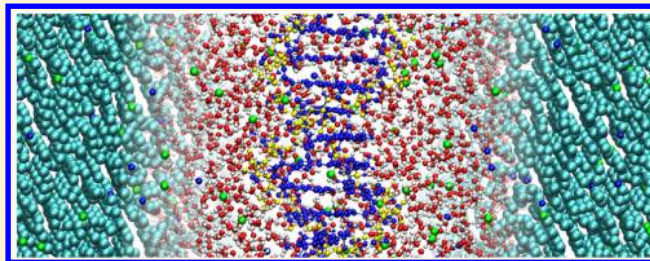


# Adaptive Resolution Simulation of a DNA Molecule in Salt Solution

Julija Zavadlav,<sup>†</sup> Rudolf Podgornik,<sup>\*,‡,§</sup> and Matej Praprotnik<sup>\*,†</sup><sup>†</sup>Laboratory for Molecular Modeling, National Institute of Chemistry, Hajdrihova 19, SI-1001 Ljubljana, Slovenia<sup>‡</sup>Department of Physics, Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia<sup>§</sup>Theoretical Physics Department, J. Stefan Institute, Jamova c. 39, SI-1000 Ljubljana, Slovenia

**ABSTRACT:** We present a multiscale simulation of a DNA molecule in 1 M NaCl salt solution environment, employing the *adaptive resolution simulation approach* that allows the solvent molecules, i.e., water and ions, to change their resolution from atomistic to coarse-grained and vice versa adaptively on-the-fly. The region of high resolution moves together with the DNA center-of-mass so that the DNA itself is always modeled at high resolution. We show that our multiscale simulations yield a stable DNA–solution system, with statistical properties similar to those produced by the conventional all-atom molecular dynamics simulation. Special attention is given to the collective properties, such as the dielectric constant, as they provide a sensitive quality measure of our multiscale approach.



## 1. INTRODUCTION

The solution environment, i.e., the surrounding ions and water, profoundly affects the structure and stability of a DNA molecule, which in turn affects the local solvent causing its properties to deviate from bulk values.<sup>1</sup> Owing to DNA high charges and polar structure, the electrostatic interactions play a fundamental role in this delicate DNA–solvent interplay.<sup>2</sup> To properly model and understand the electrostatic interactions among various parts in such complex systems and meaningfully compare computational results to existing experiments,<sup>3</sup> it is essential to determine the collective spatially dependent dielectric response of both DNA and the surrounding solvent.<sup>4–6</sup>

The dielectric constants of biological macromolecules are very difficult to determine experimentally. For this reason the DNA's dielectric constant was measured successfully only recently.<sup>3</sup> Using the electrostatic force microscopy for detection of ultraweak polarization forces of DNA encapsulated inside the T7 bacteriophage capsids, the dielectric constant was found to be  $\sim 8$ . The obtained value is significantly higher compared to the proteins with a dielectric constant of  $\sim 4$ .<sup>7</sup> Two all-atom molecular dynamics (MD) simulations have confirmed these findings with resulting dielectric constants of  $\sim 8$  and  $\sim 16$  for the Drew–Dickerson dodecamer<sup>3</sup> and DNA triple helix,<sup>8</sup> respectively. In both cases, the dielectric constant was reported to be lower for the sugar and base components and substantially higher for the phosphate backbone.

On the other hand, a DNA molecule exhibits a strong influence on the surrounding aqueous solvent.<sup>1</sup> Water is found to be considerably ordered and has a slower dynamics, due to the motional restrictions imposed by the DNA and the ions.<sup>9</sup> The resulting dielectric constant is consequently much lower compared to its bulk value. Furthermore, since both DNA grooves are binding sites for ligands, it is intriguing that a large

discrepancy between them is observed in the local dielectric behavior ( $\sim 55$  and  $\sim 20$  for the major<sup>10</sup> and minor grooves,<sup>11</sup> respectively). Understanding of the solution environment is further complicated by the presence of mobile salt ions and counterions, composing the ionic atmosphere, which respond to DNA charges and local dielectric solution properties, adding additional features to the already complicated solution environment of DNA.<sup>12</sup>

The natural DNA environment is a salt solution (the counterions that neutralize the system alone are insufficient<sup>12</sup>) which for statistical reasons requires a significant number of ions and consequently a large simulation system. Experimentally the salt concentration can be in excess of 1 M.<sup>2</sup> All-atom simulations of such systems are thus computationally very challenging. One way to overcome this limitation is to develop coarse-grained (CG) models, where irrelevant degrees of freedom are integrated out, as, for example, in refs 13–16. However, to adequately simulate the previously described spatially varying microenvironments, one has to resort to the atomic level of detail. The dielectric profile displays bulk behavior beyond the second coordination shell from the DNA molecule, and it would be therefore computationally advantageous to use in this region a simplified CG representation for the solvent whereas the atomistic (AT) resolution would be sufficient at the remainder of the system. Approaches that combine various models within a single simulation have attracted pronounced interest in recent years.<sup>17–26</sup> Among the most popular approaches is the adaptive resolution scheme (AdResS)<sup>27–33</sup> that enables one to couple different particle descriptions in a concurrent manner and has been successfully applied to a number of different biological systems.<sup>34–38</sup>

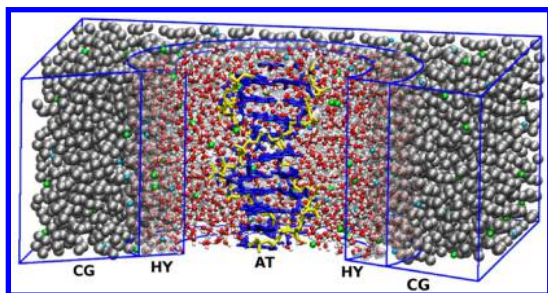
Received: June 24, 2015

Published: September 10, 2015

Here, we present an AdResS simulation of a double-stranded DNA molecule in a 1 M NaCl bathing salt solution. The DNA is at all times modeled at the full AT resolution, while water molecules and ions dynamically change their resolution between the AT and CG representations, according to their distance from the DNA molecule. Our analysis is focused on the collective dielectric properties of the DNA and the nearby solvent that could be affected in the multiscale simulations due to the anomalously long-range nature of the electrostatic interaction. Moreover, the AdResS method is able to provide additional physical insight, for instance, into the extent of the hydrogen bond network in the solvation shells of solute molecules.<sup>35,39</sup> In the present case the atomistic water is coupled to a charge neutral CG bead that is nonpolar. This allows one to assess the distance from the DNA molecule at which the inclusion of explicit atomistic water is required.

## 2. METHODS AND SIMULATION DETAILS

A schematic representation of the simulated system is depicted in Figure 1. The atomic coordinates of a B-form DNA molecule with 10 base-pair sequence (5'-CTCTCGCGCG-3') were generated using the 3D-DART Web server.<sup>40</sup>



**Figure 1.** Schematic cross-section of simulation box with cylindrical resolution regions: atomistic (AT), hybrid (HY), and coarse-grained (CG). Two levels of resolution are used for solvent molecules. A high level of resolution is used for solvent molecules within a certain radius from the DNA's CoM. Further away the water molecules are represented as single beads (gray). The Na<sup>+</sup> and Cl<sup>-</sup> atoms are shown in green and blue, respectively. The high resolution region cylinder moves with the DNA's CoM, which ensures atomistic modeling of the DNA molecule and the surrounding layer of water at all times.

Generally, the DNA molecule is simulated as a single oligonucleotide. However, in this work we employ periodic boundary condition along the DNA helix where patches, i.e., corresponding intramolecular DNA interactions defined by bond, angle, and dihedral interaction potentials, are used to connect each strand to its periodic image along the *z*-axis. This simulation setup enables us to mimic an infinitely long DNA molecule and is better suited for comparison with experimental studies where the DNA molecules are commonly ordered and densely packed.<sup>41</sup> The imposed DNA periodicity fixes the helical twist of the DNA molecule (the simulation box size is chosen such that it corresponds exactly to one B-DNA pitch, i.e., 3.4 nm) and prevents any major bending. Consequently, the DNA's persistence length, which is about 50 nm and much longer than one DNA pitch, and associated bending fluctuations do not enter our level of description. Even so, previous simulations with a periodic DNA fragment have shown stable, B-form DNA structures that still have a lot of freedom for variation of the local DNA structure.<sup>41,42</sup> The DNA

molecule is immersed in a 1 M NaCl bathing salt solution. Additional 20 Na<sup>+</sup> counterions are added to neutralize the negative charge of the DNA molecule. While the salt concentration used is higher than the physiological concentration (approximately 0.15 M) because an adequate number of ions are needed for acceptable statistics, it is still within the experimentally accessible regime.<sup>2</sup>

The DNA molecule is always modeled at the full AT resolution. The solvent's level of representation depends on the distance from the DNA's center-of-mass (CoM). At short distances we resort to the AT resolution, while at larger distances we employ a CG salt solution model<sup>34</sup> in which water molecules are represented as one-bead particles without explicit electric charges (see section 3). The multiscale simulations are carried out using the AdResS method. The total force acting on a molecule  $\alpha$  is then

$$\mathbf{F}_\alpha = \sum_{\beta \neq \alpha} w(|\mathbf{R}_\alpha - \mathbf{R}|) w(|\mathbf{R}_\beta - \mathbf{R}|) \mathbf{F}_{\alpha\beta}^{\text{at}} + \sum_{\beta \neq \alpha} [1 - w(|\mathbf{R}_\alpha - \mathbf{R}|) w(|\mathbf{R}_\beta - \mathbf{R}|)] \mathbf{F}_{\alpha\beta}^{\text{cg}} - \mathbf{F}_\alpha^{\text{TD}}(|\mathbf{R}_\alpha - \mathbf{R}|) \quad (1)$$

where  $\mathbf{F}_{\alpha\beta}^{\text{at}}$  and  $\mathbf{F}_{\alpha\beta}^{\text{cg}}$  are the forces between molecules  $\alpha$  and  $\beta$ , obtained from the AT and CG potentials, respectively. The sigmoidal function  $w \in [0,1]$  is used to smoothly couple the high and low resolution regimes, where  $\mathbf{R}_\alpha$ ,  $\mathbf{R}_\beta$ , and  $\mathbf{R}$  are two-dimensional (*x,y*) vectors of CoMs of molecules  $\alpha$  and  $\beta$ , and the DNA, respectively. The thermodynamic (TD) force  $\mathbf{F}_\alpha^{\text{TD}}$  acts on molecules' CoM in the hybrid (HY) region and ensures the thermodynamic equilibrium of the system.<sup>29,43</sup> It depends on the molecule type, i.e., we use three different ones that correspond to water molecules and Na and Cl ions,<sup>34</sup> but, as our simulation result shows, it does not depend on the AT region size (see section 3).

For a detailed derivation of  $\mathbf{F}_\alpha^{\text{TD}}$  we refer the reader to refs 29, 43, and 44. Here, for completeness we briefly summarize its derivation. To enforce a uniform density profile, we have to compensate for the chemical potential differences between the high and low resolution molecular models. Therefore,  $\mathbf{F}_\alpha^{\text{TD}}$  is defined as a negative gradient of the effective excess chemical potential due to the intermolecular interactions.<sup>44</sup> Numerically, this translates into an iterative formula<sup>29,43</sup>

$$\mathbf{F}_\alpha^{\text{TD}^{i+1}} = \mathbf{F}_\alpha^{\text{TD}^i} - C \nabla \rho^i \quad (2)$$

where  $C$  is an appropriately chosen numerical prefactor.

The geometrical boundary between resolution regions is a cylinder as it adequately reflects the shape of the DNA molecule. The center of the all-atom cylinder coincides with the DNA's CoM at all times; i.e., the resolution domains follow DNA's random translation. Such setup ensures that the DNA is always expressed in full all-atom detail and surrounded by a layer of the all-atom solvent. We investigate four sizes of atomistic cylinder radii: 1.5, 1.8, 2.1, and 2.4 nm. In all cases the width of the HY region is 0.9 nm. As a reference, we use simulation where the AT region extends across the whole simulation box, i.e., a full-blown atomistic simulation.

All MD simulations are performed using the ESPResSo++ software package.<sup>45</sup> For the integration we use the standard velocity Verlet with a time step of 1 fs. We use an orthorhombic simulation box with periodic boundary conditions and minimum image convention. The simulation box size is 8.5 nm  $\times$  8.5 nm  $\times$  3.4 nm. Thus, in the *z*-direction this

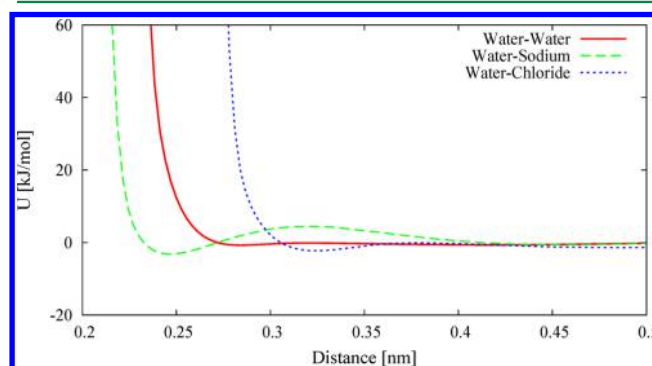
corresponds to exactly one DNA pitch. The temperature is maintained at 300 K with a local Langevin thermostat, with the value of the friction constant equal to 5.0/ps. The hydrogen atoms of the DNA molecule are constrained with the RATTLE<sup>46</sup> algorithm, while the geometry of water molecules is constrained with SETTLE.<sup>47</sup> We employ the standard TIP3P<sup>48</sup> water model for AT water in combination with the AMBER<sup>49</sup> force field for ions; i.e., we use Lennard-Jones and electrostatic interactions with standard AMBER force field parameters between atoms of water molecules and ions. The DNA molecule is modeled using the AMBER<sup>49</sup> force field. For the interactions of the DNA molecule with the solvent we also employ the same force field, which performs well in DNA simulations.<sup>50</sup> The nonbonded interactions are calculated with Lennard-Jones and Coulomb interactions within a cutoff distance  $r_c = 0.9$  nm. The generalized reaction field method<sup>51</sup> is used for the electrostatic interaction beyond the cutoff, with dielectric permittivity of the outer region equal to 80 and the Debye screening length  $\kappa = 3.25/\text{nm}$  corresponding to a 1 M salt solution. As already stated, one reason for conducting our simulations at this concentration, which is higher than the physiological concentration (approximately 0.15 M), is that an adequate number of ions is required for acceptable statistics. However, many experiments and applications in DNA nanotechnology are commonly performed at 1 M concentration. Note also that lower concentrations, e.g., 0.15 M, would not require a different method, just larger simulation systems.

The dielectric permittivity of the inner region, that is, within the cutoff distance, is equal to 1 and 80 for the AT and CG regions, respectively. This ensures that the ion–ion interactions are properly screened in the CG region, where we use the same electric charges for the salt ions as in the AT region; i.e., the ions thus interact via the same potentials in the AT and CG models aside from the changed dielectric permittivity. The specific method employed to perform computation of electrostatic interaction is an important issue in DNA simulations, and we refer the reader to ref 52 where performances of different methods, e.g., the Ewald summation, are compared. Here, we have opted for the generalized reaction field method for computing electrostatic interactions because, first, it is very convenient for implementation in AdResS; i.e., it is pairwise and short-ranged,<sup>28,53</sup> and second, we use a slightly modified version of a recently introduced multiscale model of salt solution based on the generalized reaction field approach.<sup>34</sup> For validating and comparing statistical properties of our modified multiscale salt solution model (in the absence of the DNA molecule) with the one from ref 34, we have used the trajectories of the same length as in ref 34, i.e., 10 ns. All simulations of the DNA molecule in solution are run for 15 ns after 1 ns of equilibration. These are rather short simulation runs in comparison to the microsecond time scales currently accessible.<sup>54</sup> However, we stress that here we do not simulate rare-event processes, e.g., breaking or formation of base pairs. Instead, the DNA molecule is expected to remain in one stable equilibrium configuration throughout the simulation runs.

### 3. RESULTS AND DISCUSSION

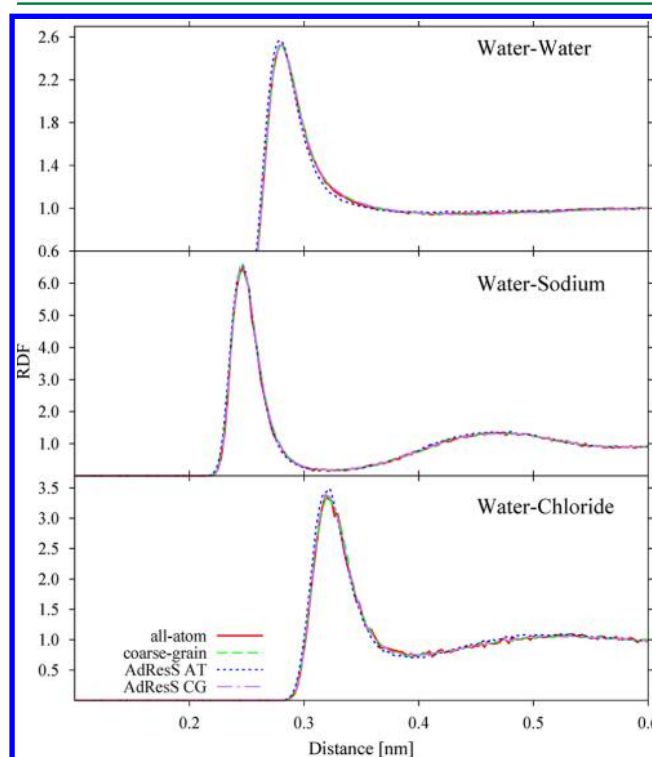
**Multiscale Model of NaCl Salt Solution.** We develop our multiscale salt solution model in the absence of the DNA molecule. Here, we modify the multiscale salt solution model developed in ref 34 to be applicable with the AMBER force field.<sup>49</sup> First, a CG model is derived and parametrized to reproduce structural properties of the full AT model. The

effective potentials for the water–water and water–ion interactions (shown in Figure 2) are obtained using the



**Figure 2.** Effective pair potential interactions between CoMs of molecules for water–water, water–sodium, and water–chloride.

iterative Boltzmann inversion method<sup>55</sup> that is incorporated into the STOCK web toolkit.<sup>56</sup> As can be seen in Figure 3, the

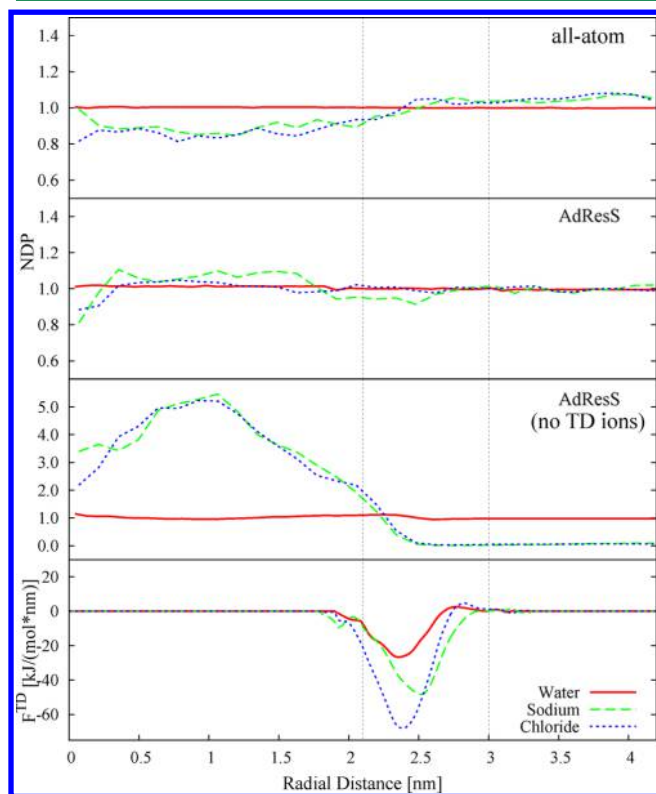


**Figure 3.** Radial distribution functions (RDFs) of water–water (top), water–sodium (middle), and water–chloride (bottom) CoMs from the CG, AdResS, and all-atom simulations. Atomistic RDF is well reproduced by the fully CG and AdResS simulations in both the AT (AdResS AT) and CG (AdResS CG) regions.

effective potentials used for the CG molecules reproduce very well the atomistic radial distribution functions (RDFs). Likewise, the correct local structure is obtained in the subdomains (labeled AdResS AT and AdResS CG) of the AdResS simulation, where only the molecules in the corresponding region are considered. The RDFs were calculated for the water–water (top), water–sodium (middle), and water–chloride (bottom) cases. All calculations are performed on the CoMs of water molecules and ions.



To compensate for the difference in the chemical potential at different levels of resolution and to achieve a uniform density profile throughout the simulation box we apply the TD forces derived as described in section 2. Figure 4 shows the TD forces

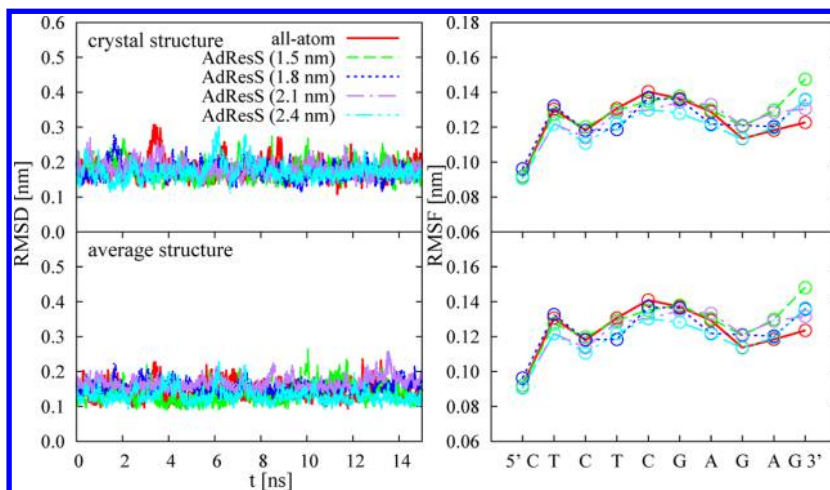


**Figure 4.** Normalized density profiles for water molecules (red) and sodium (green) and chloride (blue) ions. The results are shown for the conventional all-atom simulation and AdResS simulations with the AT region radius size of 2.1 nm. For comparison the results from an additional AdResS simulation are shown where the TD force is added only to water molecules. The bottom plot shows the TD forces applied to all three molecule types.

used for the water molecule and sodium and chloride ion. To verify this point, we plot the normalized density profiles (NDPs) for the corresponding molecule types. The NDPs are computed along the direction of the resolution change, i.e., as a function of distance from the center of the AT region. The results are shown for three simulations: the all-atom, AdResS with the TD forces acting on all molecule types, and AdResS with the TD force acting on water molecules only. The TD forces (see Figure 4) are able to flatten the density profile substantially. Larger deviations from the ideally flat profile are seen for the ions, due to the smaller number of ions in the system compared to the water. However, these deviations are comparable with those obtained from the all-atom simulation.

**Structural Properties of the DNA Molecule.** We first examine the stability of the DNA molecule structure by computing the root-mean-square deviation (rmsd) and the root-mean-square fluctuations (RMSF) of the backbone atoms with respect to the crystal and average structure.<sup>57</sup> Results depicted in Figure 5 show that the variation of the AdResS AT domain was found to have a negligible impact on the averaged DNA structure. The results are computed with respect to the crystal (top) and average (bottom) structure. The crystal structure has the same sequence and was obtained from the Protein Data Bank (pdb 196D). By comparing the values acquired by the various AdResS simulations, we conclude that the variation of the AdResS AT domain does not alter the average DNA structure.

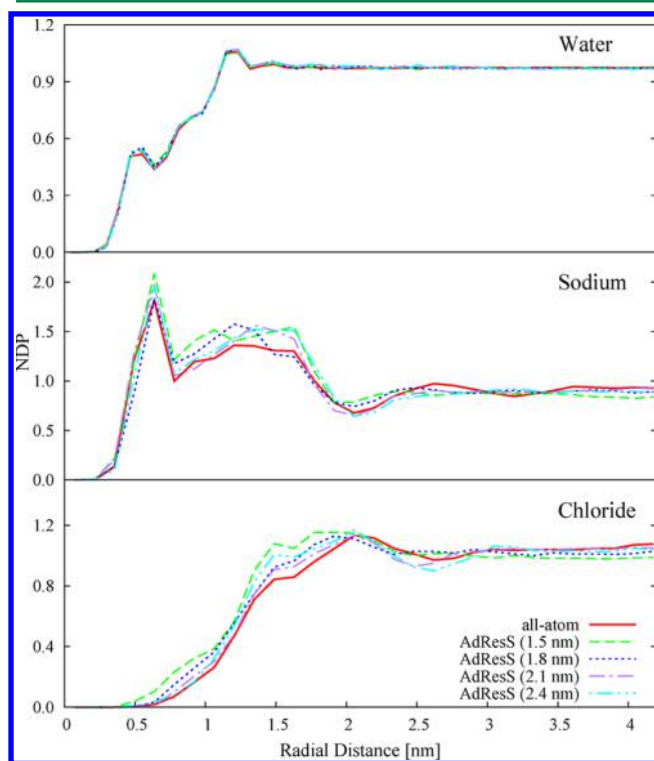
In all cases the multiscale simulation is able to reproduce the results obtained by the conventional full-blown atomistic MD simulation. The rmsd values are lower when the average DNA configuration is taken as a reference. This is to be expected since the structure of DNA molecule is sensitive to a local environment, e.g., solvent. However, the effect is not substantial in our case, and both results indicate that the DNA molecule is in a stable configuration. As mentioned previously, the imposed periodicity in the axial direction fixes the helical twist of DNA and prevents any major bending. The DNA structure is further stabilized by the absence of the terminal residues that are generally more flexible and thus have large RMSF values. Moreover, the terminal residues have an impact on the local



**Figure 5.** Root-mean-square deviation (left) and fluctuations (right) of the backbone heavy atoms with respect to the crystal structure (Protein Data Bank (pdb) entry 196D) and average structure (top and bottom, respectively). We compare the results obtained for five AT region sizes: the full atomistic and cylinders of 1.5, 1.8, 2.1, and 2.4 nm radii. The A, T, C, and G aberrations stand for the adenine, thymine, cytosine, and guanine nucleotides, respectively.

structure of neighboring residues since it has been shown that the local structure of a base pair depends on the type and structure of the flanking base pairs.<sup>58</sup>

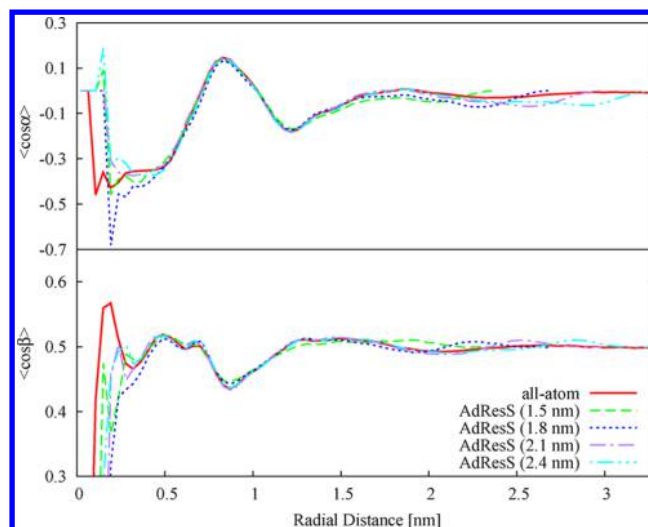
**Structural Properties of the Solvent.** Next, we check that our adaptive resolution simulations correctly reproduce the structure of the surrounding solvent. In Figure 6, we show the



**Figure 6.** Water (top), sodium (middle), and chloride (bottom) NDPs around the CoM of the DNA molecule. The results are shown for the AdResS simulations with AT region radius sizes of 1.5, 1.8, 2.1, and 2.4 nm and compared to the all-atom solvation.

water, sodium, and chloride NDPs around the CoM of the DNA molecule for different AT region sizes. The NDPs are within the error bars in agreement with the reference all-atom simulation. The domain with a decreased water density around the DNA is well within the AT region for all cases. However, the ion atmosphere (region around the DNA where the ion concentrations differ from the bulk) extends in some cases also into the hybrid region, but is nevertheless well reproduced. The plots also show that the chosen simulation box size is large enough to enable the ion concentrations to reach the bulk values at larger distances from the DNA.

Figure 7 shows the average orientations of two vectors, viz., the dipole moment of a water molecule and the vector joining the two hydrogen atoms of the molecule, as a function of the radial distance from the DNA's CoM. We denote the angles formed by these vectors and the normal vector pointing toward the CG region as  $\alpha$  and  $\beta$ , respectively. The average orientations of the vectors are quantified by the average cosine values of these angles. Note that a random orientation corresponds to  $\langle \cos \alpha \rangle = 0$  and  $\langle \cos \beta \rangle = 0.5$ . The different values are due to physically indistinguishable opposite directions for the vector joining the two hydrogen atoms whereas the dipole vector has a specific directionality.<sup>53</sup> The results show that water molecules are highly oriented in the vicinity of the DNA molecule, while further away they have a



**Figure 7.** Average cosine value of the angle formed by the dipole vector (top panel) and the vector joining the two hydrogen atoms of a molecule (bottom panel) with the interface normal vector pointing toward the CG region as a function of the radial distance from the DNA molecule.

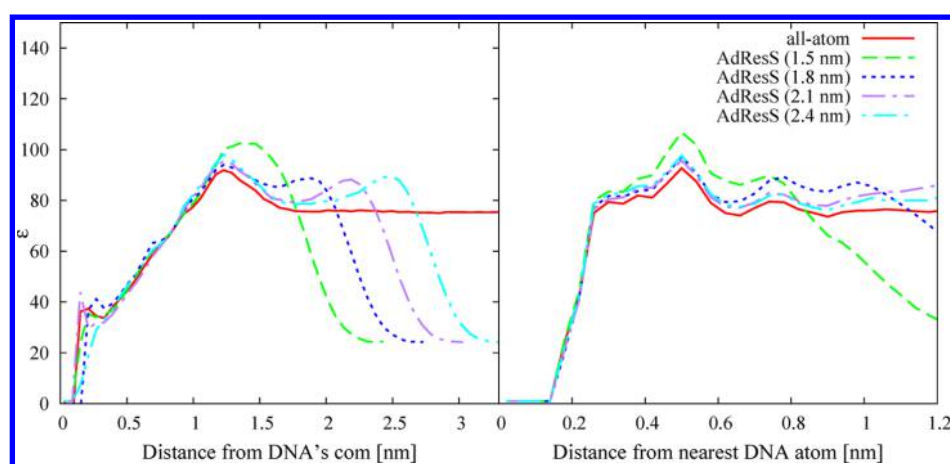
random orientation. All AdResS simulations reproduce the reference all-atom results very well. Slight orientational ordering is, however, observed in the HY region that could be induced by the HY/CG interface in a manner similar to that of a wall. The discrepancies occurring at small distances from DNA molecule are due to the poor statistics.

**Dielectric Properties of the Surrounding Solvent.** To compute the dielectric constant of water around the DNA molecule we employ Kirkwood's theory, where the dielectric constant is related to the average vector sum of the dipole moments of the individual water molecules in a spherical region  $S$ .<sup>6</sup> Using the reaction field treatment for the long-range electrostatic interaction with  $\epsilon_{RF} = 80$ , the dielectric constant is calculated as

$$\epsilon = \frac{1 + [\rho g \mu^2 / 3 \epsilon_0 k_B T] [2 \epsilon_{RF} / (2 \epsilon_{RF} + 1)]}{1 - [\rho g \mu^2 / 3 \epsilon_0 k_B T] [1 / (2 \epsilon_{RF} + 1)]};$$

$$g = 1 + \left\langle \sum_{j \in S} \cos \theta_{ij} \right\rangle_i \quad (3)$$

where  $g$  is the Kirkwood  $g$  factor,  $\mu$  is the external dipole moment of the water molecule ( $\mu = 2.337$  D for TIP3P water model),  $k_B$  is the Boltzmann constant, and  $T$  is the absolute temperature. The dipole density  $\rho$  is calculated as  $\langle N_i / (V - V_{ex,i}) \rangle_i$  where  $V$  is the volume of a Kirkwood sphere  $S$  centered on a given water molecule  $i$ ,  $N_i$  is the number of water molecules with indices  $j = 1, N_i$  in the Kirkwood sphere  $S$  around the reference water molecule  $i$ ,  $\langle \dots \rangle_i$  denotes an average over all water molecules in the set, and  $\theta_{ij}$  the angle between dipoles of water molecules  $i$  and  $j$ . Excluded volume  $V_{ex,i}$  due to DNA atoms and ions is computed for each water molecule in each configuration by a grid representation of the spherical region surrounding the reference water molecule. Grid points that fall within the van der Waals envelope of any DNA atom or ion are defined to be in the excluded volume. The radius of the Kirkwood sphere was set to  $R_c = 0.591$  nm. In accordance with previous reports<sup>6</sup> we find the magnitude of the dielectric constant to be only weakly dependent on the choice of this free parameter.



**Figure 8.** Calculated relative permittivities of water around DNA shown as a function of distance from DNA's CoM (left) and as a distance to the nearest DNA atom (right). The proximity criterion<sup>59</sup> is employed to assign reference waters to DNA atoms. The results are shown for the AdResS simulations with AT region radius sizes of 1.5, 1.8, 2.1, and 2.4 nm and compared to the all-atom solvation.

Figure 8 (left) shows the spatial dielectric profile of water around the DNA molecule. The results are computed for the water molecules in the AT and HY regions. Spatially varying dielectric constant is computed by discretizing distances from the DNA molecule into bins and taking the average  $\langle \dots \rangle_b$  in eq 3, over water molecules that fall into a corresponding bin. The dielectric constant increases with distance from the DNA's CoM and reaches the bulk value around 1.7 nm. As expected, the bulk dielectric constant is smaller compared to the value for pure water (82 for TIP3P water model) because the movement of water is restricted by the ions. Before the bulk value is reached there is a small region, around 1.2 nm away from the DNA molecule, where the dielectric constant is higher than in the bulk. This region coincides with the region of slightly increased water density (see the normalized density profile in Figure 6). In all AdResS simulations the dielectric constant is observed to decrease in the HY region. As the water molecules move from the AT to the HY region, the electrostatic interactions are gradually switched off and the molecules lose their rotational freedom. As a result of this rotational freezing the dielectric constant is lowered.

We also observe a rise in the dielectric constant at the AT/HY transition. This effect is pronounced for the case where the AT region size is equal to 1.5 nm, since the AT/HY transition is very close to the previously mentioned region of increased dielectric constant. Note that in the HY region there is a slight orientational order since the average cosine value of the angle formed by the water's dipole moment and the normal vector pointing toward the CG region is not equal to zero (see Figure 7).

To identify the dielectric constant in the first and higher hydration shells of the DNA molecule, we plot the dielectric profile also as a function of distance to the nearest DNA atom (Figure 8 right). Here we assign each water molecule to a DNA atom using the proximity criterion.<sup>59</sup> Comparing the AdResS and all-atom results, we can see that the larger the AT region is in the AdResS simulation the larger is the number of hydration shells in which the dielectric constant can be faithfully reproduced. However, the largest variations in the dielectric profile are occurring within the first hydration shell, and these are well reproduced by all AdResS simulations.

Next, we examine the dielectric behavior of water in the specific regions around the DNA, namely, the backbone and

the minor and major grooves. We take into account only the water molecules in the first hydration shell. The results are summarized in Table 1. Within the error bars no notable

**Table 1. Relative Dielectric Constant of Water Molecules in the First Hydration Shell (Within 0.3 nm Distance from the DNA Surface)<sup>a</sup>**

region	$\epsilon(1.5\text{nm})$	$\epsilon(1.8\text{nm})$	$\epsilon(2.1\text{nm})$	$\epsilon(2.4\text{nm})$	$\epsilon(\infty)$
backbone	68.6	65.3	65.8	66.8	63.7
minor groove	39.4	37.8	35.2	38.2	36.4
major groove	45.8	48.0	45.7	47.0	46.7
all	62.6	60.5	60.2	61.4	59.0

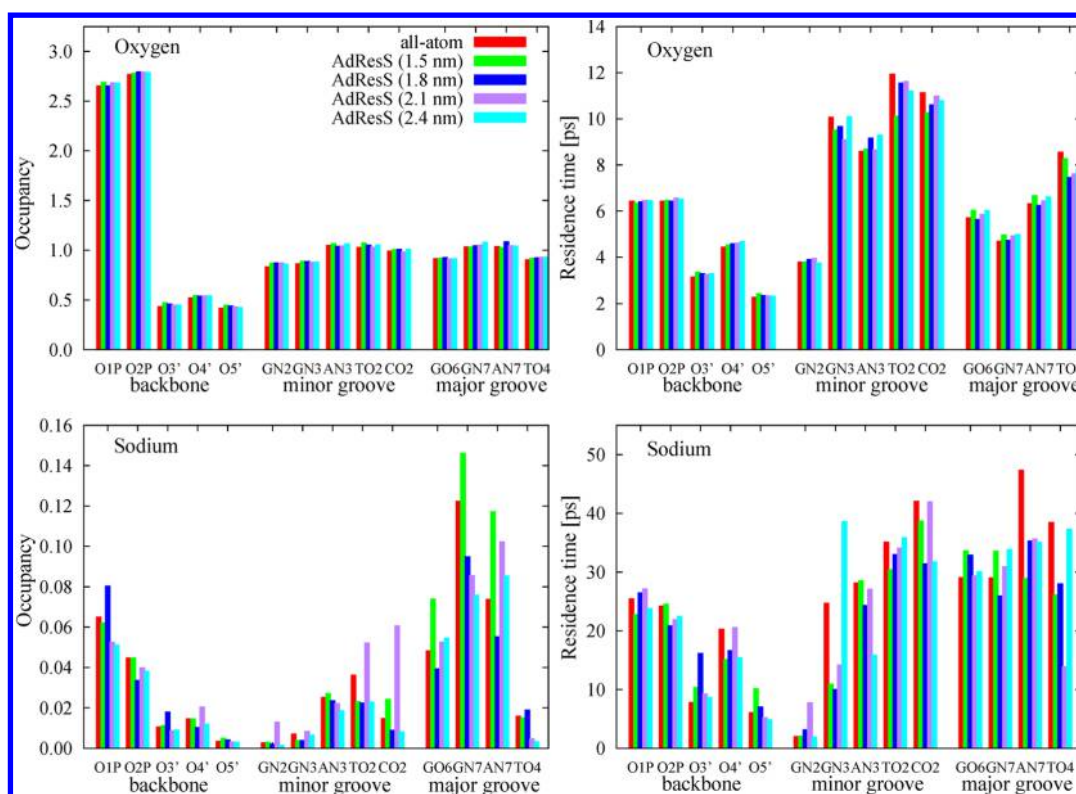
<sup>a</sup>Water molecules are appointed to DNA atoms based on the proximity criterion. We divide the DNA atoms in the following way: phosphodiester backbone {P, O1P, O2P, O5', C5', C4', O4', C1', C3', C2', O3'}, major groove {A: C8, N7, N6, C6, C5; C: C6, C5, C4, N4; G: C8, N7, C6, O6, C5; T: C7, C6, C5, O4}, and minor groove {A: C4, N3, C2, N1; C: C2, O2; G: C4, N3, N2, C2; T: N3, C2, O2}. Error bars are approximately 5%.

differences are found between the AdResS and reference all-atom simulations. In all cases the dielectric constant exhibits the following tendency:  $\epsilon(\text{phosphate backbone}) > \epsilon(\text{major groove}) > \epsilon(\text{minor groove})$ . The observed trend is in agreement with existing experimental<sup>10,11</sup> and numerical simulation results.<sup>6</sup>

It is presumed that the dielectric constant of water in the minor groove is lower because the water molecules in this region are more ordered. Our analysis of average occupancy and residence time of oxygen atoms of water in the first solvation shell of the electronegative atoms of the DNA suggests that the exchange of waters in the minor groove is slower compared to other regions. In particular, we find the minor and major grooves to be equally hydrated, but the residence times are higher in the minor groove (9–10 ps) than in the major groove (5–8 ps) or in the backbone (2–6 ps): see Figure 9.

From results shown in Figure 9 we can again conclude that the AdResS simulations are able to reproduce the atomistic results. The discrepancies between results of different simulations are small for the oxygen but somewhat larger for the Na<sup>+</sup>. The most hydrated sites are the backbone O1P and O2P atoms having on average 2.5 or more water molecules in the first hydration shell. The hydration of the minor and major





**Figure 9.** Average occupancy (left) and residence time (right) of  $\text{Na}^+$  and oxygen atoms of water in the first hydration shell of the electronegative atoms of DNA. Fast atom fluctuations ( $<1$  ps) were ignored in the calculation of the average residence time.

grooves is relatively uniform with about one water molecule. However, the residence times in the minor groove are almost 40% higher than the average residence times in the major groove and the backbone. The  $\text{Na}^+$  atoms preferentially bind to the guanine base in the major groove. Both grooves, however, exhibit similar residence times which are slightly higher than in the backbone.

Some further insight can also be obtained from the dipole autocorrelation function of water molecules that are initially located in the minor or major groove or backbone region. To this end, we compute the dipole autocorrelation function  $d_{ACF}$  for the water molecules in the first hydration shell around the DNA molecule. We consider separately the water molecules that are initially in the minor and major grooves and the backbone. From the results in Figure 10 we observe that the  $d_{ACF}$  for the water molecules in the minor groove decays significantly more slowly than does  $d_{ACF}$  for waters in the major groove or along the backbone.

**Dielectric Properties of the DNA Molecule.** Another route for computation of the static dielectric constant is the fluctuations of the total dipole moment. If the system is surrounded by an infinite dielectric medium (with  $\epsilon_{RF}$ ), then the dielectric constant  $\epsilon$  of the system is<sup>60</sup>

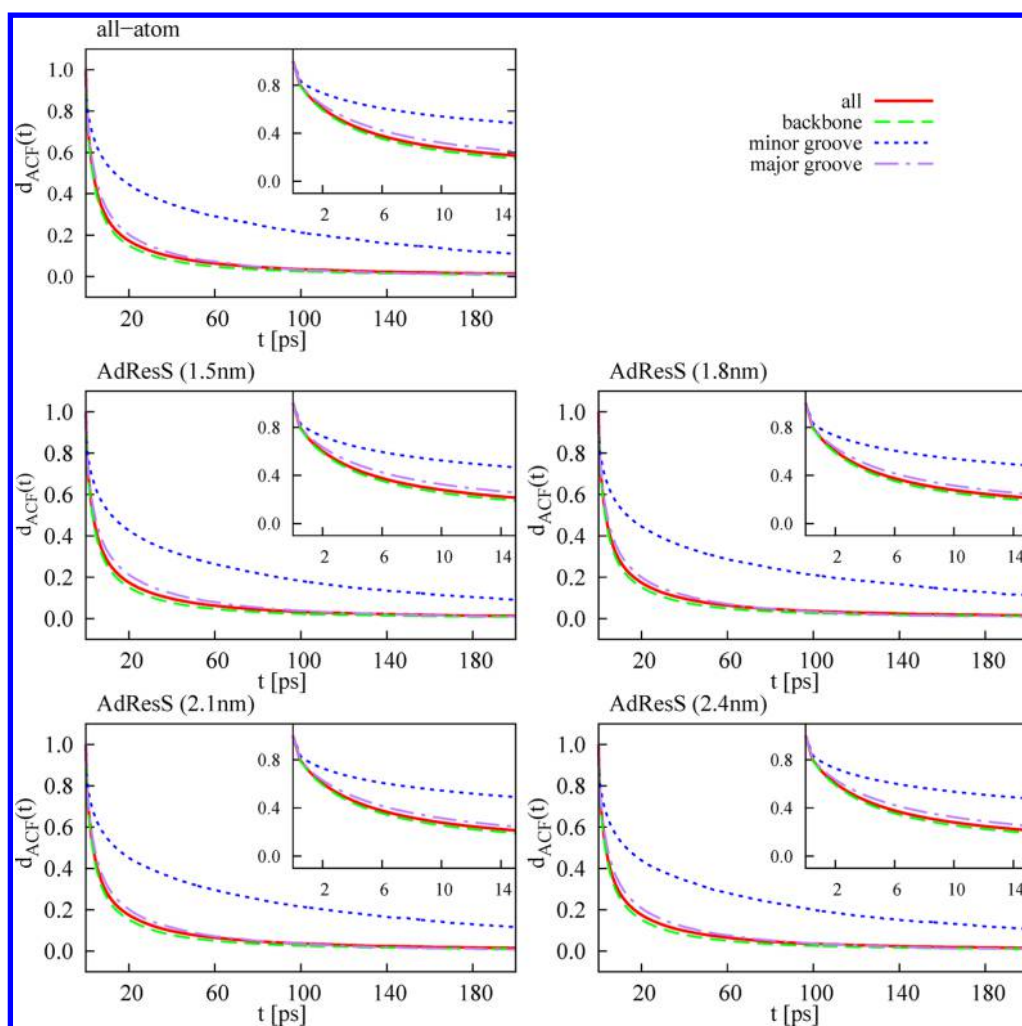
$$\epsilon = \frac{1 + [(\langle M^2 \rangle - \langle M \rangle^2) / 3\epsilon_0 V k_B T] [2\epsilon_{RF} / (2\epsilon_{RF} + 1)]}{1 - [(\langle M^2 \rangle - \langle M \rangle^2) / 3\epsilon_0 V k_B T] [1 / (2\epsilon_{RF} + 1)]} \quad (4)$$

where  $M$  and  $V$  are the total dipole moment and volume of the system. The total dipole fluctuation can be separated into groups if the cross-term between two groups is much smaller than the self-terms of the two groups.<sup>8</sup> The dielectric constants of parts of the DNA molecule can then be calculated using eq 4,

where  $V$  is the volume of a group. Prior to dipole moment evaluation the atom's trajectories have to be unfolded to avoid the discontinuities that arise due to periodic boundary conditions. In order to take into account only the atom's fluctuations, each configuration is translated so that the origin coincides with the DNA's CoM. The rotational motion of the DNA can be eliminated, as for instance in the rmsd calculation, by fitting each configuration to the initial one.

The results for the phosphate, sugar, and base components of the DNA and the total DNA molecule are shown in Table 2. In Figure 11 we plot the cumulative averages of the total dipole moment fluctuations as a function of time for the phosphate, sugar, and base component of the DNA and the whole DNA molecule. We plot the obtained values for the AdResS simulations with AT region radius sizes of 1.5, 1.8, 2.1, and 2.4 nm and the reference atomistic simulation. The results for the base component are in very good agreement while larger differences are found for the phosphate and sugar.

The relative dielectric constant for the phosphate is found to be around 17, while the base and sugar components have relative dielectric constants around 2–3. The calculated dielectric constant is somewhat different for different simulations. However, we believe that this is due to the statistical error and not due to the AdResS simulation since no specific trend is observed in the dielectric constant as the AT region is increased. For the DNA molecule our simulations give the relative dielectric constant of about 5. A possible explanation for obtaining slightly lower dielectric constant might be in the DNA environment difference; i.e., our simulations are performed at the 1 M salt concentration, while the simulations in refs 3 and 8 only include the minimum number of  $\text{Na}^+$  for charge neutrality. As for the difference with the experimental value of  $\sim 8$ ,<sup>3</sup> one needs to point out that the



**Figure 10.** Dipole autocorrelation function  $d_{ACF}$  as a function of time. We average over water molecules that are initially in the first hydration shell of DNA and over all time origins. On the basis of the proximity analysis water molecules are assigned to either minor or major groove or the backbone. The results are shown for the AdResS simulations with AT region radius sizes of 1.5, 1.8, 2.1, and 2.4 nm and the reference all-atom simulation. Insets depict a zoom into time scales below the residence times. The plots show that the water molecules in the minor groove reorient significantly more slowly than the waters in the major groove or along the backbone.

**Table 2. Relative Dielectric Constants of the DNA Molecule and Its Components Computed from the Total Dipole Fluctuations<sup>a</sup>**

group	$V$ (nm <sup>3</sup> )	$\epsilon(1.5\text{nm})$	$\epsilon(1.8\text{nm})$	$\epsilon(2.1\text{nm})$	$\epsilon(2.4\text{nm})$	$\epsilon(\infty)$
phosphate	2.48	18.3	21.1	17.3	15.7	17.0
sugar	2.51	2.7	2.8	2.7	2.4	2.6
base	3.03	2.1	2.1	2.0	2.0	2.1
DNA	8.02	5.6	5.9	5.4	4.5	5.0

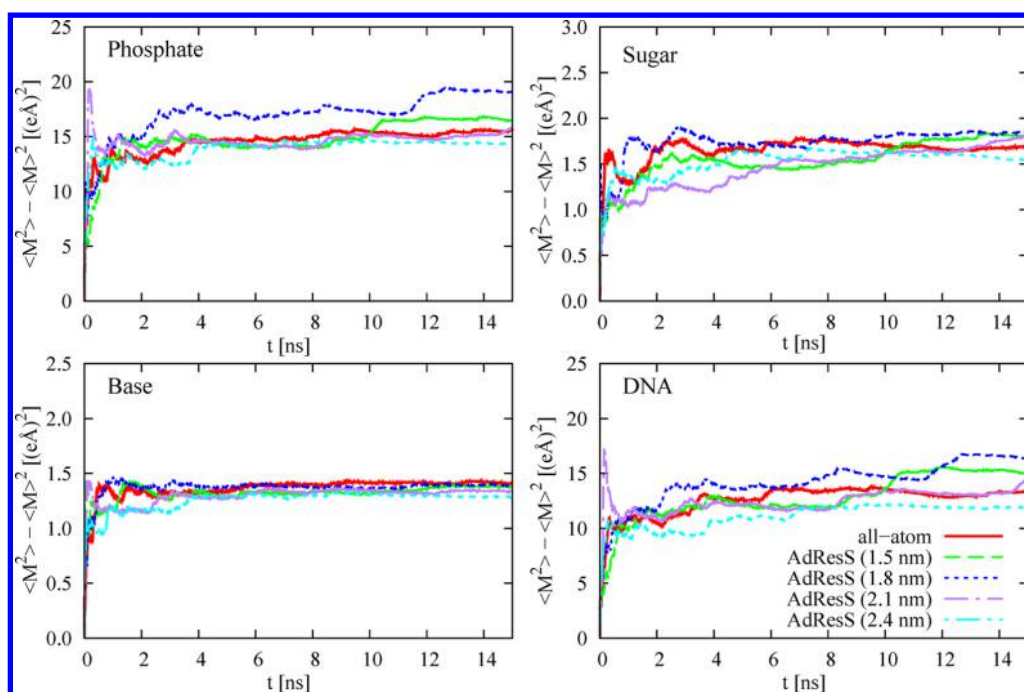
<sup>a</sup>The corresponding volumes were calculated via double cubic lattice method.<sup>61</sup> The convergence of the total dipole moment fluctuations is shown in Figure 11. The error of the dielectric constant is approximately 10%.

experiment actually gives the dielectric response of a condensed array of DNA molecules within the bacteriophage capsid,<sup>62</sup> a feature not taken into account in any simulation. The difference in the orientational fluctuations that are in fact absent in the simulation and the amount of ordered vicinal vs disordered distal water are all factors that could be implicated in the different values of the dielectric constant in the model system and the experiment, as are of course also the inaccuracies in the parametrization of the atomistic force fields. We cannot easily single out or discard a single mechanism for this discrepancy.

## 4. CONCLUSIONS

We have presented a detailed multiscale simulation of a DNA molecule in a bathing electrolyte solution employing a hybrid AT/CG solvent model. The AT domain was limited to a cylinder embedding the DNA, while a CG water model was used farther away. This leads to a speedup in comparison to an all-atom simulation due to the salt solution CG model, which reduces the number of water degrees of freedom and introduces softer interactions.<sup>34</sup> The actual speedup of AdResS simulations thus depends on the ratio between the AT and CG domain sizes. We showed that the multiscale simulations can





**Figure 11.** Accumulative mean square deviation of the total dipole moment as a function of time. The results are shown separately for the phosphate, sugar and base atoms in the DNA and for the whole DNA molecule.

reproduce very well the statistical properties of the DNA molecule. The study was centered on the collective dielectric properties of the molecular environment surrounding the DNA, as well as on the DNA molecule itself.

Since the Poisson–Boltzmann based techniques and other continuum calculations of the electrostatic potential and counterion distributions around DNA are profoundly affected by the assumed dielectric model parameters,<sup>63</sup> we believe that an accurate determination of the solvent and DNA dielectric constants presented in this work is especially useful in this particular context. The presented results could also be valuable to simulations with implicit solvent that incorporate empirical distance-dependent “effective” dielectric permittivities.<sup>64</sup>

## AUTHOR INFORMATION

### Corresponding Authors

\*E-mail: rudolf.podgornik@fmf.uni-lj.si.

\*E-mail: praprot@cmm.ki.si.

### Funding

J.Z. and M.P. acknowledge financial support through Grants J1-4134 and P1-0002 from the Slovenian Research Agency. R.P. acknowledges financial support through Grant P1-0055 from the Slovenian Research Agency.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank K. Kremer and S. Bevc for useful discussions.

## REFERENCES

- (1) Stanley, C.; Rau, D. *Curr. Opin. Colloid Interface Sci.* **2011**, *16*, 551–556.
- (2) Yasar, S.; Podgornik, R.; Valle-Orero, J.; Johnson, M.; Parsegian, V. A. *Sci. Rep.* **2014**, *4*, 6877.

- (3) Cuervo, A.; Dans, P. D.; Carrascosa, J. L.; Orozco, M.; Gomila, G.; Fumagalli, L. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, E3624–E3630.
- (4) Gavryushov, S. *J. Phys. Chem. B* **2008**, *112*, 8955–8965.
- (5) Lamm, G.; Pack, G. R. *J. Phys. Chem. B* **1997**, *101*, 959–965.
- (6) Young, M. A.; Jayaram, B.; Beveridge, D. L. *J. Phys. Chem. B* **1998**, *102*, 7666–7669.
- (7) Li, L.; Li, C.; Zhang, Z.; Alexov, E. *J. Chem. Theory Comput.* **2013**, *9*, 2126–2136.
- (8) Yang, L.; Weerasinghe, S.; Smith, P. E.; Pettitt, B. M. *Biophys. J.* **1995**, *69*, 1519–1527.
- (9) Bagchi, B. *Chem. Phys. Lett.* **2012**, *529*, 1–9.
- (10) Barawkar, K. N.; Ganesh, A. *Nucleic Acids Res.* **1995**, *23*, 159–164.
- (11) Jin, R.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 8939–8942.
- (12) Lipfert, J.; Doniach, S.; Das, R.; Herschlag, D. *Annu. Rev. Biochem.* **2014**, *83*, 813–841.
- (13) Potoyan, D. A.; Savelyev, A.; Papoian, G. A. *WIREs Comput. Mol. Sci.* **2013**, *3*, 69–83.
- (14) Snodin, B. E. K.; Randisi, F.; Mosayebi, M.; Sulc, P.; Schreck, J. S.; Romano, F.; Ouldrige, T. E.; Tsukanov, R.; Nir, E.; Louis, A. A.; Doye, J. P. K. *J. Chem. Phys.* **2015**, *142*, 234901.
- (15) Maciejczyk, M.; Spasic, A.; Liwo, A.; Scheraga, H. A. *J. Chem. Theory Comput.* **2014**, *10*, 5020–5035.
- (16) Knotts, T. A. I.; Rathore, N.; Schwartz, D. C.; de Pablo, J. J. *J. Chem. Phys.* **2007**, *126*, 084901.
- (17) Wassenaar, T. A.; Ingólfsson, H. I.; Priess, M.; Marrink, S. J.; Schaefer, L. V. *J. Phys. Chem. B* **2013**, *117*, 3516–3530.
- (18) Neri, M.; Anselmi, C.; Cascella, M.; Maritan, A.; Carloni, P. *Phys. Rev. Lett.* **2005**, *95*, 218102.
- (19) Rzepiela, A. J.; Louhivuori, M.; Peter, C.; Marrink, S. J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10437–10448.
- (20) Han, W.; Schulten, K. *J. Chem. Theory Comput.* **2012**, *8*, 4413–4424.
- (21) Sokkar, P.; Choi, S. M.; Rhee, Y. M. *J. Chem. Theory Comput.* **2013**, *9*, 3728–3739.
- (22) Gonzalez, H. C.; Darré, L.; Pantano, S. J. *Phys. Chem. B* **2013**, *117*, 14438–14448.

- (23) Nielsen, S. O.; Moore, P. B.; Ensing, B. *Phys. Rev. Lett.* **2010**, *105*, 237802.
- (24) Heyden, A.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 217–221.
- (25) Sokkar, P.; Boulanger, E.; Thiel, W.; Sanchez-Garcia, E. *J. Chem. Theory Comput.* **2015**, *11*, 1809–1818.
- (26) Rudzinski, J. F.; Noid, W. G. *J. Chem. Theory Comput.* **2015**, *11*, 1278–1291.
- (27) Praprotnik, M.; Delle Site, L.; Kremer, K. *J. Chem. Phys.* **2005**, *123*, 224106.
- (28) Praprotnik, M.; Delle Site, L.; Kremer, K. *Annu. Rev. Phys. Chem.* **2008**, *59*, 545–571.
- (29) Praprotnik, M.; Poblete, S.; Kremer, K. *J. Stat. Phys.* **2011**, *145*, 946–966.
- (30) Wang, H.; Hartmann, C.; Schütte, C.; Delle Site, L. *Phys. Rev. X* **2013**, *3*, 011018.
- (31) Agarwal, A.; Wang, H.; Schütte, C.; Delle Site, L. *J. Chem. Phys.* **2014**, *141*, 034102.
- (32) Potestio, R.; Español, P.; Delgado-Buscalioni, R.; Everaers, R.; Kremer, K.; Donadio, D. *Phys. Rev. Lett.* **2013**, *111*, 060601.
- (33) Potestio, R.; Fritsch, S.; Español, P.; Delgado-Buscalioni, R.; Kremer, K.; Everaers, R.; Donadio, D. *Phys. Rev. Lett.* **2013**, *110*, 108301.
- (34) Bevc, S.; Junghans, C.; Kremer, K.; Praprotnik, M. *New J. Phys.* **2013**, *15*, 105007.
- (35) Zavadlav, J.; Melo, M. N.; Marrink, S. J.; Praprotnik, M. *J. Chem. Phys.* **2014**, *140*, 054114.
- (36) Zavadlav, J.; Melo, M. N.; Cunha, A. V.; de Vries, A. H.; Marrink, S. J.; Praprotnik, M. *J. Chem. Theory Comput.* **2014**, *10*, 2591–2598.
- (37) Zavadlav, J.; Melo, M. N.; Marrink, S. J.; Praprotnik, M. *J. Chem. Phys.* **2015**, *142*, 244118.
- (38) Fogarty, A. C.; Potestio, R.; Kremer, K. *J. Chem. Phys.* **2015**, *142*, 195101.
- (39) Lambeth, B. P., Jr.; Junghans, C.; Kremer, K.; Clementi, C.; Delle Site, L. *J. Chem. Phys.* **2010**, *133*, 221101.
- (40) van Dijk, M.; Bonvin, A. M. J. *J. Nucleic Acids Res.* **2009**, *37*, W235–W239.
- (41) Korolev, N.; Lyubartsev, A. P.; Laaksonen, A.; Nordenskiöld, L. *Nucleic Acids Res.* **2003**, *31*, 5971–5981.
- (42) Lyubartsev, A. P.; Laaksonen, A. *J. Biomol. Struct. Dyn.* **1998**, *16*, 579–92.
- (43) Fritsch, S.; Poblete, S.; Junghans, C.; Ciccotti, G.; Delle Site, L.; Kremer, K. *Phys. Rev. Lett.* **2012**, *108*, 170602.
- (44) Poblete, S.; Praprotnik, M.; Kremer, K.; Delle Site, L. *J. Chem. Phys.* **2010**, *132*, 114101.
- (45) Halverson, J. D.; Brandes, T.; Lenz, O.; Arnold, A.; Bevc, S.; Starchenko, V.; Kremer, K.; Stuehn, T.; Reith, D. *Comput. Phys. Commun.* **2013**, *184*, 1129–1149.
- (46) Andersen, H. C. *J. Comput. Phys.* **1983**, *52*, 24–34.
- (47) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952–962.
- (48) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (49) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999–2012.
- (50) Ricci, C. G.; de Andrade, A. S. C.; Mottin, M.; Netz, P. A. *J. Phys. Chem. B* **2010**, *114*, 9882–9893.
- (51) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451–5459.
- (52) Cisneros, G. A.; Karttunen, M.; Ren, P.; Sagui, C. *Chem. Rev.* **2014**, *114*, 779–814.
- (53) Praprotnik, M.; Matysiak, S.; Delle Site, L.; Kremer, K.; Clementi, C. *J. Phys.: Condens. Matter* **2007**, *19*, 292201.
- (54) Maffeo, C.; Yoo, J.; Comer, J.; Wells, D. B.; Luan, B.; Aksimentiev, A. *J. Phys.: Condens. Matter* **2014**, *26*, 413101.
- (55) Reith, D.; Pütz, M.; Müller-Plathe, F. *J. Comput. Chem.* **2003**, *24*, 1624–1636.
- (56) Bevc, S.; Junghans, C.; Praprotnik, M. *J. Comput. Chem.* **2015**, *36*, 467–477.
- (57) Nina, M.; Simonson, T. *J. Phys. Chem. B* **2002**, *106*, 3696–3705.
- (58) Lavery, R.; Zakrzewska, K.; Beveridge, D.; Bishop, T. C.; Case, D. A.; Cheatham, T., III; Dixit, S.; Jayaram, B.; Lankas, F.; Laughton, C.; Maddocks, J. H.; Michon, A.; Osman, R.; Orozco, M.; Perez, A.; Singh, T.; Spackova, N.; Sponer, J. *Nucleic Acids Res.* **2010**, *38*, 299–313.
- (59) Mehrotra, P. K.; Beveridge, D. L. *J. Am. Chem. Soc.* **1980**, *102*, 4287–4294.
- (60) Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 3169–3174.
- (61) Eisenhaber, F.; Lijnzaad, P.; Argos, P.; Sander, C.; Scharf, M. *J. Comput. Chem.* **1995**, *16*, 273–284.
- (62) Šiber, A.; Lošdorfer Božič, A.; Podgornik, R. *Phys. Chem. Chem. Phys.* **2012**, *14*, 3746–3765.
- (63) Chu, V.; Bai, Y.; Lipfert, J.; Herschlag, D.; Doniach, S. *Curr. Opin. Chem. Biol.* **2008**, *12*, 619–625.
- (64) Cherstvy, A. *Phys. Chem. Chem. Phys.* **2011**, *13*, 9942–9968.

#### ■ NOTE ADDED AFTER ASAP PUBLICATION

This paper was published ASAP on September 17, 2015, with errors in the Results and Discussion Section following equation 3. The corrected version reposted October 2, 2015.