.

(24) Dans, P.; Zeida, A.; Machado, M.; Pantano, S. A coarse grained model for atomic-detailed DNA simulations with explicit electrostatics. *J. Chem. Theory Comput.* **2010**, *6*, 1711–1725.

(25) Head-Gordon, T.; Hura, G. Water structure from scattering experiments and simulation. *Chem. Rev.* **2002**, *102*, 2651–2670.

(26) Narten, A. H.; Danford, M. D.; Levy, H. A. X-ray diffraction study of liquid water in the temperature range 4−200°C. *Discuss. Faraday Soc.* **1967**, *43*, 97–107.

(27) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. *Intermolecular forces*; Pullman, B., Ed.; D. Reidel Publishing Company: Dordrecht, The Netherlands, 1981; pp 331−342.

(28) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(29) Mancinelli, R.; Botti, A.; Bruni, F.; Ricci, M. A.; Soper, A. K. Hydration of sodium, potassium, and chloride ions in solution and the concept of structure maker/breaker. *J. Phys. Chem. B* **2007**, *111*, 13570–13577.

(30) Mancinelli, R.; Botti, A.; Bruni, F.; Ricci, M. A.; Soper, A. K. Perturbation of water structure due to monovalent ions in solution. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2959–2967.

(31) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. Gromacs: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43–56.

(32) Lindahl, E.; Hess, B.; van der Spoel, D. Gromacs 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* **2001**, *7*, 306–317.

(33) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. Gromacs: fast, flexible and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

(34) Bekker, H.; Berendsen, H. J. C.; Dijkstra, E. J.; Achterop, S.; van Drunen, R.; van der Spoel, D.; Sijbers, A.; Keegstra, H.; Reitsma, B.; Renardus, M. K. R. *Gromacs: A parallel computer for molecular dynamics simulations*; de Groot, R. A., Nadrchal, J., Eds.; World Scientific: Singapore, 1993.

(35) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255–268.

(36) Hoover, W. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695–1697.

(37) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

(38) Nosé, S.; Klein, M. L. Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* **1983**, *50*, 1055–1076.

(39) Darden, T.; York, D.; Pedersen, L. Particle mesh ewald: an n-log(n) method for ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(40) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. A smooth particle mesh ewald potential. *J. Chem. Phys.* **1995**, *103*, 8577–8592.

(41) Winger, M.; Trzesniak, D.; Baron, R.; van Gunsteren, W. F. On using a too large integration time step in molecular dynamics simulations of coarse-grained molecular models. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1934–1941.

(42) He, X.; Shinoda, W.; DeVane, R.; Klein, M. L. Exploring the utility of coarse-grained water models for computational studies of interfacial systems. *Mol. Phys.* **2010**, *108*, 2007–2020.

(43) Herrero, C. P. Compressibility of solid helium. *J. Phys.: Condens. Matter* **2008**, *20*, 295230.

(44) van Buuren, A. R.; Marrink, S. J.; Berendsen, H. J. C. A molecular dynamics study of the decane/water interface. *J. Phys. Chem.* **1993**, *97*, 9206–9212.

(45) Mark, A. E.; van Helden, S. P.; Smith, P. E.; Janssen, L. H. M.; van Gunsteren, W. F. Convergence properties of free energy calculations: alpha-cyclodextrin complexes as a case study. *J. Am. Chem. Soc.* **1994**, *116*, 6293–6302.

(46) Beglov, D.; Roux, B. Finite representation of an infinite bulk system: solvent boundary potential for computer simulations. *J. Chem. Phys.* **1994**, *100*, 9050–9063.

(47) Luo, Y.; Roux, B. Simulation of osmotic pressure in concentrated aqueous salt solutions. *J. Phys. Chem. Lett.* **2010**, *1*, 183–189.

(48) Herrera, E. F.; Pantano, S. Salt induced asymmetry in membrane simulations by partial restriction of ionic motion. *J. Chem. Phys.* **2009**, *130*, 195105–195114.

(49) Lavery, R.; Sklenar, H. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63–91.

(50) Soper, A. K. The radial distribution functions of water and ice from 673 K and at pressures up to 400 MPa. *Chem. Phys.* **2000**, *258*, 121–137.

(51) Kell, G. S. Density, thermal expansivity, and compressibility of liquid water from 0° to 150°C: correlations and tables for atmospheric pressure and saturation reviewed and expressed on 1968 temperature scale. *J. Chem. Eng. Data* **1975**, *20*, 97–105.

(52) Holz, M.; Heil, S. R.; Sacco, A. Temperature-dependent self-diffusion coefficients of water and six selected molecular liquids for calibration in accurate H-1 NMR PFG measurements. *Phys. Chem. Chem. Phys.* **2000**, *2*, 4740–4742.

(53) Marrink, S. J.; de Vries, A. H.; Mark, A. E. Coarse grained model for semiquantitative lipid simulations. *J. Phys. Chem. B* **2004**, *108*, 750–760.

(54) Groot, R. D.; Rabone, K. L. Mesoscopic simulation of cell membrane damage, morphology change and rupture by nonionic surfactants. *Biophys. J.* **2001**, *81*, 725–736.

(55) The value of the permittivity can vary with the condition of the simulation, size of the computational box, etc.

(56) Kusalik, P. G.; Svishchev, I. M. The spatial structure in liquid water. *Science* **1994**, *265*, 1219–1221.

(57) van Maaren, P. J.; van der Spoel, D. Molecular dynamics of water with novel shell-model potentials. *J. Phys. Chem. B* **2001**, *105*, 2618–2626.

(58) van der Spoel, D.; van Maaren, P. J.; Berendsen, H. J. C. A systematic study of water models for molecular simulation: derivation of water models optimized for use with a reaction field. *J. Chem. Phys.* **1998**, *108*, 10220–10230.

(59) Mahoney, M. W.; Jorgensen, W. L. Diffusion constant of the TIP5P model of liquid water. *J. Chem. Phys.* **2001**, *114*, 363–366.

(60) Yu, H.; van Gunsteren, W. F. Charge-on-spring polarizable water models revisited: from water clusters to liquid water to ice. *J. Chem. Phys.* **2004**, *121*, 9549–9564.

(61) Yu, H.; Hansson, T.; van Gunsteren, W. F. Development of a simple self-consistent polarizable model for liquid water. *J. Chem. Phys.* **2003**, *118*, 221–234.

(62) Chen, F.; Smith, P. E. Simulated surface tensions of common water models. *J. Chem. Phys.* **2007**, *126*, 221101–221104.

(63) Wang, H.; Junghans, C.; Kremer, K. Comparative atomistic and coarse-grain study of water: what do we lose by coarse-graining. *Eur. Phys. J. E* **2009**, *28*, 221–229.

(64) Mahoney, M. W.; Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **2000**, *112*, 8910–8922.

(65) Murrell, J. N.; Jenkins, A. D. *Properties of liquids and solutions*, 2nd ed.; John Wiley & Sons: Chichester, U. K., 1994; pp 1−299.

(66) Eisenberg, D.; Kauzmann, W. *The Structure and Properties of Water*; Oxford University Press: Oxford, U.K., 1969; pp 1−308.

(67) Dilmohamud, B. A.; Seeneevaseen, J.; Rughooputh, S. D. D. V.; Ramasami, P. Surface tension and related thermodynamic parameters of alchohols using the Traube stalagmometer. *Eur. J. Phys.* **2005**, *26*, 1079.

(68) Rodnikova, M. N. A new approach to the mechanism of solvophobic interactions. *J. Mol. Liq.* **2007**, *136*, 211–213.

(69) Kalcher, I.; Horinek, D.; Netz, R. R.; Dzubiella, J. Ion specific correlations in bulk and at biointerfaces. *J. Phys.: Condens. Matter* **2009**, *21*, 424108.

(70) Shotton, M. W.; Pope, L. H.; Forsyth, V. T.; Langan, P.; Grimm, H.; Rupprecht, A.; Denny, R. C.; Fuller, W. A high-angle neutron fiber diffraction study of the hydration of B-DNA. *Physica B* **1998**, *243*, 1166–1168.

(71) Young, M. A.; Ravishanker, G.; Beveridge, D. L. A 5-ns molecular dynamics trajectory for B-DNA: analysis of structure, motions, and solvation. *Biophys. J.* **1997**, *73*, 2313–2336.

(72) Cheatham, T. E., 3rd.; Srinivasan, J.; Case, D. A.; Kollman, P. A. Molecular dynamics and continuum solvent studies of the stability of polyG-polyC and polyA-polyT DNA duplexes in solution. *J. Biomol. Struct. Dyn.* **1998**, *16*, 265–280.

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

(74) Feig, M.; Pettitt, B. M. Modeling high-resolution hydration patterns in correlation with DNA sequence and conformation. *J. Mol. Biol.* **1999**, *286*, 1075–1095.

(75) Young, M. A.; Beveridge, D. L. Molecular dynamics simulations of an oligonucleotide duplex with adenine tracts phased by a full helix turn. *J. Mol. Biol.* **1998**, *281*, 675–687.

(76) Kochoyan, M.; Leroy, J. L. Hydration and solution structure of nucleic acids. *Curr. Opin. Struct. Biol.* **1995**, *5*, 329–333.

(77) Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2179–2183.

(78) Shui, X.; Sines, C. C.; McFail-Isom, L.; VanDerveer, D.; Williams, L. D. Structure of the potassium form of CGC-GAATTCGCG: DNA deformation by electrostatic collapse around inorganic cations. *Biochemistry* **1998**, *37*, 16877–16887.

(79) Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. The B-DNA dodecamer at high resolution reveals a spine of water on sodium. *Biochemistry* **1998**, *37*, 8341–8355.

(80) Manning, G. S. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* **1978**, *11*, 179–246.

(81) Ponomarev, S. Y.; Thayer, K. M.; Beveridge, D. L. Ion motions in molecular dynamics simulations on DNA. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14771–14775.

(82) Savelyev, A.; Papoian, G. A. Electrostatic, steric, and hydration interactions favor Na+ condensation around DNA compared with K+. *J. Am. Chem. Soc.* **2006**, *128*, 14506–14518.

(83) Pérez, A.; Luque, F. J.; Orozco, M. Dynamics of B-DNA on the microsecond time scale. *J. Am. Chem. Soc.* **2007**, *129*, 14739–14745.

(84) McConnell, K. J.; Beveridge, D. L. Molecular dynamics simulations of B-DNA: sequence effects on A-tract-induced bending and flexibility. *J. Mol. Biol.* **2001**, *314*, 23–40.

(85) Feig, M.; Pettitt, B. M. Sodium and chlorine ions as part of the DNA solvation shell. *Biophys. J.* **1999**, *77*, 1769–1781.

(86) Fuoss, R. M.; Katchalsky, A.; Lifson, S. The potential of an infinite rod-like molecule and the distribution of the counter ions. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 579–589.

(87) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. A coarse grain model for phospholipid simulations. *J. Phys. Chem. B* **2001**, *105*, 4464–4470.

CT100379F