

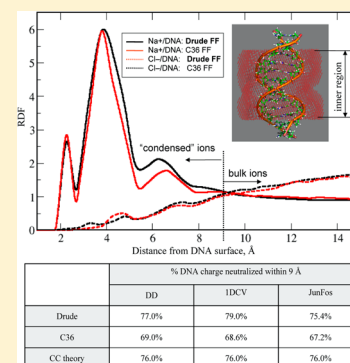
Balancing the Interactions of Ions, Water, and DNA in the Drude Polarizable Force Field

Alexey Savelyev and Alexander D. MacKerell, Jr.*

Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, Maryland 21201, United States

Supporting Information

ABSTRACT: Recently we presented a first-generation all-atom Drude polarizable force field for DNA based on the classical Drude oscillator model, focusing on optimization of key dihedral angles followed by extensive validation of the force field parameters. Presently, we describe the procedure for balancing the electrostatic interactions between ions, water, and DNA as required for development of the Drude force field for DNA. The proper balance of these interactions is shown to impact DNA stability and subtler conformational properties, including the conformational equilibrium between the BI and BII states, and the A and B forms of DNA. The parametrization efforts were simultaneously guided by gas-phase quantum mechanics (QM) data on small model compounds and condensed-phase experimental data on the hydration and osmotic properties of biologically relevant ions and their solutions, as well as theoretical predictions for ionic distribution around DNA oligomer. In addition, fine-tuning of the internal base parameters was performed to obtain the final DNA model. Notably, the Drude model is shown to more accurately reproduce counterion condensation theory predictions of DNA charge neutralization by the condensed ions as compared to the CHARMM36 additive DNA force field, indicating an improved physical description of the forces dictating the ionic solvation of DNA due to the explicit treatment of electronic polarizability. In combination with the polarizable DNA force field, the availability of Drude polarizable parameters for proteins, lipids, and carbohydrates will allow for simulation studies of heterogeneous biological systems.



INTRODUCTION

The conformational behavior of DNA in solution is driven by a subtle interplay among various intramolecular degrees of freedom and interactions with the environment.¹ For example, when in a B form, correlated motions of the backbone, sugar and glycosidic DNA torsions lead to the dynamical transitions between BI and BII states, as well as the extent of local sampling of geometries characteristic of the A form DNA as manifested in sugar puckering transitions between the south and north conformations.^{2–6} On a global scale, the balance between bonded and nonbonded intramolecular parameters along with interactions with the environment influence the conformational equilibrium between the A, B, and Z forms of DNA. Moreover, interactions with mobile ions and hydration effects are known to contribute to the overall DNA stability and regulate a number of critical polymeric properties, such as persistence length and stiffness.

Over the past decade, computational modeling of DNA had some successes in predicting selected properties of DNA and helped to rationalize many DNA related experiments. In particular, modeling allowed for atomistic investigations of DNA energetics, mechanisms of structural transitions and the impact of environmental effects on DNA^{8–12}—information that is not readily accessible to experimental techniques, such as X-ray crystallography and NMR.¹³ These studies were based on additive force fields (FF) for DNA, including CHARMM,^{11,14} AMBER,^{15,16} Bristol-Myers Squibb,² and GROMOS.¹⁷ While

successful in many respects,^{18–21} these models show a number of limitations, potentially associated with the lack of explicit electronic polarization in the model, including limited treatment in the equilibrium between the different conformations of DNA^{22–24} as well as relative energies of different conformations.²⁵ While continued developments in the additive FFs have partially overcome these limitations,^{11,26,27} it is anticipated that extending the form of the potential energy function to include the explicit treatment of electronic polarization may lead to improvements in DNA FFs, including yielding a more accurate description of the underlying physical forces dictating the structure and dynamics of DNA.

Recently, we have presented the first-generation CHARMM polarizable FF for DNA based on the classical Drude oscillator formalism.²⁸ Briefly, electronic induction in the model is represented by introducing a massless charged particle (i.e., the Drude oscillator) attached to each polarizable atom by a harmonic spring.²⁹ In practice, charge redistribution in response to changes in the local electric field is approximated by self-consistently updating the positions of the auxiliary particles to their local energy minima for any given configuration of the atoms in the system.^{30–33} This self-consistent field (SCF) approach takes into account the

Received: April 8, 2014

Revised: May 28, 2014

Published: May 29, 2014

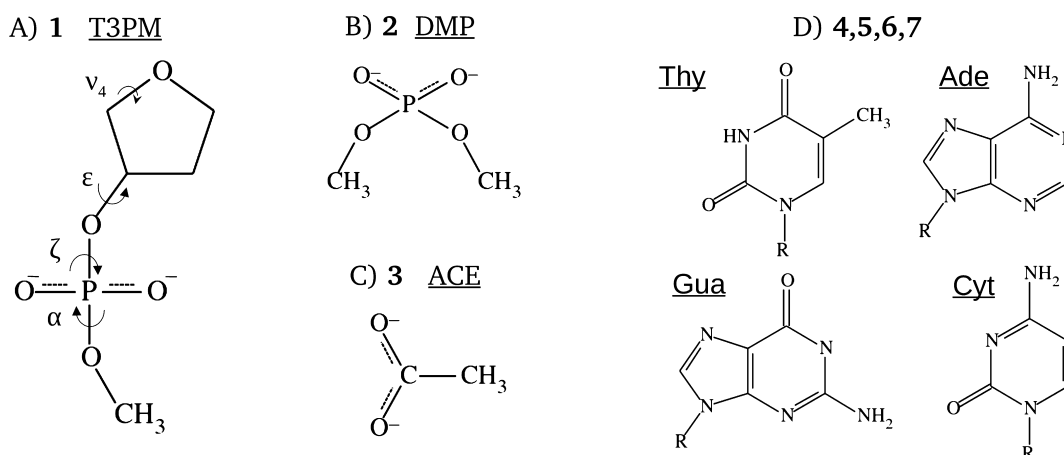


Figure 1. Model compounds used in the present study. Dihedral angles shown for **1** are defined as follows: $\alpha = \text{O}3'-\text{P}-\text{O}5'-\text{C}5'$, $\epsilon = \text{C}4'-\text{C}3'-\text{O}3'-\text{P}$, $\zeta = \text{C}3'-\text{O}3'-\text{P}-\text{O}5'$ and $\nu_4 = \text{C}3'-\text{C}4'-\text{O}4'-\text{C}1'$. In **4** and **5**, R = CH_3 or H.

permanent electric field due to the fixed charges and the contribution of the induced dipoles to the electric field. In the context of molecular dynamics (MD) simulations, the SCF calculation is substituted with an Extended Lagrangian Integrator^{34–36} that treats the Drude oscillators as dynamic variables, allowing for the computational efficiency required for long-time MD simulations. To date, the Drude model has been implemented in CHARMM,³⁷ NAMD³⁸ and ChemShell quantum mechanics/molecular mechanics (QM/MM).³⁹ As discussed elsewhere, due to the need of use of a 1 fs integration time step (versus the 2 fs time step customarily used in additive MD simulations), use of the Drude polarizable FF in simulations yields an approximately 4-fold overhead compared to additive simulations.^{40,41} In addition, a module, the “Drude Prepper” has been added to the CHARMM-GUI,⁴² facilitating the conversion of CHARMM PSF and coordinate files into the corresponding files required for Drude-based simulations including example inputs for CHARMM and NAMD.

In our published study on development of the Drude force field for DNA we focused mainly on details of parametrization of the highly correlated backbone, sugar, and glycosidic torsion angles, and on extensive validation of the final parameters by comparing MD simulation results with a wide range of experimental data on DNA in solution.²⁸ A notable observation from those simulations was the significant enhancement and variation of base dipole moments as compared to the C36 additive FF,¹¹ indicating that the underlying physical forces dictating the properties of DNA are significantly different in the polarizable model. However, the discussion in that paper covered only a small part of the optimization process, not addressing such critical aspects of the modeling as fine-tuning of the interactions between DNA and mobile ions and the balance between DNA electrostatic and hydration effects, subtle corrections in the internal base parameters, and the impact of these on the overall DNA stability and subtler conformational phenomena, including BI/BII transitions and A-to-B equilibrium.

Presently, we focus on the above aspects of the Drude FF development and additional testing of the optimized DNA parameters against a range of experimental data and theoretical estimations. Our optimization procedure involves approaches ranging from QM calculations on various model compounds and their complexes to investigating the hydration and osmotic properties of the chemically relevant ionic systems to studying

the impact of the mobile ions on DNA structural stability. We note that the final results presented here are based on the identical set of parameters developed in our earlier study on the Drude force field for DNA. Some results are also presented for a “default” parameter set, which is that based on direct transfer of parameters optimized targeting small model compounds representative of DNA with only minimal adjustment of selected dihedral parameters.

COMPUTATIONAL METHODS

Parameter Optimization Strategy. In the present study we used a number of small model systems to probe how electrostatic, hydration and other effects in the DNA duplex impact overall DNA stability, as well as subtler aspects of DNA structural behavior in solution. Those systems are shown in Figure 1.

Model compound **1**, composed of the sugar and a 3' methylated phosphate group (Figure 1) was previously used to study correlations between ϵ and ξ torsions in the context of BI/BII transitions in DNA.^{11,28} In our recent parametrization study we discussed in detail the optimization of these and other dihedral angles.²⁸ In the present study, by analyzing individual energetic contributions of **1**, we demonstrate that, in addition to the torsion parameters, adjustments of electrostatic parameters (partial charges) was necessary to achieve satisfactory agreement with both QM data on this model compound and experimental data on BI/BII transitions and A-to-B equilibrium in DNA duplexes in solution. QM data for **1**, as previously published,^{11,28} are represented by a 2D ϵ versus ξ potential energy surface based on structure optimizations at the MP2/6-31(+)-G(d) level with ϵ and ξ being sampled in increments of 15° , while restraining the overall conformation to that of the canonical B DNA form. Specifically, one of the sugar moiety dihedral angles, ν_4 , and the backbone torsion, α , were restrained.¹⁴ Analogous MM calculations were carried out using the CHARMM program⁴³ and the Drude polarizable force field.²⁸ The model was minimized first using the adopted-basis Newton–Raphson (ABNR) algorithm before performing the 2D dihedral potential energy scan. Dihedral harmonic restraints of 10^5 kcal/mol/rad² were used with minimizations performed to a gradient of 10^{-6} kcal/mol/Å².

Alteration of partial charges in **1** prompted further investigation of hydration properties of a smaller model compound, **2**, dimethylphosphate (DMP), which is representa-

tive of the phosphate group in the DNA backbone. The alteration of the electrostatic parameters to correct the behavior of **1** leads to changes in the hydration of **2**. Therefore, modifications of LJ parameters for selected atoms were necessary. The optimization was performed by targeting, simultaneously, reproduction of the hydration free energy of DMP and the gas-phase heterodimeric interactions with water. Calculation of the hydration free energy is elaborated in the next section. QM calculations on DMP–water monohydrates, using Gaussian 09,⁴⁴ were performed with DMP fixed in the geometry obtained from gas phase optimization at the MP2/6-31+G* level of theory and the water molecule fixed in the geometry of the SWM4-NDP model.⁴⁵ Three minimum energy conformations of DMP were considered: *gg*, *tt*, and *gt*, where *g* and *t* stand for “gauche” and “trans”, respectively. The relative orientations of the DMP and water molecules were also fixed, as shown in Figure 2, with only the interaction distances

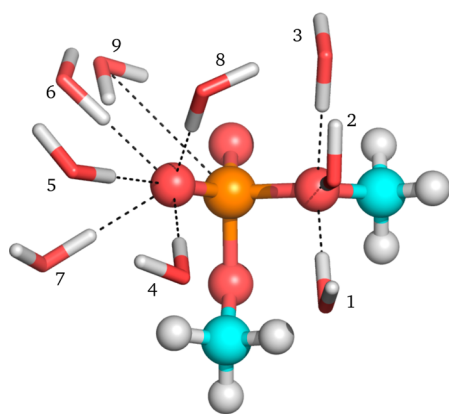


Figure 2. Schematic representation of the interaction orientation between DMP anion and water molecules. While analysis was performed on the basis of water–DMP monohydrates, in this figure all water orientations are shown simultaneously. The numbering of the water orientations corresponds to that used in Table 2 and Tables S1 and S2 of the Supporting Information.

optimized at the MP2/6-31+G* level of theory. Interaction energies were then determined via single point calculations at the RI-MP2/cc-pVQZ level of theory, performed on the MP2/6-31+G* optimized geometries. The single point calculations were performed using the program Q-Chem,^{46,47} with counterpoise correction⁴⁸ included to account for basis set superposition error (BSSE).⁴⁹ MM polarizable calculations were performed with CHARMM⁴³ using the SWM4-NDP water model⁴⁵ and restraints and orientations identical to those used in the QM calculations.

Another important aspect of DNA modeling is the ability of the FF to properly treat interactions between DNA and mobile ions. Those were tested based on two approaches. The first approach deals with the detailed analysis of the Na⁺ and Cl[−] distributions around a DNA duplex, ignoring sequence specific effects. The second approach deals with gas-phase calculations on model systems involving interactions of Na⁺ with individual nucleic acid bases, as well as with DMP. In particular, Na⁺ was positioned at distances ranging from 1 to 10 Å away from the chemically relevant atoms at various orientations relative to the base. Specifically, nitrogen atoms participating in Watson–Crick base-pairing – N1 of adenine, N3 of cytosine, N1 of guanine and N3 of thymine – were chosen for probing interactions with Na⁺. In the case of guanine and thymine,

whose N atoms are protonated, out-of-plane interactions with Na⁺ were studied. Additionally, calculations on the Ade–Ade/Na⁺ trimer model system were performed, with Na⁺ placed at distances [1–10] Å from the bases in the in-plane orientation and bases fixed in the relative orientation corresponding to B form DNA. Analogous calculations were carried out for DMP[−]/Na⁺ dimers, probing interactions among the phosphate group and Na⁺, with the latter positioned at various orientations and distances from the anionic O1P/O2P and ester O3/O4 oxygen atoms. Visual representations of the model systems are provided in the Results and Discussion section along with the corresponding QM and MM energy profiles. The distance increment in positioning Na⁺ from the atom of interest was 0.1 Å in the range of [1–4] Å, and 1 Å at larger separations. QM potential energy profiles for all complexes were based on gas phase structures optimized at the MP2/6-31+G(2df,2pd) level of theory using Gaussian 09⁴⁴ with counterpoise corrections⁴⁸ applied to account for BSSE.⁴⁹ Equivalent single point calculations were then performed using the Drude polarizable force field for nucleic acids.

Additional analysis of the nucleic acid bases was performed on their geometries and vibrational spectra, motivated by preliminary simulations of the Drew Dickerson (DD or EcoRI) dodecamer,^{50–52} indicating that the conformational behavior of the individual nucleic acid bases may contribute to DNA instability. This prompted a partial reoptimization of the internal parameters for the nucleic acid bases. As discussed below, this mainly involved manual adjustment of the base valence angles leading to a better representation of the vibrational frequencies and improvements in the geometries. Target data for bonds and angles involving hydrogen atoms were obtained from the previously published QM geometries optimized at the MP2/6-31+G* level of theory,⁵³ while all other bond length and angle parameters were optimized to reproduce target crystallographic geometries obtained from the Cambridge Structural Database survey.⁵⁴ CHARMM vibrational spectra were calculated using the MOLVIB utility⁵⁵ of the CHARMM program,⁴³ with the target infrared spectra calculated *ab initio* at the MP2/6-31+G* level of theory. Internal coordinate assignment was performed according to the method of Pulay et al.,⁵⁶ and a scale factor of 0.9434 was applied to QM calculated normal modes to account for limitations in the level of theory used.⁵⁷

Calculation of Hydration Free Energies. Hydration free energies, ΔG_{hyd} , were calculated via the free energy perturbation (FEP) method using the protocol of Deng and Roux.⁵⁸ In this approach, the LJ potential is separated into purely repulsive and attractive components using the Weeks–Chandler–Anderson (WCA) decomposition scheme,⁵⁹ as has been previously described.⁶⁰ Simulation systems involved a periodic box containing a single DMP molecule solvated by 250 SWM4-NDP⁴⁵ water molecules. Simulations were performed using CHARMM⁴³ and the velocity Verlet integrator,⁶¹ which includes treatment of Drude particles via an extended Lagrangian formalism.³⁷ The SHAKE algorithm⁶² was employed to constrain covalent bonds to hydrogen, and an integration time step of 1 fs was used. Simulations were performed in the NPT ensemble at 298.15 K and 1 atm, using a Nosé–Hoover thermostat^{63,64} and a modified Andersen–Hoover barostat.⁶⁵ Electrostatic interactions were treated using particle mesh Ewald (PME) summation⁶⁶ with a coupling parameter of 0.34 and a sixth-order spline for mesh interpolation. The system was equilibrated for 100 ps at each

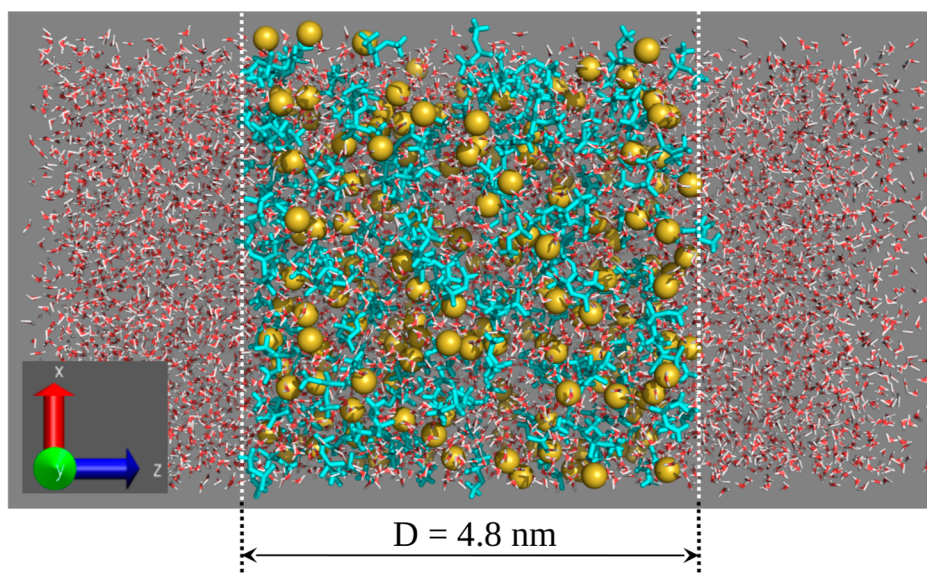


Figure 3. Snapshot from an MD simulation of the ~ 3 M aqueous solution of the DMP^- and Na^+ ions. Sodium ions are yellow, DMP ions are cyan, and water is red/white.

value of the thermodynamic coupling parameter, λ , after which the system was simulated for 200 ps of production MD. For the calculation of dispersive and electrostatic contributions to ΔG_{hyd} , λ was assigned values ranging from 0.0 to 1.0 in equally spaced increments of 0.1. For the repulsive term, extra windows were required in the “low λ ” region, such that λ took additional values of 0.05, 0.15, 0.25, and 0.35. The dispersive and electrostatic contributions to ΔG_{hyd} were calculated using thermodynamic integration (TI),⁶⁷ with the repulsive contribution calculated using a soft-core scheme,⁵⁸ and unbiased using the weighted histogram analysis method (WHAM).⁶⁸ Because these simulations were performed under periodic boundary conditions using the particle sum (P-sum) convention,⁶⁹ the interface potential between the bulk liquid and the external dielectric medium is zero. In reality, solvation of a charged species across a liquid/vacuum interface also depends on a nonzero and water-model-dependent interface potential ($\varphi_{\text{I}\rightarrow\text{v}}$). Therefore, an additional contribution to the free energy, equal to $q_{\text{ion}}\varphi_{\text{I}\rightarrow\text{v}}$, must be accounted for when performing calculations using the P-sum convention under periodic boundary conditions.⁷⁰ When considering a system consisting of a monovalent anion solvated with SWM4-NDP water, this contribution is 12.5 kcal/mol. The gas phase contribution to the hydration free energy was calculated using an identical procedure, but considered only an isolated DMP molecule in the gas phase, rather than a single molecule solvated in a water box.

Osmotic Pressure Calculations for the Bulk Electrolyte Solutions. System setup, MD simulation protocol, and the calculation of osmotic pressures for the bulk electrolyte solutions closely follow the method developed by Roux and co-workers.^{71,72} Three aqueous electrolyte solutions of (1) Na^+ and Cl^- ions, (2) Na^+ and DMP^- ions, and (3) Na^+ and Acetate^- ions were simulated at two *molal* concentrations each, approximately corresponding to solutions having molarities of 1 M and 3M, respectively (see below). Each simulation system contains three compartments separated by two virtual semipermeable membranes aligned with the xy -plane. The electrolyte compartment centered on the absolute origin of coordinates is enveloped by compartments containing pure

water. Water can freely pass through the membranes separating electrolyte and water solutions, while the ions are subject to the force of the half-harmonic planar potential on each side of the ionic compartment whose action mimics the ideal behavior of a semipermeable membrane inducing osmosis. A representative snapshot from MD simulation of the Na^+/DMP^- electrolyte solution at ~ 3 M concentration is shown in Figure 3. The instantaneous osmotic pressure, $\pi = F_{\text{memb}}/A$, equals the instantaneous total force on the membranes,

$$F_{\text{memb}} = - \sum_i k(z_i \pm D/2), \quad |z_i| > D/2 \quad (1)$$

divided by the total area of the membranes, A . In this expression, k is the force constant of the membrane potential, 10 kcal/(mol·Å), $D = 48$ Å is the linear size of the cubic ionic compartment, and z_i is the z -coordinate of the i ion. Values of the instantaneous force were recorded every 0.5 ps, and the total MD simulation run was for 70 ns for each simulated electrolyte system. The convergence of MD simulations was verified by calculating osmotic pressures from the two halves of the MD trajectory. Standard deviations around the averaged values were obtained based on seven 10 ns blocks of the total MD trajectories.

Given the values of experimentally measured osmotic pressure coefficient, $\varphi = \pi/\pi_{\text{id}}$, the experimental value of the osmotic pressure at molal concentration, m , is determined as follows:

$$\pi(m) = \varphi \cdot \pi_{\text{id}}(m) \quad (2)$$

$$\pi_{\text{id}}(m) = \frac{RT}{V_m} \nu \cdot m \quad (3)$$

where R is the gas constant, T is temperature in Kelvin, V_m is the molar volume of pure water at 1 bar and room temperature, and ν is the total number of cations and anions per formula unit (in our case of 1/1 electrolyte solutions $\nu = 2$).⁷³ It is important to note that, following the convention used in experimental measurements, the osmotic pressure of an ideal solution in the above expression is proportional to the *molal* concentration of the electrolyte, m , which differs from the traditionally used

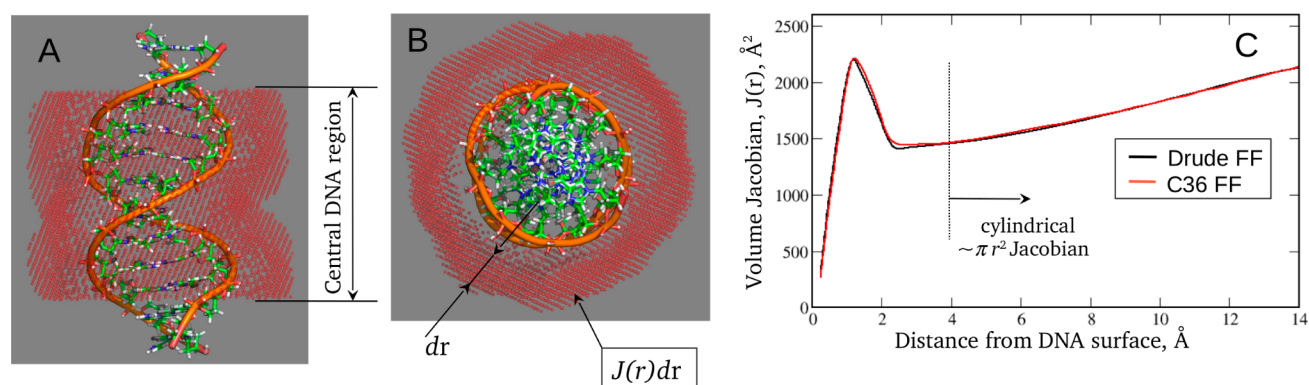


Figure 4. Side (A) and top (B) views of the DNA and equidistant shell (in red) from its surface. The shell's thickness is dr , and its volume is $J(r) dr$. $J(r)$ is the volume Jacobian whose numerical value as a function of the distance from DNA surface is shown in C. It is seen that the function increases monotonically only after ~ 4 Å from the DNA, from which it can be approximated by the cylindrical Jacobian often used for calculation of the ionic distributions around DNA. To avoid end effects, ionic distributions were analyzed in the central region of the DNA, as shown in A.

molar concentration M . Therefore, to enable direct comparison with the experiment, the actual molal concentrations need to be computed from MD simulations based on the values of mean densities of the cations, anions, and pure water in the ionic compartment as follows:

$$m = \frac{\bar{\rho}_{\text{salt}}/M_{\text{salt}}}{\bar{\rho}_{\text{wat}}}, \quad \bar{\rho}_{\text{salt}} = \bar{\rho}_{\text{cation}} + \bar{\rho}_{\text{anion}} \quad (4)$$

where overbars indicate ensemble averages, and M_{salt} is the molar concentration of the salt. The resulting molal concentration is then used to obtain the osmotic pressure of an ideal solution via eqs 2 and 3. An identical approach for computing molal concentrations and osmotic pressures is used in ref 74.

MD simulation protocol for studying electrolyte solutions is analogous to that described below for the DNA duplexes. A few differences are represented by the use of 'useFlexibleCel' and 'useConstantArea' options of NAMD.⁷⁵ The first option allows for the three dimensions of the periodic cell to fluctuate independently, while the second one keeps the ratio of the unit cell constant in the xy -plane while allowing fluctuations along the z axis in the course of MD simulation. In addition, the "TclForce" script of NAMD was utilized to impose semi-harmonic restraints on ionic species during MD simulations, as elaborated above. Simulations were carried out with the CHARMM Drude polarizable and the additive C36 force fields. Each simulated system included ~ 45 000 particles, including atoms, lone pairs, and Drude particles, in the case of the polarizable model and ~ 26 000 atoms in the additive model.

MD Simulation Protocol for DNA Duplexes in an Aqueous Salt Environment. Simulated systems included the EcoRI dodecamer^{50–52} and Junfos^{24,76} 16-mer whose structural behavior was extensively investigated in our original study.²⁸ Additionally, we studied a shorter 10-base-pair DNA oligomer (1DCV).⁷⁷ All DNA molecules were solvated with the polarizable SWM4-NDP⁴⁵ water model in cubic boxes with water molecules extending at least ~ 15 Å from the DNA surface, with addition of neutralizing Na^+ ions and an extra ~ 120 mM of NaCl salt.^{72,78} Initial configurations were generated by simulating the analogous additive CHARMM36 (C36) systems for several nanoseconds according to the protocol described elsewhere,¹¹ taking the last snapshot from these runs as inputs for subsequent Drude simulations.

CHARMM was used for generating the polarizable Drude system, self-consistent relaxation of the Drude particle positions, and short equilibration of the whole system, as elaborated elsewhere.²⁸ For the production MD simulation NAMD⁷⁵ (version 2.9) was used whose self-consistent treatment of the Drude polarizable model is based on a dual thermostating scheme and Langevin dynamics.³⁸ In particular, real atoms and polarizable degrees of freedom (Drude particles) were coupled to the thermostats at 300 and 1 K, respectively. Electrostatic interactions were treated using the PME summation⁶⁶ with a coupling parameter of 0.34 and a sixth-order spline for mesh interpolation. Nonbonded pair lists were maintained out to 16 Å, and a real space cutoff of 12 Å was used for the electrostatic and Lennard-Jones (LJ) terms. All covalent bonds involving hydrogen as well as the intramolecular geometries of water molecules were constrained using the SETTLE algorithm.⁷⁹ The "HARDWALL" feature enabled the use of a 1 fs time step in MD simulations.²⁰ As previously described, this feature is associated with a "hard wall" reflective term in the potential energy function that has been added to resolve the polarization catastrophe problem in Drude MD simulations.^{40,41} This term was invoked only when Drude particles moved >0.2 Å away from their parent nuclei during MD simulations.

Studying the Ionic Atmosphere Around DNA Duplex Based on Ion/DNA Radial Distribution Functions. To analyze in detail the Na^+ and Cl^- distributions around a DNA duplex, we calculated their radial distribution functions (RDFs)⁸⁰ from MD simulations following the approach developed by Savelyev et al.⁸¹ Each RDF was based on first defining the ion–DNA separation as the closest distance between the DNA molecule and the particular ion. These distances were used to construct DNA–ion distance histograms from each snapshot of the MD simulation. To obtain the RDFs, the histograms were normalized by the volume Jacobian, playing the role of the configurational entropy of ions structured around DNA. The number of neighbors within a distance r from a given object is

$$n(r) = \rho \int_0^r g(r)J(r) dr \quad (5)$$

where ρ is the bulk particle concentration, $g(r)$ is the RDF (pair correlation function), and $J(r)$ is the volume Jacobian. The latter is defined as the volume of a shell, equidistant from the

DNA surface (see Figure 4a,b), which was numerically calculated as a function of a distance from the DNA surface (Figure 4b). To minimize end effects, only the inner region of the DNA duplex (eight internal base pairs in the case of the 12-basis-pair *EcoRI* dodecamer) was used for computing the RDFs. This was achieved by “clipping” the periodic boundary cell with two parallel planes that pass through the centers of the upper four and the lower four DNA base pairs and perpendicular to the line connecting these centers. The above procedure was repeated for each snapshot along the MD trajectory.

It is worth mentioning that the present method of computing $J(r)$ and RDF differs from the standard procedures relying on either spherical or cylindrical symmetries.⁸² For example, although DNA is on average cylindrically symmetric and the use of the cylindrical Jacobian is reasonable, such an approach leads to an overestimation of the counterion condensation at small distances from the DNA surface. This is important, in particular, when calculating the absolute number of ions contributing to the various RDF peaks and subsequently determining the fraction of the DNA charge neutralized by condensed counterions—a critical test for evaluating the quality of the force field used to model ion/DNA interactions (see below). The numerical volume Jacobian, taking into account the complex and irregular shape of the DNA segment, is characterized by an unusual *nonmonotonic* behavior in the vicinity of the DNA oligomer (Figure 4c). It is seen that only at distances more than ~ 5 Å from the DNA surface can the monotonically increasing Jacobian be approximated by the cylindrical one.

Ion–DNA distance histograms, computed over snapshots saved every 4 ps from the 100 ns simulations, were normalized by the corresponding numerical Jacobian. Three-dimensional grids with lattice spacings of 0.5 Å were used to calculate both the ion–DNA distance histograms and the volume Jacobian. The biochemical algorithm library (BALL)⁸³ was used to implement the computational analysis subroutines. Additionally, BALL was used for writing code for reimaging/recentering of MD trajectories in cases when DNA strands (treated as independent segments) appeared separated in the course of simulation because of the periodic boundary conditions. This postprocessing was necessary for subsequent calculations of the volume Jacobian and RDFs, and also other overall structural DNA properties, such as root-mean-square (RMS) deviations and Watson–Crick base-pairing.

RESULTS AND DISCUSSION

Earlier stages of the Drude force field optimization for DNA were aimed at resolving overall duplex instability, while later published efforts²⁸ addressed finer details of the DNA structural behavior after the duplex stability was achieved. The resulting Drude DNA parameters were subjected to an extensive validation focusing on the details of the key backbone, sugar and glycosidic dihedral angles, whose correlated motions influence two critical aspects of the DNA structural behavior in solution—the proper treatment of the BI/BII transitions and the equilibrium between A and B forms DNA.²⁸ In the present work we present a description of earlier parametrization of selected bonded parameters in the bases and nonbonded parameters regulating electrostatic and van der Waals interactions between mobile ions, DNA atoms and water molecules. These interactions are strongly interdependent and must be properly balanced, with the balance extending to the rest of the FF. For example, mitigating electrostatic interactions

between selected atoms along the DNA backbone by varying their default partial charges proved necessary for achieving experimentally and computationally predicted BI-to-BII transition frequencies^{8,84} and the dynamical puckering transitions of the sugars,^{85–87} as well as better consistency with gas-phase QM data.¹¹ Such alterations, however, disrupted the existing balance of various energetic contributions having an effect on, e.g., DNA hydration properties, requiring adjustments of LJ parameters that impact specific interactions with water molecules. Those adjustments, in turn, prompted calculations of the hydration free energies and osmotic pressure coefficients for a number of biologically relevant ions in aqueous solution and comparison of MD simulation results with available experimental data.

Another complication that was encountered involved the interactions between the DNA and mobile ions. While the individual ions^{72,78} and base nonbond parameters⁵³ were carefully optimized, that optimization did not involve direct interactions of the ions with the bases. In the FF, these “off-diagonal” interactions are treated with the electrostatic parameters for the individual moieties along with LJ parameters obtained from combining rules, the Lorentz–Berthelot combining rules⁸⁸ in the present case. With the additive force field, this approximation generally worked well, although limitations in the approach have been discussed.^{51,60,89,90} However, with the Drude force field, the direct use of the electrostatic and combining rule LJ parameters has been shown to be more problematic, leading to optimization of both terms in specific cases via the use of through-space Thole interactions (NBTHOLE)^{91,92} and pair-specific LJ parameters (NBFIX),⁴³ respectively. For example, NBTHOLE terms have been used to fine-tune the interactions of counterion pairs based on osmotic pressures⁷² and the pair-specific LJ parameters have been used to improve free energies of hydration.⁶⁰ With DNA, an imbalance between the base and counterions leads to the interactions between them being significantly overestimated, leading to destabilization of the DNA. To correct this, gas phase QM calculations of interactions of the nucleic acid bases and Na^+ ion were used to initially optimize the off-diagonal parameters, followed by MD simulations of the fully solvated DNA duplex in a salt buffer with analysis focused on reproduction of predictions from counterion condensation theory.⁹³ Following several iterations, pair-specific LJ parameters for the ions and nucleic acid base hydrogen bond acceptor atoms were obtained that achieve satisfactory agreement with both the QM and counterion condensation theory target data.

Tuning the Interactions Between Nucleic Acid Bases and Na^+ Counterions. Initial nonbond parameters for the DNA FF involved the published base⁵³ and atomic ion⁷² parameters; however, their application lead to instabilities in simulations of duplex DNA. As shown in Figure S1 of the Supporting Information, simulations of the *EcoRI* dodecamer resulted in RMS deviations from canonical B conformation which increased irreversibly after ~ 10 ns and ~ 15 ns of MD simulations, respectively, when only neutralizing Na^+ ions and Na^+ ions together with an extra ~ 120 mM NaCl salt were included in the simulation system. The direct impact of the salt concentration on the longevity of the DNA’s stability domain suggested the potential cause for such behavior, which turned out to be associated with the anomalously favorable interactions among Na^+ ions and the DNA oligomer.

Identification of the problem proceeded as follows. We first computed the DNA/ Na^+ RDFs and compared the results from

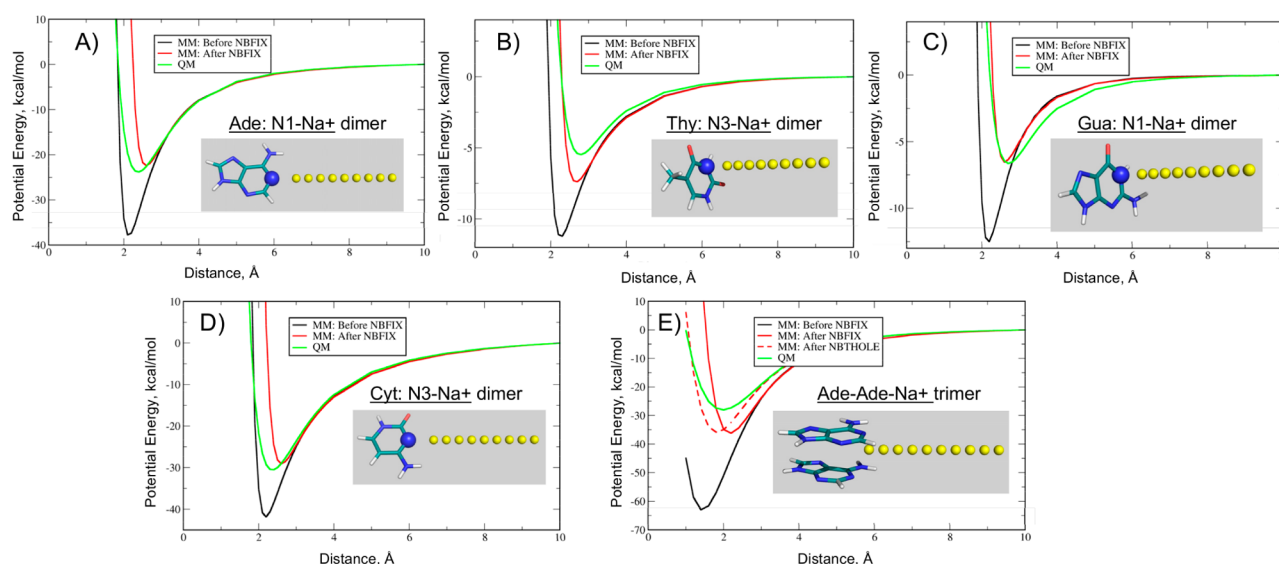


Figure 5. QM and empirical Drude potential energy surfaces as functions of the distance between Na^+ ion and selected atoms of the nucleic acid bases (A–D), or of A–A stacked bases (E).

polarizable Drude and analogous additive C36 MD simulations (see Figure S2, Supporting Information). Such comparison revealed that Na^+ condensed on DNA to a much greater extent in the polarizable model, with that increase largely involving direct contact of Na^+ with DNA. Subsequent RDF analysis of specific sites on the DNA (Figure S3, Supporting Information) identified stronger binding of Na^+ in the central AATT tract of the *EcoRI* oligomer. Analysis was then performed on the frequency of Na^+ binding to these four internal residues (Figure S4) revealing that while, on average, Na^+ ions interact more frequently with the phosphate groups in the additive model (Figure S4a), there is an enhanced binding of the sodium to the interior of the DNA in the Drude model, namely, to the N1/N6 atoms of the adenine and N3/O4 atoms of the thymine (Figure S4b, Supporting Information). Because these atoms participate in the Thy–Ade Watson–Crick base-pairing (N3–H \cdots N1 and O4 \cdots H–N6 hydrogen bonds), it became clear that the cause for the DNA structural disruption was Na^+ competing for those interactions.

The above analysis prompted the calculation of QM and MM interactions between the individual nucleic acid bases and Na^+ . Additionally, as the AA stacked pair was indicated to be problematic (Figure S4c, Supporting Information), calculations on the Ade–Ade/ Na^+ trimer complex were performed, with the adenine bases fixed in the B DNA conformation. The QM and MM potential energy scans are shown in Figure 5. As is evident, the default Drude parameters largely overestimate the attraction between Na^+ and selected atoms of the bases at short distances. The MM and QM energies differ by ~ 12 – 15 kcal/mol for Ade and Cyt bases (Figure 5a, d), and by ~ 8 kcal/mol for Thy and Gua bases (Figure 5b, c). The discrepancy is even bigger in the case of the Ade–Ade/ Na^+ trimer for which MM and QM differ by nearly 40 kcal/mol (Figure 5e). Clearly, these results point to the default Drude parameters leading to overestimation of the base– Na^+ interactions ultimately destabilizing duplex DNA.

To overcome the overestimation of the base– Na^+ interactions, we considered the use of both the NBTHOLE⁷⁴ and pair-specific LJ (NBFIX option in CHARMM)⁴³ approaches. While both approaches could alleviate the problem, as shown in Figure 5e, use of pair-specific LJ terms was considered to be

more practical as a number of other pair-specific LJ terms are implemented in the Drude FF. The NBFIX approach is based on assigning small repulsive cores (i.e., LJ radii) to the particles of interest to mitigate Coulomb interaction among them. To resolve the problem, we modified only the values of R_{min} and did not alter the LJ depth (ϵ) from that obtained from the combining rules. Such modifications have been made for LJ interactions between Na^+ and the following atoms: (1) N1, N3, N6 of adenine; (2) N1, N2, N3 of guanine; (3) N1, N3, N6 of thymine; and (4) N4, N3 of cytosine. In practice, because of the repeating atom types in all four bases, only three unique pair-specific LJ terms (e.g., NBFIXes) were introduced. As seen from Figure 5, this improved the consistency with the QM data. As anticipated, these adjustments led to improved stability of duplex DNA in solution, as judged from RMS differences versus B-form DNA for the *EcoRI* dodecamer (see Figure S5).

Adjustment of the Internal Parameters for Nucleic Acid Bases. During preliminary duplex DNA MD simulations, significant out-of-plane distortions of individual bases lead to DNA instability on times exceeding ~ 30 ns (Figure S5, Supporting Information). Visual examination of the base geometries during the course of the simulations revealed that excessive out-of-plane distortions were evident for non-hydrogen atoms as well as for the amino groups. The combined effect of these distortions was the formation of hydrogen bond-like interactions between the stacked bases in the duplex DNA, in particular, the amino groups with the carbonyl oxygen atoms. Interestingly, similar deformations were recently observed from gas-phase QM calculations on model systems composed of the two stacked bases, whose overall geometries were fixed in either B or A DNA conformations based on selected sets of internal coordinates, but allowed to relax otherwise.⁹⁴ While the effect of these out-of-plane distortions in duplex DNA in solution is not clear,⁹⁵ such base deformations clearly need to be limited to achieve a stable DNA structure.

To understand the cause of the distortions, analysis of the amino groups in the adenine and cytosine bases demonstrated that the out-of-plane displacement of the NH_2 group hydrogen atoms was associated with the sum of the valence angles around nitrogen atoms being smaller than 360° (Figure S6, Supporting

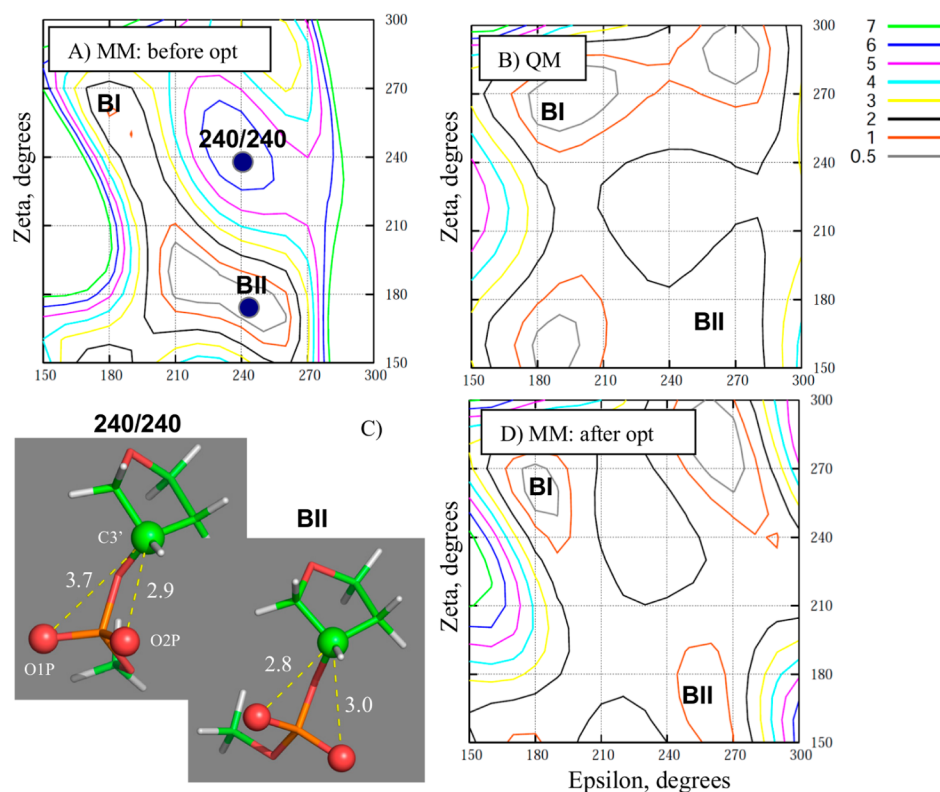


Figure 6. QM and empirical Drude 2D potential energy surfaces of ϵ versus ζ for model compound **1** (T3PS). Also depicted are structures of the lowest (BII) and highest (240/240) energy conformations (C) of landscape A generated by nonoptimized Drude parameters. Some of the 1D representative ϵ and ζ energy profiles are provided in the Supporting Information, Figure S7.

Information). A similar analysis was also performed for the base's 5- and 6-membered rings whose planarity is characterized by the sum of all inner valence angles being 540° and 720° , respectively (not shown). This indicated the valence angles required additional optimization.

Adjustments of the internal base parameters were performed targeting crystal survey geometries⁵ and QM geometries and vibrational spectra of individual bases with their geometries restrained to be planar. Based on the above observations, optimization focused on the equilibrium values for the valence angles, subjected to the following restraints: (1) the sum of the valence angles around the planar centers, such as amino groups, should be 360° ; and (2) the sum of the inner valence angles in the 5- and 6-membered rings should be 540° and 720° , respectively. Although these were just a few restraints, performing such adjustments posed a challenge because of the repeating atom types in chemically different bases, making it difficult for these conditions to be satisfied for all in-plane centers and/or rings. However, the final parameters significantly improved both the agreement of the Drude model with both the experimental geometries and QM frequencies of the bases, as shown in Tables S1–S8 of the Supporting Information. Following these corrections, the DNA duplex was observed to be stable on time scales of 100 ns based on MD simulations of the EcoR1 dodecamer (Figure S5), as well as simulations of other duplexes presented in our previous study.²⁸

Addressing Fine Details of the DNA Structural Behavior. Once extended simulations of DNA in solution were possible, more subtle conformational effects were addressed. These included transitions between the BI and BII substates of the canonical B form DNA and pucker transitions

of the sugar groups from the dominant south conformation to the occasionally visited north conformation in B form simulations.²⁸ Details of the optimization process follow.

Model compound **1** was designed to study energetics of the correlated motions among the ϵ and ζ backbone dihedral angles in the context of the BI/BII transitions in duplex DNA.^{11,28} QM energetic data on **1** for 2D energy scans of ϵ versus ζ torsions along with experimentally determined populations of the BI and BII states were used as target data for optimization of the ϵ and ζ torsions in the context of the initial Drude parameters. As the optimization proceeded, it was necessary to significantly sacrifice the quality of agreement with the QM data to obtain satisfactory sampling of the BI and BII states. The resulting Drude 2D ϵ versus ζ energy surface showed the model to significantly overstabilize the BII state of **1** as compared to the QM surface, as shown in Figure 6a,b, respectively. In addition to the overall shapes of the QM and MM landscapes differing dramatically, the low energy pathways connecting BI and BII states were in poor agreement. Moreover, while BI-to-BII transitions were occurring, no experimentally observed transitions in sugar pucker from the south to north conformation,⁸⁵ an important prerequisite for stabilizing A-form DNA under appropriate conditions,^{9,11,12,28,96} occurred. This is important as the A/B and BI/BII conformational equilibrium are strongly coupled, with the BII substate known to be strongly anticorrelated with sampling of the north sugar conformation characteristic of the A form DNA.^{3,11,28}

This involved energetic analysis of conformations of **1** having the smallest and the largest relative energies; those states are indicated by blue circles in Figure 6a with the approximate values of the ϵ/ζ angles being $240^\circ/240^\circ$ for the highest and $245^\circ/175^\circ$ (BII) for the lowest energetic states. Comparison of

individual energetic contributions (bonded, LJ and electrostatic) in these conformations revealed that their electrostatic contributions differ by ~ 8 kcal/mol in favor of the BII state, a contribution almost entirely responsible for the total energetic difference of ~ 7 kcal/mol. Examination of the geometries of these conformations suggested this difference to be attributed to the attraction between the anionic O1P/O2P atoms of the phosphate and the positively charged C3' carbon of the sugar ring. This distance is shorter by ~ 1 Å in the BII conformation, as shown in Figure 6c. Due to this favorable interaction, the C3'–O3'–P valence angle is significantly deformed, being $\sim 109^\circ$ instead of $\sim 120^\circ$ as observed in a survey of DNA crystal structures.⁵ Despite being a local effect, the polymeric nature of DNA would lead to this difference potentially causing an elongated DNA duplex.

To overcome this issue, we altered the default electrostatic parameters by making the partial charges for the O1P/O2P atoms less negative by 0.1e, and the C3' atom less positive by 0.2e. Following this change and additional optimization, it was observed that the agreement between the QM and MM data for **1** was significantly improved (Figure 6d) as was the agreement with experiment for the BI/BII equilibrium and puckering of the sugar moieties in duplex DNA. In addition, the MM model better reproduced experimental data for selected valence and dihedral angles in duplex DNA.²⁸ In the empirical ϵ versus ζ potential energy surface, the flattening of the landscape was achieved by mitigating the electrostatic interactions. To facilitate the comparison between QM and MM outcomes, we also included in the Supporting Information several representative 1D potential energy profiles as a function of ϵ or ζ obtained for model compound **1** being in either BI or BII conformation (Figure S7). Figure 7 shows time series of the ϵ and ζ torsions sampled in the *EcoRI* dodecamer using the final parameters, demonstrating frequent transitions in a bimodal fashion corresponding to BI/BII transitions. Additionally, time series for the sugar pseudorotation angles indicate pucker transitions from the north to south conformation occurring on the experimentally observed time scale of ~ 200 ps.⁸⁵ Figure S8 of Supporting Information also demonstrates BI/BII transitions and sugar pucker transitions to occur on the level of a single nucleotide. Thus, adjustments of the partial atomic charges lead to significant improvements in the conformational properties in the context of the gas phase model compound QM data and the condensed phase duplex DNA data.

Optimization of LJ Parameters for Interactions Between the DNA Phosphate Group and Water. Due to the changes in DNA backbone electrostatics it was necessary to reevaluate the hydration of the phosphodiester backbone. This was performed based on the ability of the force field to reproduce the hydration free energy as well as the QM minimum interactions energies with individual water molecules using DMP as the model compound (Figure 1).

The experimental hydration free energy of DMP is not available in the literature, but it can be estimated using the thermodynamic cycle shown in Figure 8. ΔG_1 is determined to be -325 ± 4 kcal/mol from gas phase acidity data,⁹⁷ and $\Delta G_3 = 1.75$ kcal/mol is obtained from the pK_a of protonated DMP (HDMP).⁹⁸ To estimate ΔG_2 , the hydration free energy of HDMP, we start with the known hydration free energy of a closely related compound, trimethylphosphate (TMP), which has a hydration free energy of -8.7 kcal/mol.⁹⁹ We then use QM to compute the relative hydration free energy between HDMP and TMP: a QM AMSOL^{100,101} calculation reveals it to

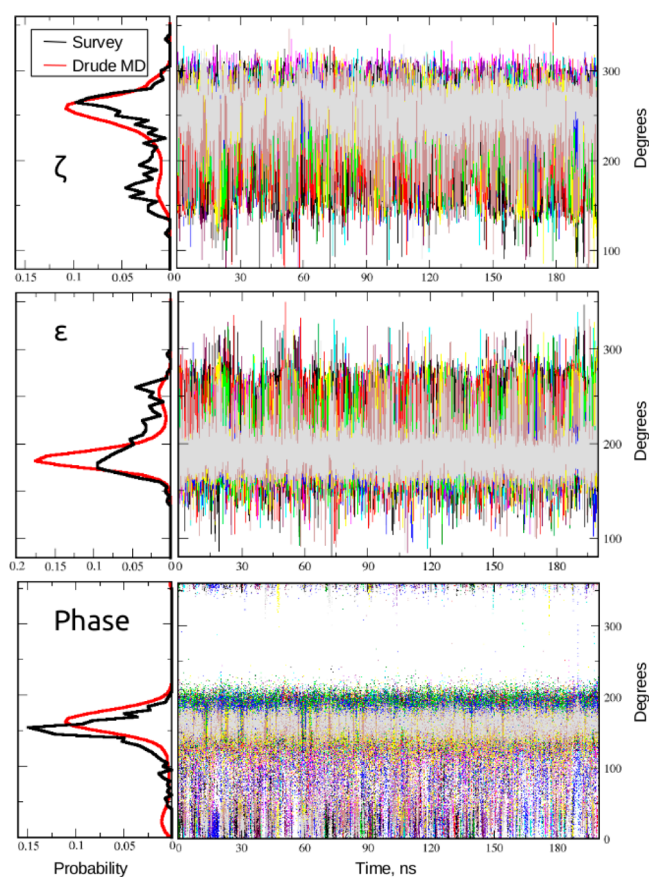


Figure 7. Time series and probability distribution functions from the MD simulation of the *EcoRI* dodecamer for ζ and ϵ backbone torsions, as well as for the sugar pucker obtained for the Drude parameters optimized based on model compound **2**. Transitions between BI ($\epsilon \approx 190^\circ$, $\zeta \approx 270^\circ$) and BII ($\epsilon \approx 270^\circ$, $\zeta \approx 180^\circ$) states, and north ($P \approx 15^\circ$) and south ($P \approx 175^\circ$) sugar conformations are evident. Based on these time series, both BI/BII transitions and sugar pucker transitions occur on time scales of ~ 200 ps, consistent with the QM data for model compound **2** and NMR experimental data, respectively.

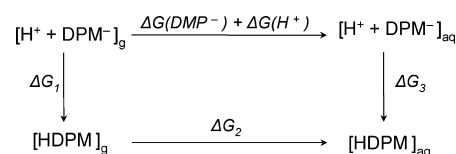


Figure 8. Thermodynamic cycle used to determine the experimental hydration free energy of DMP.

be -3.4 kcal/mol, giving an estimate of -12.1 kcal/mol for ΔG_2 . Alternatively, the relative hydration free energy between HDMP and TMP may be determined from the free energies of hydration of acetic acid, -6.7 kcal/mol, and methyl acetate, -3.32 kcal/mol,⁹⁹ resulting in the same value for $\Delta G_2 = -12.1$ kcal/mol. For the hydration free energy of the proton, $\Delta G(\text{H}^+)$, a value of -258.8 kcal/mol is used.⁷⁸ Putting all the contributions together gives an estimate for the experimental hydration free energy of DMP of -78.4 ± 5 kcal/mol.

Table 1 shows the hydration free energy for DMP computed at different stages of the parameter optimization. It is seen that the above-described change in charges for the O1P/O2P oxygen and C3' carbon atoms resulted in the hydration free energies being less favorable by ~ 14 kcal/mol, significantly

Table 1. Hydration Free Energies (kcal/mol) for DMP at Different Stages of the Optimization Procedure

experiment	default parameters	altered electrostatic parameters	altered electrostatic/LJ parameters
-78.4	-82.1	-64.3	-76.2

lower than the experimental estimate. To correct this, we systematically modified LJ parameters of selected atoms in DMP, varying both the well depth (ϵ) and R_{\min} terms. The final value of the hydration free energy obtained using the optimized Drude model parameters was -74 kcal/mol, in good agreement with the experimental estimate.

To further test the quality of the model, the gas phase QM and MM minimum interaction energies for DMP–water monohydrates in different orientation were determined. This analysis is important as it indicates whether the relative energies of hydrogen bonds with different atoms on DMP are correctly reproduced, information that the hydration free energy cannot provide. A total of nine different water positions were considered, as shown in Figure 2. The minimum interaction energies and distances with DMP in the *gg* conformation, corresponding to the geometry of the phosphate group encountered in DNA duplexes, are shown in Table 2. Minimum interaction energies for the other two minimum energy conformations of DMP, *gt* and *tt*, are provided in Tables S9 and S10, respectively, of the Supporting Information. It is clear from the data that the gas phase interactions with water are uniformly well reproduced. In general, the results indicate that an appropriate balance between the electrostatic and LJ parameters has been achieved.

Optimization of the Na⁺–Phosphate Interactions Targeting Osmotic Pressure Calculations. As evidenced above for the base–Na⁺ interactions, the proper balance of the nonbond portion of the force field is essential. To investigate this for the phosphodiester backbone and the NaCl salt, osmotic pressure coefficients were calculated for selected biologically relevant electrolyte solutions. Osmotic pressure coefficient is a measure of the solution “nonideality”, describing how much the behavior of the real electrolyte solution deviates from that of an ideal solution at a particular concentration.

Thermodynamically, osmotic pressure is related to the second osmotic coefficient in the osmotic virial series for a nonideal solution, with that coefficient, in turn, related to the effective interaction potential (potential of mean force) between a pair of ions.¹⁰² Thus, reproduction of osmotic pressures is an important requirement directly related to the FF’s ability to properly model interionic interactions and ensure those interactions are properly balanced with the ion–water interactions.

Systems studied included aqueous solutions of Na⁺ and Cl[−] ions, Na⁺ and DMP[−] ions, and Na⁺ and Acetate[−] ions, each at low and high molal concentrations (~ 1 M and 3M). Experimental values for the osmotic pressures of Na⁺/Cl[−] electrolytes are available for a wide range of concentrations, [0–6 M].^{12,103} For Na⁺/DMP[−] solution, experimental osmotic pressures are known only at low salt concentrations (≤ 1 M).¹⁰⁴ Therefore, an additional electrolyte system composed of Na⁺ and Acetate[−], due to the acetate oxygen atoms being chemically similar to the anionic oxygens of DMP[−], was simulated at ~ 1 M and 3M, concentrations for which experimental osmotic pressure coefficients are available. Based on the osmotic pressure calculations, optimization of selected pair-specific LJ parameters between Na⁺ and DMP was performed.

Table 3 shows the osmotic pressures produced by the final Drude parameters along with experimental values and results from the additive C36 FF. Table 4 contains the actual *molal* concentrations in the ionic slab from the MD simulations, required for direct determination with the experimental data. Comparison of the results in Table 3 shows the level of agreement to be quite good. The Drude model is typically in improved agreement with experiment over C36, although the additive model is in better agreement for the Na⁺/Acetate[−] solution. Given the overall good agreement between the Drude and experimental data, the current parametrization provides a plausible description of osmotic properties of biologically relevant ionic solutions.

While experimental data for osmotic pressure of the Na⁺/DMP[−] system is not available at high concentrations, the value obtained from a MD simulation study of this system at ~ 3 M is presented in Table 3. The value was determined in a study

Table 2. QM and MM Interactions between the *gg* Conformation of DMP and a Single Water Molecule in Different Orientations^a

dimer conformation	distance ^b			energy		
	QM	MM	difference: MM–QM	QM	MM	difference: MM–QM
1 o11_lp2				-2.65	-2.20	0.44
2 o11_bis	1.87	1.61	-0.26	-9.35	-9.73	-0.38
3 o11_lp1	1.92	1.65	-0.27	-8.79	-8.43	0.35
4 o13_lp2	1.84	1.73	-0.11	-11.50	-11.37	0.12
5 o13_180	1.81	1.70	-0.11	-12.74	-13.14	-0.40
6 o13_lp1	1.86	1.70	-0.16	-13.63	-13.59	0.03
7 o13_lp_d270				-11.94	-11.74	0.19
8 o13_lp_d90				-13.23	-13.24	-0.01
9 p_hoh	3.34	3.20	-0.14	-14.53	-14.46	0.06
AVE			-0.17			0.04
RMSD			0.06			0.27

^aOrientations are as shown in Figure 2. All energies are in kcal/mol. Distances for selected conformations are in Å. In QM calculations, the structure of the DMP was first optimized at MP2/6-31+G* level of theory. Next, the distances between water molecule and DMP were optimized at MP2/6-31+G* level of theory with the rest of the dimer fixed. QM interaction energies were calculated at the RI-MP2/cc-pVQZ level of theory using counterpoise correction. See Methods section for details. ^bDistances for the QM optimized geometries 2–6 are between the water hydrogen and o11/o13 oxygen atom of DMP. For geometry 9, the distance is between water oxygen and the DMP phosphorus.

Table 3. Experimental and MD Simulation Results for the Osmotic Pressures (Bars) for Different Electrolyte Solutions at ~ 1 M and ~ 3 M Concentrations

	experiment		Drude FF		additive C36 FF	
	~ 1 M	~ 3 M	~ 1 M	~ 3 M	~ 1 M	~ 3 M
Na–Cl	42.67	152.01	42.14 ± 3.52	153.36 ± 4.21	44.01 ± 4.24	149.58 ± 6.41
Na–DMP	47.41	226.8 [MD] ^a	48.66 ± 4.21	281.13 ± 21.66	45.39 ± 5.11	228.30 ± 18.61
Na–Acetate	49.70	195.20	47.42 ± 3.12	174.62 ± 8.15	44.23 ± 2.21	181.88 ± 9.10

^aThis value is from the computational study of Aksimentiev et al. utilizing AMBER force field.⁷⁴

Table 4. Actual Molal Concentrations for All Studied Electrolyte Solutions Computed from MD Simulations According to Eq 4

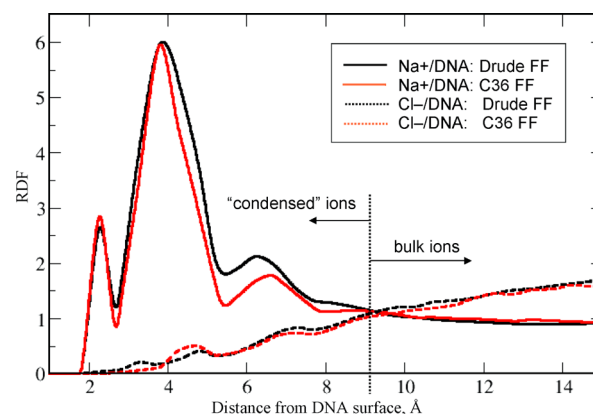
	Drude FF		additive FF	
	~ 1 M	~ 3 M	~ 1 M	~ 3 M
Na–Cl	0.85	3.04	0.79	2.82
Na–DMP	0.89	3.37	0.91	3.55
Na–Acetate	0.99	3.24	0.91	3.00

aimed at improved LJ parametrization of specific ion pairs in the AMBER force field.⁷⁴ As can be seen, the result from that study is in a good agreement with the additive C36 result, while the Drude simulation produced a noticeably higher value. Given the underestimation of the osmotic pressure for the Na⁺/Acetate[−] system, this result indicates that the experimental value may be significantly higher than that obtained from the additive FFs. This result serves as a prediction by the Drude FF.

Analysis of the Ionic Distribution Around DNA Duplex. Our final test of the Drude model is focused on details of the ionic structuring around the DNA duplex. Mobile ions impact DNA structure and dynamics through neutralization of its residual charge.⁷ As demonstrated above, excessive condensation of counterions to the DNA may lead to disruption of the DNA structure. Numerous prior experimental^{105–112} and theoretical^{81,82,113–116} studies focused on sequence specific ionic binding, such as determining whether and how long counterions reside in DNA grooves, or competitive ionic binding to specific DNA sites. While these aspects will be addressed in future studies, for validation we will focus on the overall ionic condensation.

Ionic distribution around DNA was characterized by computing RDFs, with averaging over the central region of the DNA duplexes.⁸¹ Such an approach allows for a direct comparison of the MD simulation results with the counterion condensation (CC) theory.⁹³ In their seminal work, Onsager and Manning¹¹⁷ demonstrated that CC must occur on linear rods that are uniformly charged above some critical charge density. The DNA molecule, being linear and carrying high charge per unit length, is well above this threshold.¹¹⁸ One can estimate from these simple arguments that 76% of the total DNA charge is expected to be neutralized by condensed counterions within a ~ 9 Å Manning radius.^{93,119}

RDFs for Na⁺ and Cl[−] ions shown in Figure 9 were computed from MD simulations of the *EcoRI* dodecamer in the ~ 120 mM NaCl aqueous solution using both the polarizable Drude and additive C36 models. Analogous data for another two DNA systems, a longer Junfos (16-base-pair) and a shorter 1DCV (10-base-pair) duplex are provided in the Supporting Information (Figure S9). In all graphs, Na⁺ RDFs are characterized by three prominent peaks, indicating the formation of ionic “shells” around polyanionic DNA. The onset of such an ionic structuring when approaching DNA from

**Figure 9.** Na⁺/DNA and Cl[−]/DNA RDFs from the Drude (black) and additive (red) MD simulations of the *EcoRI* dodecamer computed based on the closest approach, as elaborated in the Methods section.

the bulk is commonly associated with the Manning’s radius of CC theory. As seen from the plots, counterions are structured within ~ 9 Å from DNA surface in both FFs, coinciding with the CC predictions. This result is consistent with the definition of Beveridge and co-workers⁸² for the condensed counterions as lying within a 9 Å region from the DNA surface, where ions are structured with respect to DNA.

The computed ionic RDFs were used to calculate the fraction of the DNA residual charge neutralized by the ions. To this end, RDFs for Na⁺ and Cl[−] were integrated from 0 to 9 Å, according to eq 5, and the resulting number of co-ions and counterions were summed and divided by DNA residual charge. Drude and additive result for DD, Junfos, and 1DCV DNA systems are shown in Table 5. The agreement with the

Table 5. Comparison between Additive and Drude MD Simulation Results and Counterion Condensation Theory Based on Analysis of Ion/DNA RDFs

	% DNA Charge Neutralized within 9 Å		
	DD	1DCV	JunFos
Drude	77.0%	79.0%	75.4%
C36	69.0%	68.6%	67.2%
CC theory	76.0%	76.0%	76.0%

CC estimates is clearly better for the Drude FF simulations, which predict $\sim 77\%$, $\sim 79\%$, and $\sim 75\%$ of the DNA residual charge to be neutralized by ions within the Manning radius for *EcoRI*, 1DCV and Junfos, respectively. These results indicate that the Drude model faithfully captures basic features of the interplay between DNA duplexes and the surrounding ionic atmosphere, offering an improvement over the additive C36 FF, and further indicating the quality of the nonbonded parameters for DNA and mobile ions.

CONCLUSIONS

In the present study we addressed a number of critically important aspects of the development of the Drude polarizable force field for DNA. These issues included treatment of base- Na^+ interactions, proper treatment of the base intramolecular parameters, balancing DNA electrostatic and hydration effects, and how they impact the overall DNA stability and regulate subtler aspects of DNA conformational behavior. For example, when Drude parameters developed for the nucleic acid bases were applied to a DNA duplex in solution, the DNA structure became unstable after ~ 25 ns of MD simulation time because of significant out-of-plane deformations of individual bases. The problem was overcome by a partial reoptimization of base valence angle parameter based on reproduction of experimental geometries and QM geometries and vibrational spectra. Next, based on the QM model compound calculations, altering the default electrostatic charges of selected backbone DNA atoms was shown to be necessary to improve BI/BII conformational equilibrium and to better reproduce the dynamical sugar pucker transitions observed experimentally.⁸⁵ Such changes, however, required readjustment of selected LJ parameters dictating interactions between DNA and water, as well as DNA and mobile ions. These modifications were based on simultaneous targeting a range of experimental and QM data, such as the hydration free energy for DMP and the QM gas-phase interactions of DMP with water molecules, QM data on small model compounds representative of the bases and DMP-interacting with Na^+ ion, and experimental data on osmotic pressure measurements. A critical test of the FFs ability to capture basic features of the distribution of mobile ions around DNA involved comparing predictions from counterion condensation theory⁹³ with those based on computing ion/DNA RDFs obtained from MD simulations. Our results indicate that Drude simulations generate a physically plausible picture of ionic distribution around DNA, reproducing the theoretical estimates of the fraction of neutralized residual DNA charge by the condensed ions. Notably, the Drude results show an improvement over the C36 FF, indicating the improved physical description of the forces dictating the ionic solvation of DNA due to the explicit treatment of electronic polarizability.

In summary, we focused on balancing the electrostatics and hydration of the DNA backbone by optimizing selected Drude parameters for interactions between DNA, mobile ions and water molecules. Our initial efforts were aimed at stabilizing an overall DNA structure in solution, which included improvements in the base valence angle parameters, while later parameter modifications addressed finer details of the DNA conformational behavior, such as BI/BII transitions and the equilibrium between north and south states of the sugar moiety.^{11,28,96} Force field validation presented in this study, together with the earlier demonstrated ability of the force field to adequately model all aspects of the B form DNA, as well as to stabilize the A form in low water activity environment,²⁸ indicate that the current Drude polarizable model for DNA is suitable for computational studies of various DNA molecules under varying conditions on the 100 ns time scale. In addition, the availability of Drude polarizable parameters for proteins,⁴⁰ lipids,⁴¹ and carbohydrates^{53,92,120–128} will allow for simulation studies of heterogeneous biological systems.

ASSOCIATED CONTENT

Supporting Information

Details on analysis of the ionic binding to DNA (Figures S1–S4); RMSD plots characterizing DNA stability (Figure S5); time series for selected internal nucleic base parameters (Figure S6); 1D QM and MM potential energy profiles as functions of the ϵ and ζ dihedral angles for model compound **1** in BI and BII conformations (Figure S7); time series of the ϵ and ζ torsions as well as the sugar phase for selected nucleotide of the EcoR1 sequence (Figure S8); Na^+ /DNA and Cl^- /DNA RDFs from Drude MD simulations of the 1DCV and JunFos sequences (Figure S9); selected experimental and geometrical parameters (Tables S1–S4) and QM and MM vibrational frequencies (Tables S5–S8) for methylated nucleic acid bases; interaction energies for selected conformations of the DMP anion and water molecules (Tables S9, S10). This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Mailing Address: 20 Penn Street, Room 633, Baltimore, MD 21201. E-mail: alex@outerbanks.umaryland.edu. Phone: (410) 706-7442. Fax: (410) 706-5017.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The NIH (GM051501) is thanked for financial support. The University of Maryland Computer-Aided Drug Design Center and the XSEDE resources are gratefully acknowledged for their generous allocations of computer time. We also thank Dr. B. Roux and Dr. A. Aksimentiev for helpful discussions.

REFERENCES

- (1) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
- (2) Langley, D. R. Molecular dynamics simulations of environment and sequence dependent DNA conformation: The development of the BMS nucleic acid force field and comparison with experimental results. *J. Biomol. Struct. Dyn.* **1998**, *16*, 487–509.
- (3) Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M. L. The anatomy of A-, B-, and Z-DNA. *Science* **1982**, *216*, 475–85.
- (4) Hartmann, B.; Lavery, R. DNA structural forms. *Q. Rev. Biophys.* **1996**, *29*, 309–368.
- (5) Gelbin, A.; Schneider, B.; Clowney, L.; Hsieh, S. H.; Olson, W. K.; Berman, H. M. Geometric parameters in nucleic acids: Sugar and phosphate constituents. *J. Am. Chem. Soc.* **1996**, *118*, 519–529.
- (6) Dickerson, R. E. DNA structure from A to Z. *Method. Enzymol.* **1992**, *211*, 67–111.
- (7) Peters, J. P.; Maher, L. J., III. DNA curvature and flexibility in vitro and in vivo. *Q. Rev. Biophys.* **2010**, *43*, 23–63.
- (8) Banavali, N. K.; Roux, B. Free energy landscape of A-DNA to B-DNA conversion in aqueous solution. *J. Am. Chem. Soc.* **2005**, *127*, 6866–6876.
- (9) Cheatham, T. E.; Kollman, P. A. Insight into the stabilization of A-DNA by specific ion association: Spontaneous B-DNA to A-DNA transitions observed in molecular dynamics simulations of dACCCGCGGGT (2) in the presence of hexaamminecobalt(III). *Structure* **1997**, *5*, 1297–1311.
- (10) Pastor, N. The B- to A-DNA transition and the reorganization of solvent at the DNA surface. *Biophys. J.* **2005**, *88*, 3262–3275.
- (11) Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; MacKerell, A. D., Jr. Optimization of the CHARMM Additive Force

Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. *J. Chem. Theory Comput.* **2012**, *8*, 348–362.

(12) Cheatham, T. E.; Crowley, M. F.; Fox, T.; Kollman, P. A. A molecular level picture of the stabilization of A-DNA in mixed ethanol-water solutions. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 9626–9630.

(13) Cheatham, T. E., III; Case, D. A. Twenty-Five Years of Nucleic Acid Simulations. *Biopolymers* **2013**, *99*, 969–977.

(14) Foloppe, N.; MacKerell, A. D., Jr. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem.* **2000**, *21*, 86–104.

(15) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

(16) Cheatham, T. E., III; Cieplak, P.; Kollman, P. A. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845–861.

(17) Soares, T. A.; Hunenberger, P. H.; Kastenholz, M. A.; Kraeutler, V.; Lenz, T.; Lins, R.; Oostenbrink, C.; van Gunsteren, W. An improved nucleic acid parameter set for the GROMOS force field. *J. Comput. Chem.* **2005**, *26*, 725–737.

(18) MacKerell, A. D., Jr.; Nilsson, L., In *Computational Studies of DNA and RNA: Molecular Dynamics, Quantum Chemistry, Mesoscopic Modeling*, Spomer, J.; Lankas, F., Eds.; Springer: Dordrecht, The Netherlands, 2006; pp 73–94.

(19) Lavery, R.; Zakrzewska, K.; Beveridge, D.; Bishop, T. C.; Case, D. A.; Cheatham, T., III; Dixit, S.; Jayaram, B.; Lankas, F.; Loughton, C.; Maddocks, J. H.; Michon, A.; Osman, R.; Orozco, M.; Perez, A.; Singh, T.; Spackova, N.; Spomer, J. A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucleic Acids Res.* **2010**, *38*, 299–313.

(20) Foloppe, N.; Nilsson, L. Toward a Full Characterization of Nucleic Acid Components in Aqueous Solution: Simulations of Nucleosides. *J. Phys. Chem. B* **2005**, *109*, 9119–9131.

(21) Svozil, D.; Kalina, J.; Omelka, M.; Schneider, B. DNA conformations and their sequence preferences. *Nucl. Acid Res.* **2008**, *36*, 3690–3706.

(22) Pérez, A.; Marchán, I.; Svozil, D.; Spomer, J.; Cheatham, T. E., III; Loughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophys. J.* **2007**, *92*, 3817–3829.

(23) Zgarbova, M.; Otyepka, M.; Spomer, J.; Mladek, A.; Banas, P.; Cheatham, T. E., 3rd; Jurecka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.* **2011**, *7*, 2886–2902.

(24) Heddi, B.; Foloppe, N.; Bouchemal, N.; Hantz, E.; Hartmann, B. Quantification of DNA BI/BII backbone states in solution. Implications for DNA overall structure and recognition. *J. Am. Chem. Soc.* **2006**, *128*, 9170–7.

(25) Spomer, J.; Mladek, A.; Spackova, N.; Cang, X.; Cheatham, T. E., III; Grimme, S. Relative Stability of Different DNA Guanine Quadruplex Stem Topologies Derived Using Large-Scale Quantum-Chemical Computations. *J. Am. Chem. Soc.* **2013**, *135*, 9785–9796.

(26) Zgarbova, M.; Javier Luque, F.; Spomer, J.; Cheatham, T. E., III; Otyepka, M.; Jurecka, P. Toward Improved Description of DNA Backbone: Revisiting Epsilon and Zeta Torsion Force Field Parameters. *J. Chem. Theory Comput.* **2013**, *9*, 2339–2354.

(27) Krepl, M.; Zgarbova, M.; Stadlbauer, P.; Otyepka, M.; Banas, P.; Koca, J.; Cheatham, T.; Jurecka, P.; Spomer, J. Reference Simulations of Noncanonical Nucleic Acids with Different Variants of the AMBER Force Field: Quadruplex DNA, Quadruplex RNA, and Z-DNA. *J. Chem. Theory Comput.* **2012**, *8*, 2506–2520.

(28) Savelyev, A.; MacKerell, A. D., Jr. All-Atom Polarizable Force Field for DNA Based on the Classical Drude Oscillator Model. *J. Comput. Chem.* **2014**, *35*, 1219–1239.

(29) Drude, P.; Millikan, R. A.; Mann, R. C. *The Theory of Optics*; Longmans, Green, and Co.: New York, 1902; p 588.

(30) Jacucci, G.; McDonald, I. R.; Singer, K. Introduction of the shell model of ionic polarizability into molecular dynamics calculations. *Phys. Lett. A* **1974**, *50*, 141–143.

(31) Jacucci, G.; McDonald, I. R.; Rahman, A. Effects of polarization on equilibrium and dynamic properties of ionic systems. *Phys. Rev. A* **1976**, *13*, 1581.

(32) Lindan, P. J. D.; Gillan, M. J. Shell-model molecular dynamics simulation of superionic conduction in CaF₂. *J. Phys.: Condens. Matter* **1993**, *5*, 1019–1030.

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

(34) Sprik, M.; Klein, M. L. A Polarizable Model for Water Using Distributed Charge Sites. *J. Chem. Phys.* **1988**, *89*, 7556–7560.

(35) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Explicit reversible integrators for extended systems dynamics. *Mol. Phys.* **1996**, *87*, 1117–1157.

(36) Tuckerman, M. E.; Martyna, G. J. Understanding modern molecular dynamics: Techniques and applications. *J. Phys. Chem. B* **2000**, *104*, 159–178.

(37) Lamoureux, G.; Roux, B. Modeling induced polarization with classical Drude oscillators: Theory and molecular dynamics simulation algorithm. *J. Chem. Phys.* **2003**, *119*, 3025–3039.

(38) Jiang, W.; Hardy, D. J.; Phillips, J. C.; Mackerell, A. D., Jr.; Schulten, K.; Roux, B. High-performance scalable molecular dynamics simulations of a polarizable force field based on classical Drude oscillators in NAMD. *J. Phys. Chem. Lett.* **2011**, *2*, 87–92.

(39) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schafer, A.; Lennartz, C. QUASI: A general purpose implementation of the QM/MM approach and its application to problems in catalysis. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 1–28.

(40) Lopes, P. E. M.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, A. D., Jr. Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2013**, *9*, 5430–5449.

(41) Chowdhary, J.; Harder, E.; Lopes, P. E.; Huang, L.; MacKerell, A. D., Jr.; Roux, B. A Polarizable Force Field of Dipalmitoylphosphatidylcholine Based on the Classical Drude Model for Molecular Dynamics Simulations of Lipids. *J. Phys. Chem. B* **2013**, *117*, 9142–9160.

(42) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–65.

(43) Brooks, B. R.; Brooks, C. L., III; MacKerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caffisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoseck, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545–614.

(44) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.;

Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, 2009.

(45) Lamoureux, G.; Harder, E.; Vorobyov, I. V.; Roux, B.; MacKerell, A. D., Jr. A polarizable model of water for molecular dynamics simulations of biomolecules. *Chem. Phys. Lett.* **2006**, *418*, 245–249.

(46) Kong, J.; White, C. A.; Krylov, A. I.; Sherrill, C. D.; Adamson, R. D.; Furlani, T. R.; Lee, M. S.; Lee, A. M.; Gwaltney, S. R.; Adams, T. R.; Ochsenfeld, C.; Gilbert, A. T. B.; Kedziora, G. S.; Rassolov, V. A.; Maurice, D. R.; Nair, N.; Shao, Y.; Besley, N. A.; Maslen, P. E.; Dombroski, J. P.; Daschel, H.; Zhang, W.; Korambath, P. P.; Baker, J.; Byrd, E. F. C.; Voorhis, T. V.; Oumi, M.; Hirata, S.; Hsu, C.-P.; Ishikawa, N.; Florian, J.; Warshel, A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. Q-Chem 2.0: A high-performance ab initio electronic structure program. *J. Comput. Chem.* **2000**, *21*, 1532–1548.

(47) Shao, Y.; Fusti-Molnar, L.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; Jr, R. A. D.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Voorhis, T. V.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsels, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; III, H. L. W.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; III, H. F. S.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M.; Gan, Z.; Zhao, Y.; Schultz, N. E.; Truhlar, D.; Epifanovsky, E.; Oana, M. *Q-Chem 3.1*; Q-Chem, Inc.: Pittsburgh, PA, 2007.

(48) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors (Reprinted from *Molecular Physics*, vol 19, pg 553–566, 1970). *Mol. Phys.* **2002**, *100*, 65–73.

(49) Ransil, B. J. Studies in Molecular Structure. IV. Potential Curve for the Interaction of Two Helium Atoms in Single-Configuration LCAO MO SCF Approximation. *J. Chem. Phys.* **1961**, *34*, 2109–2118.

(50) Drew, H. R.; Dickerson, R. E. Structure of a B-DNA Dodecamer III. Geometry of Hydration. *J. Mol. Biol.* **1981**, *151*, 535–556.

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

(52) Drew, H. R.; Samson, S.; Dickerson, R. E. Structure of a B-DNA Dodecamer at 16 K. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4040–4044.

(53) Baker, C. M.; Anisimov, V. M.; MacKerell, A. D., Jr. Development of CHARMM polarizable force field for nucleic acid bases based on the classical Drude oscillator model. *J. Phys. Chem. B* **2011**, *115*, 580–96.

(54) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S.-H.; Srinivasan, A. R.; Schneider, B. The Nucleic Acid Database: A comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys. J.* **1992**, *63*, 751–759.

(55) Wiorkiewicz-Kuczera, J.; Kuczera, K.; Karplus, M. *MOLVIB*; Harvard University: Cambridge, MA, 1989.

(56) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. Systematic ab initio gradient calculation of molecular geometries, force constants, and dipole moment derivatives. *J. Am. Chem. Soc.* **1979**, *101*, 2550–2560.

(57) Scott, A. P.; Radom, L. Harmonic Vibrational Frequencies: An Evaluation of Hartree-Fock, Møller–Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempirical Scale Factors. *J. Phys. Chem.* **1996**, *100*, 16502–16513.

(58) Deng, Y.; Roux, B. Hydration of amino acid side chains: Non-polar and electrostatic contributions calculated from staged molecular

dynamics free energy simulations with explicit water molecules. *J. Phys. Chem. B* **2004**, *108*, 16567–16576.

(59) Weeks, J. D.; Chandler, D.; Andersen, H. C. Role of repulsive forces in determining the equilibrium structure of simple liquids. *J. Chem. Phys.* **1971**, *54*, 5237–5247.

(60) Baker, C. M.; Lopes, P. E. M.; Zhu, X.; Roux, B.; MacKerell, A. D., Jr. Accurate Calculation of Hydration Free Energies Using Pair-Specific Lennard-Jones Parameters in the CHARMM Drude Polarizable Force Field. *J. Chem. Theory Comput.* **2010**, *6*, 1181–1198.

(61) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.* **1982**, *76*, 637–649.

(62) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

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

(64) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695.

(65) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant-pressure molecular dynamics algorithms. *J. Chem. Phys.* **1994**, *101*, 4177–4189.

(66) Darden, T. A.; York, D.; Pedersen, L. G. Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(67) Leach, A. R. *Molecular Modelling: Principles and Applications*; Prentice Hall: Harlow, England; New York, 2001.

(68) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(69) Hummer, G.; Pratt, L. R.; Garcia, A. E.; Garde, S.; Berne, B. J.; Rick, S. W. Reply to comment on “Electrostatic potentials and free energies of solvation of polar and charged molecules”. *J. Phys. Chem. B* **1998**, *102*, 3841–3843.

(70) Harder, E.; Roux, B. On the origin of the electrostatic potential difference at a liquid-vacuum interface. *J. Chem. Phys.* **2008**, *129*.

(71) Luo, Y.; Roux, B. Simulation of Osmotic Pressure in Concentrated Aqueous Salt Solutions. *J. Phys. Chem. Lett.* **2010**, *1*, 183–189.

(72) Luo, Y.; Jiang, W.; Yu, H. A.; MacKerell, A. D., Jr.; Roux, B. Simulation study of ion pairing in concentrated aqueous salt solutions with a polarizable force field. *Faraday Discuss.* **2013**, *160*, 135–149.

(73) McQuarrie, D. A.; Simon, J. D. *Physical Chemistry: A Molecular Approach*; University Science Books: Sausalito, CA, 1997.

(74) Yoo, J.; Aksimentiev, A. Improved Parametrization of Li^+ , Na^+ , K^+ , and Mg^{2+} Ions for All-Atom Molecular Dynamics Simulations of Nucleic Acid Systems. *J. Phys. Chem. Lett.* **2012**, *3*, 45–50.

(75) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(76) Heddi, B.; Foloppe, N.; Oguey, C.; Hartmann, B. Importance of accurate DNA structures in solution: The Jun-Fos model. *J. Mol. Biol.* **2008**, *382*, 956–70.

(77) Eichman, B. F.; Vargason, J. M.; Mooers, B. H. M.; Ho, P. S. The Holliday junction in an inverted repeat DNA sequence: Sequence effects on the structure of four-way junctions. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3971–3976.

(78) Yu, H.; Whitfield, T. W.; Harder, E.; Lamoureux, G.; Vorobyov, I.; Anisimov, V. M.; MacKerell, A. D., Jr.; Roux, B. Simulating Monovalent and Divalent Ions in Aqueous Solution Using a Drude Polarizable Force Field. *J. Chem. Theory Comput.* **2010**, *6*, 774–786.

(79) Miyamoto, S.; Kollman, P. SETTLE: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13*, 952–962.

(80) Barrat, J.-L.; Hansen, J.-P. *Basic Concepts for Simple and Complex Liquids*; Cambridge University Press: Cambridge, U.K., 2003.

- (81) 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.
- (82) 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.
- (83) Kohlbacher, O.; Lenhof, H. P. BALL - Rapid software prototyping in computational molecular biology. *Bioinformatics* **2000**, *16*, 815–824.
- (84) Djuranovic, D.; Hartmann, B. Conformational Characteristics and Correlations in Crystal Structures of Nucleic Acid Oligonucleotides: Evidence for Sub-states. *J. Biomol. Struct. Dyn.* **2003**, *20*, 771–788.
- (85) Duchardt, E.; Nilsson, L.; Schleucher, J. Cytosine ribose flexibility in DNA: A combined NMR ¹³C spin relaxation and molecular dynamics simulation study. *Nucleic Acids Res.* **2008**, *36*, 4211–9.
- (86) Bax, A.; Lerner, L. Measurement of ¹H–¹H coupling constants in DNA fragments by 2D NMR. *J. Magn. Reson.* **1988**, *79*, 429–438.
- (87) Wu, Z. G.; Delaglio, F.; Tjandra, N.; Zhurkin, V. B.; Bax, A. Overall structure and sugar dynamics of a DNA dodecamer from homoand heteronuclear dipolar couplings and ³¹P chemical shift anisotropy. *J. Biomol. NMR* **2003**, *26*, 297–315.
- (88) Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: New York, 1997.
- (89) Halgren, T. A. Representation of van der Waals (vdW) Interactions In Molecular Mechanics Force-Fields - Potential Form, Combination Rules, And vdW Parameters. *J. Am. Chem. Soc.* **1992**, *114*, 7827–7843.
- (90) Halgren, T. A. Merck molecular force field 0.2. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J. Comput. Chem.* **1996**, *17*, 520–552.
- (91) Thole, B. T. Molecular polarizabilities calculated with a modified dipole interaction. *Chem. Phys.* **1981**, *59*, 341.
- (92) Harder, E.; Anisimov, V. M.; Whitfield, T. W.; MacKerell, A. D., Jr.; Roux, B. Understanding the dielectric properties of liquid amides from a polarizable force field. *J. Phys. Chem. B* **2008**, *112*, 3509–3521.
- (93) Manning, G. S. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* **1978**, *11*, 179–246.
- (94) McDonald, A. R.; Denning, E. J.; MacKerell, A. D., Jr. Impact of Geometry Optimization on Base-Base Stacking Interaction Energies in the Canonical A- and B-Forms of DNA. *J. Phys. Chem. A* **2013**, *117*, 1560–1568.
- (95) Sychrovsky, V.; Foldynova-Trantirkova, S.; Spackova, N.; Robeyns, K.; Van Meervelt, L.; Blankenfeldt, W.; Vokacova, Z.; Sponer, J.; Trantirek, L. Revisiting the planarity of nucleic acid bases: Pyrimidilization at glycosidic nitrogen in purine bases is modulated by orientation of glycosidic torsion. *Nucleic Acids Res.* **2009**, *37*, 7321–7331.
- (96) MacKerell, A. D., Jr.; Banavali, N. K. All-atom empirical force field for nucleic acids: 2) Application to solution MD simulations of DNA. *J. Comput. Chem.* **2000**, *21*, 105–120.
- (97) Lum, R. C.; Grabowski, J. J. Trimethyl phosphate: the intrinsic reactivity of carbon versus phosphorus sites with anionic nucleophiles. *J. Am. Chem. Soc.* **1992**, *114*, 8619–8627.
- (98) Lopez, X.; Schaefer, M.; Dejaegere, A.; Karplus, M. Theoretical evaluation of pK_a in phosphoranes: Implications for phosphate ester hydrolysis. *J. Am. Chem. Soc.* **2002**, *124*, 5010–5018.
- (99) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. SM6: A density functional theory continuum solvation model for calculating aqueous solvation free energies of neutrals, ions, and solute–water clusters. *J. Chem. Theory Comput.* **2005**, *1*, 1133–1152.
- (100) Dolney, D. M.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. Universal solvation model based on conductor-like screening model. *J. Comput. Chem.* **2000**, *21*, 340–366.
- (101) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Universal quantum mechanical model for solvation free energies based on gas-phase geometries. *J. Phys. Chem. B* **1998**, *102*, 3257–3271.
- (102) McMillan, W. G.; Mayer, J. E. The Statistical Thermodynamics of Multicomponent Systems. *J. Chem. Phys.* **1945**, *13*, 276–305.
- (103) Robinson, R. A.; Stokes, R. H. *Electrolyte Solutions*; Dover Publications, Inc: 2002.
- (104) Tamaki, K.; Suga, K.; Tanihara, E. Solution Properties of Dialkyl Phosphate Salts - Apparent Molar Volumes, Viscosity B-Coefficients, Heats of Solution, And Osmotic Coefficients. *Bull. Chem. Soc. Jpn.* **1987**, *60*, 1225–1229.
- (105) Shui, X. Q.; 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.
- (106) Tereshko, V.; Minasov, G.; Egli, M. A “hydrat-ion” spine in a B-DNA minor groove. *J. Am. Chem. Soc.* **1999**, *121*, 3590–3595.
- (107) Tereshko, V.; Minasov, G.; Egli, M. The Dickerson–Drew B-DNA dodecamer revisited at atomic resolution. *J. Am. Chem. Soc.* **1999**, *121*, 470–471.
- (108) Tereshko, V.; Wilds, C. J.; Minasov, G.; Prakash, T. P.; Maier, M. A.; Howard, A.; Wawrzak, Z.; Manoharan, M.; Egli, M. Detection of alkali metal ions in DNA crystals using state-of-the-art X-ray diffraction experiments. *Nucleic Acids Res.* **2001**, *29*, 1208–1215.
- (109) Denisov, V. P.; Halle, B. Sequence-specific binding of counterions to B-DNA. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 629–633.
- (110) Halle, B.; Denisov, V. P. Water and monovalent ions in the minor groove of B-DNA oligonucleotides as seen by NMR. *Biopolymers* **1998**, *48*, 210–233.
- (111) Marincola, F. C.; Denisov, V. P.; Halle, B. Competitive Na⁺ and Rb⁺ binding in the minor groove of DNA. *J. Am. Chem. Soc.* **2004**, *126*, 6739–6750.
- (112) Hud, N. V.; Feigon, J. Characterization of divalent cation localization in the minor groove of the A_nT_n and T_nA_n DNA sequence elements by ¹H NMR spectroscopy and manganese(II). *Biochemistry* **2002**, *41*, 9900–9910.
- (113) Varnai, P.; Zakrzewska, K. DNA and its counterions: A molecular dynamics study. *Nucleic Acids Res.* **2004**, *32*, 4269–4280.
- (114) Rueda, M.; Cubero, E.; Laughton, C. A.; Orozco, M. Exploring the counterion atmosphere around DNA: What can be learned from molecular dynamics simulations? *Biophys. J.* **2004**, *87*, 800–811.
- (115) Feig, M.; Pettitt, B. M. Sodium and chlorine ions as part of the DNA solvation shell. *Biophys. J.* **1999**, *77*, 1769–1781.
- (116) Hamelberg, D.; Williams, L. D.; Wilson, W. D. Influence of the dynamic positions of cations on the structure of the DNA minor groove: Sequence-dependent effects. *J. Am. Chem. Soc.* **2001**, *123*, 7745–7755.
- (117) Manning, G. S. Limiting Laws and Counterion Condensation in Polyelectrolyte Solutions I. Colligative Properties. *J. Chem. Phys.* **1969**, *51*, 924–933.
- (118) Manning, G. S. Is a small number of charge neutralizations sufficient to bend nucleosome core DNA onto its superhelical ramp? *J. Am. Chem. Soc.* **2003**, *125*, 15087–15092.
- (119) Deserno, M.; Holm, C.; May, S. Fraction of condensed counterions around a charged rod: Comparison of Poisson–Boltzmann theory and computer simulations. *Macromolecules* **2000**, *33*, 199–206.
- (120) Vorobyov, I.; Anisimov, V. M.; Greene, S.; Venable, R. M.; Moser, A.; Pastor, R. W.; MacKerell, A. D., Jr. Additive and classical drude polarizable force fields for linear and cyclic ethers. *J. Chem. Theory Comput.* **2007**, *3*, 1120–1133.
- (121) Lopes, P. E. M.; Lamoureux, G.; Roux, B.; MacKerell, A. D., Jr. Polarizable empirical force field for aromatic compounds based on the classical drude oscillator. *J. Phys. Chem. B* **2007**, *111*, 2873–2885.
- (122) Anisimov, V. M.; Vorobyov, I. V.; Roux, B.; MacKerell, A. D., Jr. Polarizable empirical force field for the primary and secondary alcohol series based on the classical drude model. *J. Chem. Theory Comput.* **2007**, *3*, 1927–1946.
- (123) Harder, E.; Anisimov, V. M.; Vorobyov, I. V.; Lopes, P. E. M.; Noskov, S. Y.; MacKerell, A. D., Jr.; Roux, B. Atomic level anisotropy in the electrostatic modeling of lone pairs for a polarizable force field

based on the classical Drude oscillator. *J. Chem. Theory Comput.* **2006**, *2*, 1587–1597.

(124) Zhu, X.; MacKerell, A. D., Jr. Polarizable empirical force field for sulfur-containing compounds based on the classical Drude oscillator model. *J. Comput. Chem.* **2010**, *31*, 2330–41.

(125) Baker, C. M.; MacKerell, A. D., Jr. Polarizability rescaling and atom-based Thole scaling in the CHARMM Drude polarizable force field for ethers. *J. Mol. Model.* **2009**, *16*, 567–576.

(126) Lopes, P. E. M.; Lamoureux, G.; MacKerell, A. D., Jr. Polarizable Empirical Force Field for Nitrogen-Containing Heteroaromatic Compounds Based on the Classical Drude Oscillator. *J. Comput. Chem.* **2009**, *30*, 1821–1838.

(127) He, X.; Lopes, P. E. M.; MacKerell, A. D., Jr. Polarizable Empirical Force Field for Acyclic Polyalcohols Based on the Classical Drude Oscillator. *Biopolymers* **2013**, *99*, 724–738.

(128) Patel, D. S.; He, X.; MacKerell, A. D., Jr. Polarizable Empirical Force Field for Hexopyranose Monosaccharides based on the Classical Drude Oscillator. *J. Phys. Chem. B* **2014**, DOI: 10.1021/jp412696m.