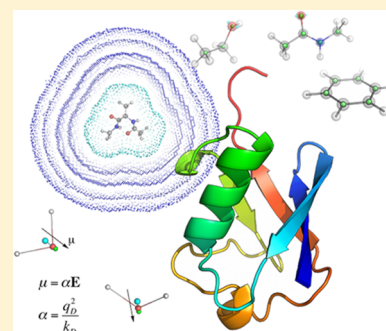# An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications

Justin A. Lemkul,[†] Jing Huang,[†] Benoît Roux,[‡] and Alexander D. MacKerell, Jr.*,[†]

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

[‡]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637, United States

**ABSTRACT:** Molecular mechanics force fields that explicitly account for induced polarization represent the next generation of physical models for molecular dynamics simulations. Several methods exist for modeling induced polarization, and here we review the classical Drude oscillator model, in which electronic degrees of freedom are modeled by charged particles attached to the nuclei of their core atoms by harmonic springs. We describe the latest developments in Drude force field parametrization and application, primarily in the last 15 years. Emphasis is placed on the Drude-2013 polarizable force field for proteins, DNA, lipids, and carbohydrates. We discuss its parametrization protocol, development history, and recent simulations of biologically interesting systems, highlighting specific studies in which induced polarization plays a critical role in reproducing experimental observables and understanding physical behavior. As the Drude oscillator model is computationally tractable and available in a wide range of simulation packages, it is anticipated that use of these more complex physical models will lead to new and important discoveries of the physical forces driving a range of chemical and biological phenomena.

## CONTENTS

## 1. INTRODUCTION

Atomistic molecular dynamics (MD) simulations have become an integral component of the tool set used to examine biomolecular systems. Since the first MD simulation of a protein in 1977,[1] the time scales and sizes of computationally tractable simulations have grown by orders of magnitude. Systems may now be simulated on the microsecond[2] or millisecond[3] time scale and contain over a million atoms[4,5] due to increased computing power and ever-improving algorithms for parallel and graphical processing unit (GPU)-based computing. Such capabilities also allow for more rigorous testing of the accuracy of the models used for such MD simulations.

Quantum mechanics (QM) provides a highly accurate representation of the energetics of molecules.[6−8] While some progress has been made in the application of QM methods to

biological macromolecules, for example, hybrid quantum mechanics/molecular mechanics (QM/MM) methods that have been utilized for decades,[9−11] applying QM to large biomolecular systems remains infeasible in most cases. As a result, MD simulations have relied upon empirical potential energy functions (so-called "force fields") to calculate the energies of configurations and, thus, the forces on all constituent atoms required to perform dynamic simulations. Thus, the predictive ability of MD simulations relies on the accuracy of the underlying force fields that describe the nature of the interactions between all components in the simulation system.

Modeling molecular processes accurately is made challenging by the wide range of interactions that atoms or molecules can experience in condensed-phase environments. For instance, a protein in a cell may interact with other proteins with polar and nonpolar surfaces, amphipathic lipid membranes, and highly charged species like ions and nucleic acids (DNA and RNA), all while diffusing through a predominantly polar, aqueous environment. Small molecules (xenobiotics, metabolites, etc.) may partition across hydrophobic membranes or bind to specific transporters before being delivered to intracellular receptors or enzymes. Thus, their dynamics must also be modeled in a manner that accounts for these diverse environments. Describing all of these processes using computationally feasible functional forms remains a challenge in the area of force field development. To date, the majority of force fields for molecular simulations have been based on pairwise or additive force fields in which the partial atomic charges are fixed. However, this convention represents a significant approximation that limits the ability of the force field to be accurate in the diverse range of environments present in biological and chemical condensed-phase systems. To overcome this limitation, more complex force fields that include explicit electronic polarization that allow the charge distribution to vary as a function of environment are being developed. In this review, the development of a force field based on the Drude oscillator model[12] will be described, including the theoretical background and parametrization methodology, as well as applications to biomolecular systems in which explicit polarization has provided unique insights into functional dynamics. Traditional additive force fields will be described briefly as an introduction to the fundamentals of force field development. The review will conclude with an outlook on the next steps in Drude polarizable force field development and systems to which we anticipate the inclusion of polarizability will be critical. We note that the review will focus on the Drude-2013 polarizable force field, which has its origins in the CHARMM additive force field, with the development of the force field initiated in 2000 with the first polarizable water model published in 2003.[13] While there are efforts by other groups to develop polarizable models, a comprehensive evaluation of all of these force fields is beyond the scope of the present review. To provide perspective on the field, we will briefly discuss efforts by other investigators using alternate models in section 3.1.

## 2. ADDITIVE FORCE FIELDS

Most MD simulations to date have used fixed-charge (additive) force fields. A force field has two components, a functional form for calculating energies, and a parameter set that populates the variables in the functional form. These force fields are characterized by point charges on each of the atoms, centered on the atomic nucleus. The sum of these pairwise interactions gives the total electrostatic energy, hence the name "additive." There are many popular atomistic force fields for biomolecular simulations, the most widely used of which are AMBER,[14−22] CHARMM,[23−34] GROMOS,[35−38] and OPLS.[39−41] As an example, the functional form of the CHARMM force field is as follows:

$$
\begin{aligned}
U(\vec{R}) = &\sum_{\text{bonds}} k_{\text{b}}(b - b_0)^2 + \sum_{\text{angles}} k_\theta(\theta - \theta_0)^2 \\
&+ \sum_{\text{Urey−Bradley}} k_{\text{UB}}(S - S_0)^2 \\
&+ \sum_{\text{dihedrals}} k_\phi(1 + \cos(n\phi - \delta)) \\
&+ \sum_{\text{residues}} U_{\text{CMAP}}(\phi, \psi) + \sum_{\text{impropers}} k_\omega(\omega - \omega_0)^2 \\
&+ \sum_{\text{nonbonded pairs}} \left\{ \varepsilon_{ij}^{\text{min}} \left[ \left( \frac{R_{ij}^{\text{min}}}{r_{ij}} \right)^{12} - 2 \left( \frac{R_{ij}^{\text{min}}}{r_{ij}} \right)^6 \right] \right. \\
&\left. + \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon r_{ij}} \right\}
\end{aligned}
\tag{1}
$$

The calculated potential energy of a configuration, $U(\vec{R})$, is thus a sum of internal (bonded) and noncovalent (nonbonded) interactions. Bonds ($b$), valence angles ($\theta$), Urey−Bradley ($S$), dihedral angles (or torsions) ($\phi$), CMAP backbone torsion corrections ($\phi, \psi$),[42] and improper dihedrals ($\omega$) comprise the internal terms. Most of these terms are described by simple harmonic functions with a force constant, $k$, with respect to an equilibrium value (subscript 0). The dihedral term describing bond rotations is not harmonic; instead, it is a periodic term characterized by a force constant, multiplicity ($n$), and phase shift ($\delta$). Urey−Bradley cross-terms describe interactions between atoms $i$ and $k$ (so-called "1−3 interactions") in a valence angle formed by atoms $i−j−k$. Improper dihedrals describe out-of-plane vibrations of planar groups. The non-bonded interactions are comprised of van der Waals and electrostatic terms. The van der Waals component is described by a Lennard-Jones (LJ or "6−12 potential") form, and electrostatic interactions are calculated using Coulomb's law with point charges $q_i$ and $q_j$ on atoms $i$ and $j$, respectively. This potential energy function shares many common features with those of the other force fields.

While additive force fields have enjoyed many years of development and success in describing macromolecular dynamics,[43] there are inherent problems and limitations with these models. Additive force fields are generally parametrized using a mean-field approximation, such that they respond in an average way to the different environments that a molecule might experience during the course of a simulation, a consideration that is especially important in terms of the electrostatic parameters. A common strategy for implicitly representing polarization response in additive force fields is to overestimate the gas-phase dipole moments from QM calculations, typically on the order of 20% or more.[44] While this approach is tractable from the perspective of parametrization and computational cost, this underlying approximation has important consequences.[45,46] Many biological processes involve the interaction of molecules that pass

through or interact with the predominantly aqueous milieu as well as less polar environments, such as substrate binding to a soluble enzyme (which requires at least partial desolvation prior to entering the binding site) or passage of small molecules through membranes (passing into and through the hydrophobic core of the membrane). The electronic distribution of these molecules will change in response to variations in the local electric field, but assigning fixed charges prevents such a response and is inherently a deficient description of the interactions in the system that limits accuracy.[47] That is to say, the inaccuracies in additive models are not necessarily due to inadequate parameters; rather, they reflect a fundamental limitation in the functional form. It is for this reason that an extended description of the physical model, one that includes explicit polarization, is desirable.

## 3. POLARIZABLE FORCE FIELDS

### 3.1. Methods for Modeling Explicit Polarization

Several methods have been developed to treat polarization explicitly in MD simulations. Though the present review focuses on the classical Drude oscillator model, it is important to briefly summarize other approaches that have been developed before moving on to a more in-depth discussion of the Drude-2013 polarizable force field. More comprehensive reviews of these polarization methods have been presented previously.[45,48−51]

Given that empirical force fields are approximations of QM representations, modeling induced polarization via QM as an extension to the empirical energy function is an intuitive approach. The X-POL potential[52] is one such approach that is related to QM/MM simulations. Unlike traditional QM/MM, in which part of the simulation is treated via QM, in the X-POL potential, the addition of a QM layer to the MM system accounts for polarization effects. The application of a semiempirical QM method reduces the computational cost relative to ab initio approaches, but the cost of such simulations remains very high. As such, it is attractive to represent many-body effects classically, as a direct extension or modification of the MM potential. The three primary methods for treating electronic polarization classically are induced point dipole,[9,53−58] fluctuating charge,[59−63] and the Drude oscillator (also called "shell"[64] or "charge-on-a-spring"[65]) model.[12] An outgrowth of the Drude oscillator model is the QMPFF general force field,[66,67] which treats polarization via the attachment of electron clouds to atomic cores. Rather than using a discrete particle to represent electronic degrees of freedom, polarization is modeled on the basis of the distance between the cloud center and the core atom.

In the induced point dipole representation, each atomic site is assigned a dipole that is calculated using an iterative self-consistent field (SCF) procedure (see section 3.2.2 below). The electrostatic energy of the system is then calculated from the charge−charge, charge−dipole, and dipole−dipole interactions. The most notable example of an induced dipole model is the AMOEBA force field,[57,58] which also includes multipoles up to quadrupoles in the treatment of the electrostatic interactions. Other induced dipole models include those presented by Berne, Friesner, and co-workers,[53,54] as well as the POSSIM model,[68] among others.[69−87]

The fluctuating charge model treats the point charges on the atoms as dynamical variables, such that the topology can vary throughout the simulation. Charges on individual molecules can redistribute according to the electronegativity of each atom, though the overall charge on the molecule is maintained. The CHeq force field for proteins,[62,63] lipids,[88] and carbohydrates[89] is one such model that uses fluctuating charges, and other fluctuating charge models have been reported.[90,91] The CHeq model has been successfully employed in studies of protein−ligand binding,[92] ion solvation,[93] and diffusion of arginine across a model lipid membrane.[94] The main limitation with fluctuating charge models is that they cannot account for out-of-plane polarization in conjugated systems, though, in principle, the model could be extended to account for this phenomenon, e.g., by using a semiempirical QM/MM approach[95] or in a purely MM approximation by either combination with the Drude oscillator model[96] or the use of off-center virtual sites to which charge can be distributed.

The Drude oscillator model is a direct extension of additive models and retains many of the pairwise features of its functional form. It simply consists of attaching an auxiliary particle carrying a charge to an atom via a harmonic spring to simulate induced polarization via its displacement under the influence of an electric field. In early studies of ion solvation with the fluctuating charge model for water, the Drude formalism was used for the ions,[96] since for a monatomic ion there is no other way to redistribute charge in the fluctuating charge method. For these reasons, its implementation in MD codes to simulate chemical and biological systems is fairly straightforward, and the approach has been used in a number of laboratories[64,65,96−98] in addition to our own efforts.

With all polarizable models, the requirement to solve the magnitude of all the induced dipoles self-consistently is computationally demanding and is typically a limiting factor in the efficiency of such simulations. A number of approaches have been contemplated to overcome the limitation of the SCF procedure. The most common method has been to attribute an effective mass and kinetic energy to the degrees of freedom associated with the induced dipoles via an extended Lagrangian framework. This concept goes back to the so-called Car−Parrinello MD method to propagate a system in density functional theory.[99] An alternative route has been adopted by the POSSIM force field[68] and the iAMOEBA model for water,[100] whereby only a single evaluation of mutual polarization is carried out. This idea of explicitly treating first-order induction explicitly while accounting for higher-order contributions approximately via an empirical modification of force field parameters was first introduced by Straatsma and McCammon.[101] In this manner, the induced dipoles are solved in the electric field of the fixed charges and multipoles in the system, and the costly iterative SCF procedure is avoided. This approach is computationally more efficient than a full SCF procedure, but it is an approximation that neglects the explicit many-body effects of the other induced dipoles in the system. Consequently, the parametrization of the model accounts for an average effect of induced dipoles, an approximation analogous to traditional assumptions in additive force fields. Though the iAMOEBA approach is likely adequate for addressing polarization effects in homogeneous systems such as water, it remains to be seen if such an approximation is valid for heterogeneous environments or in the presence of highly charged particles. For example, the coordination of an ion is controlled by a delicate balance involving the interactions between the induced dipoles of the ligands, an effect that is neglected in a first-order approximation. A recent alternate formulation of the induced dipole approach treats the

polarization response using perturbation theory rather than a variational approach to achieve the SCF condition, yielding an improvement in computational efficiency.[102]

The remainder of this review will focus on the Drude oscillator model of induced polarization. It will cover the functional form of the Drude model, parametrization strategy, and development history of the Drude-2013 force field and its applications to biological systems and will conclude with a prospective view of future studies that will benefit from this polarizable model. While focusing on Drude-2013, it is important to note other recent Drude-based polarizable FF studies include the use of multipole expansion to simulate ionic liquids,[97] with the polarization effects yielding important improvements in the translational and rotational behavior of the molecular ions, and similar behavior was observed in a Drude model.[98] Explicit treatment of polarizability in these systems led to important observations about contributions to the dielectric constant that are not possible with nonpolarizable force fields, which, by virtue of their fixed charges, often retard dynamics.[98] Polarizability led to enhanced fluidity and reduced long-range structure using both induced point-dipole and Drude models, each with comparable accuracy.[103]

### 3.2. The Drude Oscillator Model

The central feature of the atomic force field based on the classical Drude model that we have been developing for molecular simulations is the inclusion of electronic polarizability on all non-hydrogen atoms, as described in detail in the next section. In addition, the model includes virtual particles representative of lone pairs, typically located on hydrogen-bond-acceptor atoms. The remainder of the functional form in the Drude-2013 polarizable force field shares many common elements with that of the additive model shown in eq 1.

**3.2.1. The Polarizable Functional Form.** The isotropic polarizability, $\alpha$, of a given atom is described by the distribution of the total charge on the atom, $q$, between the core atom and its Drude oscillator, which is attached to the core atom via a harmonic spring, as shown schematically in Figure 1A. The charges assigned to the Drude oscillator ($q_D$) and to the core atom ($q_A$) are such that $q = q_A + q_D$. The atomic polarizability is defined as

$$\alpha = \frac{q_D^2}{k_D} \tag{2}$$

where $k_D$ is the force constant of the Drude−atom harmonic bond. In principle, $q_D$, $k_D$, or both can be tuned to achieve an appropriate polarization response. In the present Drude-2013 force field, the value of $k_D$ is set to 1000 kcal mol$^{-1}$ Å$^{-2}$ for all Drude−atom pairs. This convention reduces the complexity of parametrization by reducing the number of parameters to be fit. It follows from this description of polarization that there exists some displacement, $\mathbf{d}$, between the Drude oscillator and the core atom in response to an electric field, $\mathbf{E}$, according to

$$\mathbf{d} = \frac{q_D \mathbf{E}}{k_D} \tag{3}$$

from which the induced atomic dipole can be calculated as

$$\mu = \frac{q_D^2 \mathbf{E}}{k_D} \tag{4}$$

The electrostatic component of the system's potential energy includes a harmonic self-polarization term, $U_{self}$, which is
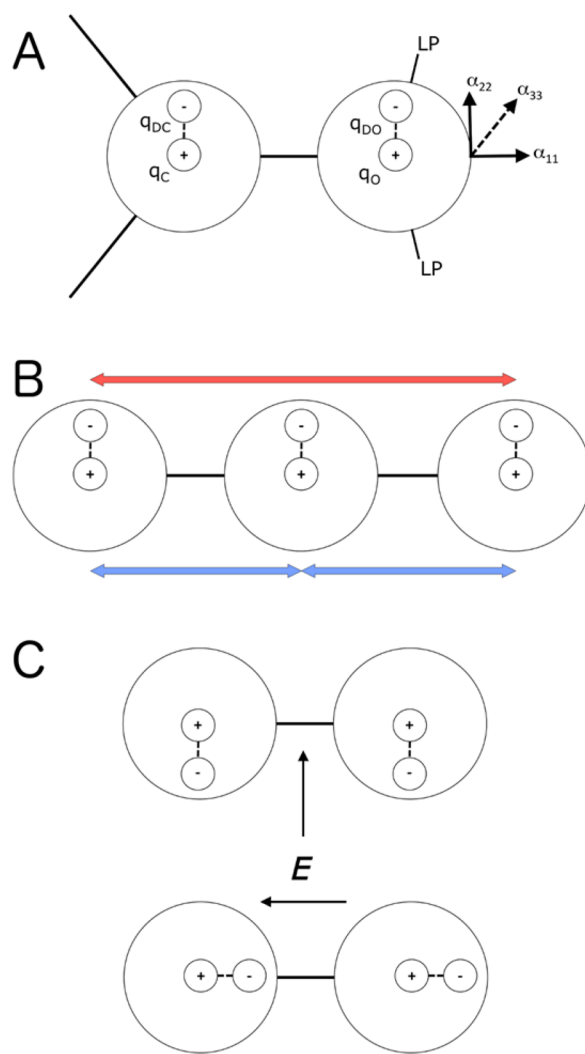


**Figure 1.** Schematic of the Drude oscillator model. (A) The addition of Drude oscillators to carbon and oxygen atoms via harmonic springs and the subsequent distribution of charge between the atoms ($q_C$ and $q_O$) and their respective Drude oscillators ($q_{DC}$ and $q_{DO}$). Lone pairs on the oxygen atom are labeled "LP" and the axes of the local molecular frame used for anisotropic polarization are shown. Values of $\alpha_{11}$, $\alpha_{22}$, and $\alpha_{33}$ are the tensor components of the anisotropic atomic polarizability along the indicated axes. (B) 1−2 and 1−3 dipole−dipole interactions are indicated by blue and red double-headed arrows, respectively. These interactions are not present in additive force fields but are treated explicitly via Thole screening in the Drude-2013 force field, contributing to anisotropic molecular polarizability. (C) The directional response of the Drude oscillators to external electric fields (**E**) is shown for a simple diatomic where, due to the 1−2 dipole−dipole interactions when the electric field is perpendicular to the bond, the 1−2 dipoles damp each other, leading to a smaller polarization response compared to the electric field being parallel to the bond leading to the 1−2 dipoles enhancing each other.

calculated on the basis of the harmonic bond energy between the Drude oscillators and their core atoms. Self-polarization is thus expressed as

$$U_{self}(\mathbf{d}) = \frac{1}{2} k_D \mathbf{d}^2 \tag{5}$$

with $k_D$ and $\mathbf{d}$ defined above. As implied by eq 3, the displacement vector is the result of many-body effects induced

by the electric field **E** resulting from the configuration of the charged particles in the system.

The polarizability of a given atom need not be isotropic; that is, the Drude oscillator can deform differentially in all three spatial dimensions. Anisotropic polarization was introduced in the Drude-2013 force field to more accurately describe functional groups acting as hydrogen-bond acceptors.[104] The anisotropic form is achieved by modifying $U_{self}$, which is written as

$$U_{self} = \frac{1}{2}([\mathbf{K}_{11}^{(D)}]d_1^2 + [\mathbf{K}_{22}^{(D)}]d_2^2 + [\mathbf{K}_{33}^{(D)}]d_3^2) \tag{6}$$

in which $d_1$, $d_2$, and $d_3$ represent the components of the Drude–atom displacement vector **d** in a coordinate frame based on the local atomic geometry (Figure 1A). A unit vector $\hat{n}_{i,j}$ is defined between atoms $i$ and $j$ from which each component of the displacement vector is calculated, e.g., $d_1 = \hat{n}_{i,j} \cdot \mathbf{d}$ along the first axis. The isotropic force constant $k_D$ is expanded to a tensor, $\mathbf{K}^{(D)}$, with off-diagonal elements set to zero. Consistent with the isotropic $k_D$, the anisotropic polarization tensor is chosen such that $\text{tr}[\mathbf{K}^{(D)}]/3 = 1000$ kcal mol$^{-1}$ Å$^{-2}$.[104] As mentioned above, atoms described by anisotropic polarization are also generally augmented by "lone pairs" (virtual sites carrying negative charge) to further improve the description of electronic distribution around the polarizable heavy atom (Figure 1A). The combination of lone pairs and anisotropic polarization leads to an improved description of hydrogen bonding in polar compounds and interactions with ions as a function of orientation.[104]

One significant difference between the polarizable and additive functional forms is the explicit inclusion of dipole–dipole interactions for atoms within three bonds (Figure 1B). Normally, in additive force fields, nonbonded interactions between atoms connected by one or two bonds (so-called 1–2 and 1–3 pairs) are excluded and their effects are implicitly represented in the parametrization of bond-stretching and angle-bending terms. In the Drude polarizable model, interactions between induced dipoles (but not charge–dipole interactions) are explicitly included for 1–2 and 1–3 atom pairs, with those electrostatic interactions screened[105] by

$$S_{ij}(r_{ij}) = 1 - \left[\left(1 + \frac{ar_{ij}}{2(\alpha_i\alpha_j)^{1/6}}\right)\right]\exp\left[\frac{-ar_{ij}}{(\alpha_i\alpha_j)^{1/6}}\right] \tag{7}$$

where $r_{ij}$ is the distance between atoms $i$ and $j$, $\alpha_i$ and $\alpha_j$ are respective atomic polarizabilities, and $a$ is a damping constant. Equation 7 is equivalent to a smeared charge distribution described by

$$\rho = \frac{a^3}{8\pi}\exp\left[-a\frac{r_{ij}}{(\alpha_i\alpha_j)^{1/6}}\right] \tag{8}$$

as originally proposed by Thole.[105] A value of 2.6 for $a$, similar to the 2.089 suggested by Thole, was found to reproduce the polarizability anisotropy ratio of benzene[106] and was used initially as the default value in the Drude-2013 force field. Subsequently, atom-specific values were implemented, as they may be assigned to fine-tune the molecular polarizability tensor[107]

$$a = a_i + a_j \tag{9}$$

where $a_i$ and $a_j$ are respective atomistic Thole factors. Interactions between atoms and Drude oscillators separated by three or more bonds are calculated according to the normal (unscreened) Coulombic potential. Topological bonds involving only the real atoms determine bonded connectivity; atom–Drude bonds do not count in making this determination. It is worth pointing out that without such 1–2 and 1–3 dipole–dipole interactions, the polarizability tensor of molecular groups is trivially diagonal by default (i.e., the polarizability of the molecule in Figure 1C would be the same, whether the field is parallel or orthogonal to the chemical bond).

In the fitting procedure used to optimize the parameters, the starting values for $\alpha$ are taken from molecular polarizabilities from the Miller atomic hybrid polarizability (ahp) model.[108] These values are based on the hybridization state of each atom. Given that the Drude-2013 force field does not treat hydrogen atoms as polarizable, the $\alpha$ values of any hydrogen atoms bonded to a heavy atom are added to the $\alpha$ value of that heavy atom. Initial guesses for atomic charges and bonded parameters originated from the additive CHARMM force field.[34] From these initial guesses, various fitting and refinement procedures have been used to obtain agreement with a wide range of target data (see section 3.3).

**3.2.2. Molecular Dynamics Algorithm.** A practical consideration in carrying out polarizable MD simulations is the efficiency of the integration scheme. There are two principal methods by which simulations using the Drude oscillator model can be carried out. The first is the SCF regime, which is an intuitive extension of the Born–Oppenheimer approximation. That is, the electronic degrees of freedom relax to their ground state immediately upon any change in nuclear configuration. In an MD simulation with the SCF approach, the positions of the Drude oscillators are solved via energy minimization at each integration step such that

$$\frac{\partial U}{\partial \mathbf{d}_i} = 0 \tag{10}$$

where the index $i$ runs over all the Drude–atom pairs in the simulation system. This approach is computationally demanding, as multiple force evaluations are required at each MD time step, and convergence is not guaranteed with limited precision. Numerical errors could also be accumulated depending on the convergence criterion, initial guess, and minimization algorithms.[109]

As an alternative to the SCF regime, an extended Lagrangian approach[59,110] was developed by Lamoureux and Roux for the Drude model,[111] in which each Drude oscillator is ascribed a small mass $(m_D)$ that is subtracted from the parent atom such that the total mass of the Drude–atom pair remained equal to the atomic mass. In practice, this mass is set to 0.4 amu. Since the Drude oscillators are not massless, their positions can be integrated on equal dynamical footing with the real atoms of the system. A velocity–Verlet algorithm is used to numerically integrate the equations of motion, which use the absolute positions of the atoms and their Drude oscillators ($\mathbf{r}_i$ and $\mathbf{r}_{D,i}$, respectively) to compute forces on the Drude–atom centers of mass ($\mathbf{R}_i$) and Drude–atom displacements ($\mathbf{d}_i$), which are

$$\mathbf{F}_{\mathbf{R},i} = -\frac{\partial U}{\partial \mathbf{r}_i} - \frac{\partial U}{\partial \mathbf{r}_{D,i}} \tag{11}$$

$$\mathbf{F}_{\mathbf{d},i} = -\left(1 - \frac{m_{\mathrm{D}}}{m_i}\right)\frac{\partial U}{\partial \mathbf{r}_{\mathrm{D},i}} + \left(\frac{m_{\mathrm{D}}}{m_i}\right)\frac{\partial U}{\partial \mathbf{r}_i} \qquad (12)$$

To approximate the SCF energy surface, dual Nosé−Hoover thermostats[112,113] are used to regulate the temperature, $T$, of the real atoms (the "physical" thermostat), and the relative velocities of the Drude oscillators are regulated at some low temperature, $T_*$, such that $T_* \ll T$. The equations of motion in the canonical ($NVT$) ensemble are thus expressed as

$$m_i \ddot{\mathbf{R}}_i = \mathbf{F}_{\mathbf{R},i} - m_i \dot{\mathbf{R}}_i \dot{\eta} \qquad (13)$$

$$m_i' \ddot{\mathbf{d}}_i = \mathbf{F}_{\mathbf{d},i} - m_i' \dot{\mathbf{d}}_i \dot{\eta}_* \qquad (14)$$

$$Q\ddot{\eta} = \sum_j m_j \dot{R}_j^{\,2} - N_f k_{\mathrm{B}} T \qquad (15)$$

$$Q_*\ddot{\eta}_* = \sum_j m_j' \dot{d}_j^{\,2} - N_{f*} k_{\mathrm{B}} T_* \qquad (16)$$

where $\mathbf{F}_{\mathbf{R},i}$ and $\mathbf{F}_{\mathbf{d},i}$ are given by eqs 11 and 12. The reduced mass of each Drude−atom pair, $m_i'$, is calculated as $m_i' = (1 - m_{\mathrm{D}}/m_i)$. $N_f$ and $N_{f*}$ are the number of degrees of freedom in the physical thermostat and the low-temperature Drude oscillator thermostat, respectively. The forces acting on each thermostat are dynamical variables that are used to calculate the friction coefficients $\dot{\eta}$ and $\dot{\eta}_*$ on each, per eqs 15 and 16. These friction coefficients are used to scale center-of-mass and relative velocities of the Drude−atom pairs. Each integration step propagates the velocity−Verlet term and the multistep Nosé−Hoover terms according to the procedure of Martyna et al.[114] The isothermal−isobaric ($NPT$) ensemble commonly used in biomolecular MD simulations is also supported.[111]

The dual-thermostat extended Lagrangian integration scheme has been implemented in CHARMM,[111] NAMD (as Langevin dynamics rather than velocity−Verlet),[115] ChemShell QM/MM,[116] OpenMM,[117] and GROMACS[118] and is more efficient than the SCF approach. The use of an extended Lagrangian allows for highly scalable MD simulations[115,118] that are only about 2 times slower than additive simulations with a comparable number of atoms, assuming the same integration time step, However, Drude simulations are typically performed with a 1 fs time step versus 2 fs for additive simulations such that the overall difference between the additive and Drude models is approximately a factor of 4. The stability of the integration is enhanced by the use of a "hard-wall constraint"[119] to prevent polarization catastrophe, which can occur when the forces acting on the Drude oscillator cause large displacement from its core atom. The constraint reflects the Drude oscillator back toward its core atom, along the Drude−atom bond, if it reaches a predefined limit, typically set to 0.2 Å. This approach is a simpler form than what was proposed by Donchev et al. in QMPFF, which uses a distance-dependent force constant to prevent polarization catastrophe.[66] Importantly, the efficiency of the Drude extended Lagrangian integration algorithms allows for simulations of biologically relevant systems on the microsecond scale.[120]

### 3.3. Drude-2013 Polarizable Force Field Development

**3.3.1. Parametrization Strategy.** The general parametrization scheme for the Drude-2013 polarizable force field is summarized in Figure 2. The protocol relies on a strong connection between target data based on QM calculations for initial parametrization and experimental, condensed-phase
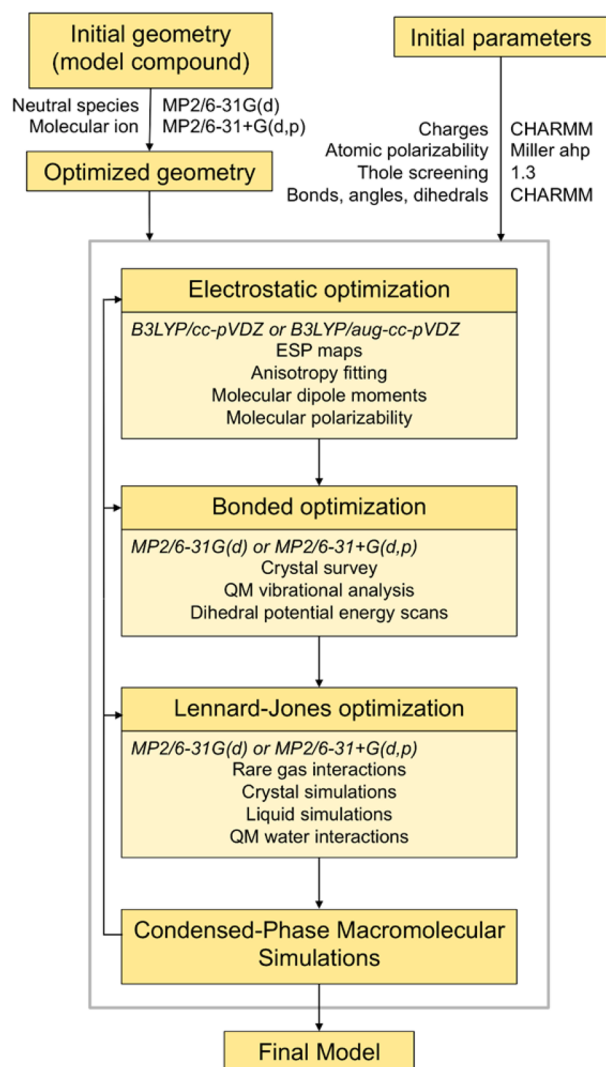


**Figure 2.** Overview of the parametrization protocol for the Drude polarizable force field.

observables for parameter validation and refinement. All QM target data are generated using structures optimized at the MP2/6-31G(d) level of theory for neutral molecules and MP2/6-31+G(d,p) for molecular ions.[121−124] The fitting of electrostatic parameters has been described in detail previously,[104,125] but a brief discussion is appropriate here. In contrast to the electrostatic component of additive force fields that requires only partial charges to be fit against target data, the Drude polarizable model requires that the total charge on the Drude−atom pair, atomic polarizability ($\alpha$, eq 2), and Thole screening factors ($a$, eq 7) all be parametrized. Internal (bonded) terms are parametrized according to crystal survey data and QM vibrational calculations (for bonds, angles, selected dihedrals, and improper dihedrals) and 1-D dihedral potential energy scans (in the case of proper dihedrals) to obtain correct conformational energetics. Potential energy surfaces as a function of backbone ($\phi$,$\psi$) dihedrals are obtained for model dipeptides to parametrize CMAP terms (see below).

Partial atomic charges in additive force fields are often assigned by targeting the QM electrostatic potential (ESP) around the molecule of interest. In the parametrization of the Drude-2013 force field, in addition to this unperturbed ESP map, the polarization response is calculated by generating

multiple ESP maps in the presence of individual perturbing point charges with a magnitude of +0.5 $e$. The QM calculations of all ESP are performed on the MP2-optimized geometries using the B3LYP hybrid functional[126−128] with cc-pVDZ or aug-cc-pVDZ basis sets.[129] B3LYP/cc-pVDZ systematically underestimates dipole moments relative to experimental values, whereas B3LYP/aug-cc-pVDZ yields better agreement and thus is preferred. For large molecules, using the augmented basis set may be too computationally demanding, and for this reason the B3LYP/cc-pVDZ values can be used in conjunction with a well-established value (1/0.83) for scaling of the initial atomic polarizability values; calculations done with the aug-cc-pVDZ basis set require no initial scaling.[125] Following fitting of electrostatic parameters, a further scaling factor of 0.85 is generally applied to the fitted atomic polarizabilities. This value represents an approximate mean obtained from the empirical optimization of the scale factor targeting dielectric constants for a range of small molecules (see section 3.3.3 below) and is introduced to account for overlap of electron clouds in the condensed phase that opposes induction[130−135] and/or as a compensation for an inhomogeneous electric field resulting from the excluded volume of water around the solute.[136] The postfitting scaling factors are not universal, as will be described in the following sections, but remain an important guideline and consideration in the development of the Drude polarizable force field. Scaling factors for specific types of functional groups are discussed in section 3.3.3 below.

The perturbed QM ESP maps become target data for a fitting function that is described in detail elsewhere.[125] During fitting, charge penalties are introduced to prevent the fitting function from yielding charges with poor chemical significance, thereby favoring "chemically intuitive" results; the restraint is also necessary because the charge-fitting problem is underdetermined. This approach is known as restrained electrostatic potential (RESP) fitting.[137] For ESP calculations, rather than constructing a cubic grid around the molecule, which is computationally expensive and suffers from nonunique axes, or a random grid, which impairs reproducibility, a Connolly surface[138] is generated around the molecule being parametrized onto which grid points are placed. The advantages to this approach are that there is shape complementarity to the molecule of interest and the total number of grid points is reduced relative to a cubic grid. Since the surface is generated as a function of the atomic radii, multiple nonoverlapping surfaces can be generated at increasing distance from the atomic centers by multiplying the radii by size factors, $f$ (Figure 3). Electrostatic parameter fitting in the Drude-2013 force field relies on these multiple surfaces, which are composed of either perturbing charges or ESP grid points. In practice, five such surfaces are generated; multiple perturbing charge layers are needed to account for the orientation dependence of the polarizability, and multiple ESP surfaces are needed to calculate the polarization response and the molecular dipole moment. The innermost layer ($f$ = 2.2) consists of perturbing charges at typical hydrogen-bonding distance and geometry, with additional charges interspersed at vertices between these points. The next layer ($f$ = 3.0) is an ESP surface to calculate the response to these charges. To account for the orientational dependence of the molecular polarizability, an additional layer of charges ($f$ = 4.0) and another ESP surface ($f$ = 5.0) are subsequently added. The final layer is placed at the longest distance ($f$ = 6.0) from molecule, to resolve the molecular dipole moment.
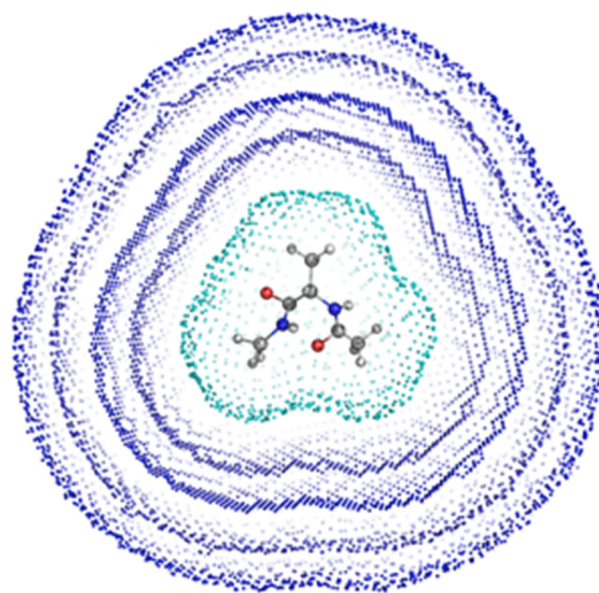


**Figure 3.** Cross sections of Connolly surfaces at increasing $f$-values around the alanine dipeptide.

Following initial generation of the charges, polarizabilities, and Thole factors, the anisotropic polarizabilities are optimized, typically only for hydrogen-bond acceptors. The anisotropic polarization tensor is fit to ESP calculated at the B3LYP/aug-cc-pVDZ level of theory in the presence of perturbing charges of +0.5 $e$ along in- and out-of-plane arcs typically around the acceptor atom.[104] These arcs are typically 2 Å from the acceptor atom, spanning 90°−180° with grid points every 15°.

After optimization of the electrostatic parameters, the internal (bonded) parameters are optimized by targeting both QM and experimental data. QM vibrational analysis is performed to determine force constants for internal parameters, supplemented with dihedral potential energy scans carried out for flexible torsions being parametrized. Equilibrium bond lengths and valence angles target available crystal survey data for model compounds. In addition, for protein ($\phi,\psi$) conformational properties, a high-level QM 2-D ($\phi,\psi$) potential energy surface is obtained for calculation of the CMAP potential.[42] QM potential energy surfaces for dipeptides are obtained at the MP2/6-311G(d,p)//RIMP2/CBS model chemistry[139] and serve as additional target data for protein force field parametrization. The CMAP corrections to ($\phi,\psi$) dihedral cross-terms are a grid-based correction map that are an extension of the force field to obtain better agreement in the energy landscape. In the Drude-2013 protein force field, the alanine dipeptide CMAP is applied to all amino acids, except glycine and proline, for which specific CMAPs are determined on the basis of the respective dipeptides.

LJ parameters are refined and validated using condensed-phase simulations. These simulations generally target crystal data (molecular volume, $V_m$; lattice parameters; and heat of sublimation, $\Delta H_{sub}$) and properties of neat liquids ($V_m$ and heat of vaporization, $\Delta H_{vap}$). Additional target data include QM water interaction energies,[34,140] as well as free energies of hydration. Optimization of the LJ parameters typically requires multiple iterations and represents one of the most difficult aspects of force field parametrization.

The final step toward the finished model consists of condensed-phase simulations of macromolecules. As will be

shown below, combining the model compound parameters to construct polymers is nontrivial in the case of the Drude polarizable force field. Additional effects (particularly dipole–dipole interactions) may present themselves only in macromolecules, such that multiple electrostatic models must be considered, and additional dihedral and CMAP refinements are often required. The entire refinement protocol is iterative, as implied by Figure 2, such that adjustments at any stage of the protocol require validation in all other steps, until sufficient agreement is achieved for all target data.

**3.3.2. Development of Polarizable Water Models.** As most systems of biological interest are immersed in water, the natural starting point for the creation of any force field is the development of a suitable water model. A central challenge in developing an accurate water model is balancing structural, thermodynamic, and dynamic properties.

Several polarizable water models have been developed for use with the Drude-2013 force field, the first of which was a four-site model called SWM4-DP,[13] featuring a rigid configuration of HOH atoms, a negatively charged virtual site ("M site"), and a Drude oscillator attached to the O atom. The M site is placed such that the gas-phase dipole moment and quadrupole moments are correctly reproduced. The Drude oscillator and its parent O atom have charges that are equal in magnitude but opposite in sign such that the total charge of the oxygen is zero. In SWM4-DP ("simple water model with Drude polarization"), the partial charge on the Drude oscillator is positive. Following the development of SWM4-DP, a newer model was introduced, called SWM4-NDP (Figure 4),[141] with the N denoting a
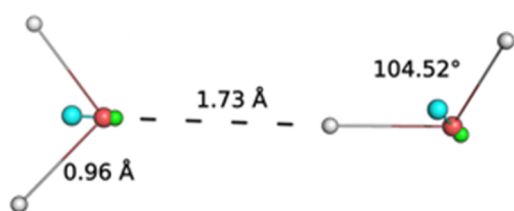


**Figure 4.** SWM4-NDP dimer, extracted from a simulation of bulk water, with intermolecular and intramolecular geometric properties labeled. Atoms are colored by element (oxygen = red, hydrogen = white), the M sites are in cyan, and the Drude oscillators are green. The oxygen–Drude displacements are 0.09 Å, demonstrating the polarization response along the hydrogen bond.

negatively charged Drude oscillator (hence "simple water model with negative Drude polarization"). Given that the Drude oscillator is intended to represent electronic degrees of freedom, assigning it a negative charge is more intuitive and may be better suited to reproducing higher-order polarization effects. As such, this convention of assigning a negative charge to all Drude oscillators has been used throughout the development of the Drude-2013 force field.

Both SWM4-DP and SWM4-NDP were parametrized to reproduce a variety of gas- and condensed-phase properties. Unlike fixed-charge water models, which must inherently make compromises between properties in gas and condensed phases, the explicit inclusion of polarizability allows for better response of the polarizable models to changes in environment and therefore better reproduction of thermodynamic and transport properties. The dipole moment of SWM4-DP and SWM4-NDP serves as an example of this phenomenon. Whereas an isolated water molecule reproduces the experimental dipole moment of 1.85 D, transfer to the liquid state induces an increase in the

dipole moment to a value of 2.46 D, in good agreement with experimental results for water in a hydrogen-bonded environment.[142] The SWM4-NDP dipole moment is also sensitive to the presence of charged groups (Figure 5). In contrast, the
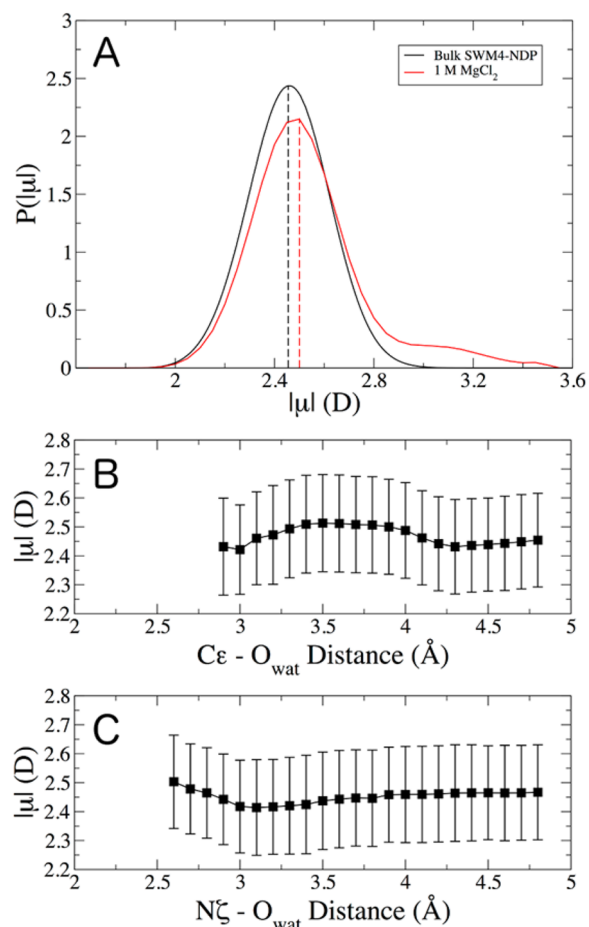


**Figure 5.** Response of SWM4-NDP water dipole moments to charged groups. (A) Distributions of water dipole moments in bulk water ($\langle|\mu|\rangle$ = 2.46 D) and in a 1 M solution of $MgCl_2$ ($\langle|\mu|\rangle$ = 2.53 D). Averages are indicated by vertical, dashed lines. Dipole moments as a function of distance from (B) glutamate-64 and (C) lysine-11 in ubiquitin. Error bars in panels B and C are the root-mean-square fluctuations over the simulation.

commonly used TIP3P water model[143] has a fixed dipole moment of 2.347 D, irrespective of environment. Other properties, such as the change in internal energy upon liquefaction, static dielectric constant, and molecular volume, are also more accurately reproduced using the SWM4-DP and SWM4-NDP models relative to TIP3P.[141] Importantly, the polarizable models also reproduce the self-diffusion constant and viscosity of water much more accurately than fixed-charge models, indicating that the time scales of biomolecular phenomena may be better represented using polarizable models. We note that the water models yield self-diffusion constants that are slightly overestimated when the Yeh and Hummer correction for periodicity[144] is taken into account due to the models being developed prior to publication of that correction. Though SWM4-DP and SWM4-NDP were parametrized using similar methods, and produce similar physical properties, the SWM4-NDP model is preferred because of its intuitive connection between the negatively charged Drude
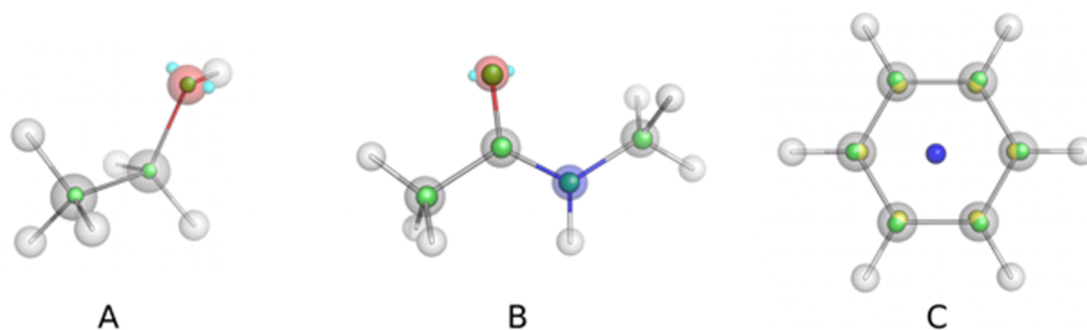
**Figure 6.** Examples of model compounds used in the development of the Drude polarizable force field: (A) ethanol, (B) *N*-methylacetamide, and (C) benzene. Atoms are colored by element (O = red, H = white, C = gray, N = blue). Lone pairs attached to atoms treated with anisotropic polarization in panels A and B are cyan. Drude oscillators in panels A and B are shown in green. In the case of benzene in panel C, a perturbing charge of +0.5 *e* (dark blue sphere) is placed below the center-of-mass of the ring to illustrate the response of the Drude oscillators (colored green in the unperturbed system and yellow when perturbed). The Drude oscillators move toward the center of the benzene ring, and also out-of-plane, toward the perturbing charge.

oscillators and electronic degrees of freedom, and it is the water model that has been used for the development of the remainder of the Drude polarizable force field.

It is also worth noting a more recent polarizable water model in this series, known as SWM6.[145] Like SWM4-NDP, SWM6 includes a negatively charged Drude oscillator and an M site, but also adds two lone pairs out of the H−O−H plane. The addition of these two lone pairs and subsequent reoptimization of the electrostatic parameters of the existing SWM4-NDP model were an effort to minimally perturb the properties of SWM4-NDP while achieving better agreement with structural properties of water,[145] which would better capture both microscopic and bulk behavior. The SWM6 model better reproduces the structure and dipole moment of a water dimer, while preserving or improving upon the dynamic and thermodynamic properties of SWM4-NDP. Both SWM4-NDP and SWM6 capture the polarization effects in water and yield better agreement with a wide variety of physical properties than additive water models commonly used in most MD simulations, like TIP3P, TIP4P, and TIP5P,[143,146] with an exception being the temperature-dependence of the density of water.

**3.3.3. Polarizable Organic Compounds.** The development of any force field is generally approached in a progressive manner; that is, parameters are initially optimized by targeting small model compounds (see examples in Figure 6), with those parameters acting as the foundation for the development of parameters for larger entities, such as amino acids, nucleotides, and ultimately proteins and nucleic acids. In this section, the development of Drude polarizable models for organic compounds that are the basis for the protein, nucleic acid, lipid, and carbohydrate aspects of the force field will be discussed in detail, with specific emphasis on target data used for optimization of the nonbonded parameters and the role of explicit polarization in the accuracy of the physical model. In general, target data include QM (gas-phase) properties, such as molecular dipole and quadrupole moments, water interactions (energies and distances), and dihedral energy scans, as well as condensed-phase properties with experimental target data, such as diffusion coefficients, heats of vaporization ($\Delta H_{vap}$) for liquids, heats of sublimation ($\Delta H_{sub}$) for crystals, molecular volumes ($V_m$) for liquids and crystals, free energies of hydration ($\Delta G_{hydr}$), and static dielectric constants ($\varepsilon$) of neat liquids. The dielectric constant is a crucial consideration when evaluating

parameter sets, as it largely dictates the extent to which gas-phase polarizabilities are scaled, as mentioned above. In practice, the scaling factor becomes a freely tunable parameter during optimization, as will be discussed below.

*3.3.3.1. Alkanes.* Aliphatic groups are a major component of biological macromolecules, including hydrophobic lipid tails, amino acid side chains, and the backbone of carbohydrates. Alkanes are a fundamental component of any force field, but additive models suffer from an inability to respond to a high-frequency oscillating field that gives rise to the optical dielectric constant, $\varepsilon_{inf}$, according to the Clausius−Mossotti equation.[147] As a consequence, additive models lead to a systematic underestimation of alkane dielectric constants, as low as 1. While the dielectric constants of alkanes are very low, this underestimation can have important consequences for free energies of solvation, which scale with $(1 − 1/\varepsilon)$ according to the Born approximation.[148] Given that electronic polarization is central to the dielectric response in low-polarity media like alkanes, its inclusion in the functional form is likely to improve these physical properties. For this reason, the next logical step toward a fully functional polarizable force field was the development of a polarizable model for alkanes.[149]

These alkane parameters were optimized according to the method outlined in section 3.3.1, with two minor modifications. During parameter fitting, charges tended to accumulate asymmetrically on carbon atoms. While this outcome is not without physical basis, it represents a conformation-dependent solution to the fitting algorithm and reduces transferability between compounds. Given that computing ESP maps for a large number of rotamers of long alkanes was not computationally tractable, a restraint was imposed on the fitting algorithm such that the charge of any carbon atom ($q_C$) was restricted to $q_C = −xq_H$, where $x$ is the number of hydrogen atoms (each carrying a charge $q_H$) bonded to the carbon in the $CH_x$ group, which by convention is electroneutral. Final parameters for $CH_x$ groups were obtained by averaging over several alkanes in multiple rotameric states. The second modification of the fitting protocol concerns the scaling of atomic polarizability values. The general protocol described above prescribes a scaling factor of 0.85 to be applied if the ESP maps are calculated at the B3LYP/aug-cc-pVDZ level of theory. In the case of alkanes, applying this scaling led to poor agreement between modeled and experimental observables (i.e., with the alkanes a scale factor of 1 was applied), indicating that

a single scaling factor for all classes of molecules is not appropriate (see Table 1).[149]

**Table 1. Values of Atomic Polarizability Scaling Factors for Small Molecules in the Drude-2013 Force Field**

| model compound | atomic polarizability scale factor |
| --- | --- |
| alkanes | 1.0 |
| aromatics | 0.724 |
| heteroaromatics | 0.85 |
| alcohols | 0.70 (C), 1.0 (O) |
| SWM4-NDP water | 0.724 |
| ethers | 0.85 |
| thiols | 0.70 |
| disulfide | 0.85 |
| ethyl methyl sulfide | 0.60 |
| amides | 0.70 |

Having developed a transferable set of parameters for ethane, propane, butane, isobutene, and pentane, the force field was assessed in terms of thermodynamic, structural, and dynamic properties. Specifically, the condensed-phase behavior of these alkanes was evaluated in terms of $\Delta H_{vap}$, $V_m$, static dielectric constants, and $\Delta G_{hydr}$. Notably, the polarizable model improved upon $\Delta H_{vap}$ values substantially, achieving agreement within 2% of experimental data, whereas the additive models underestimated $\Delta H_{vap}$ by up to 18%. The enhanced interaction due to electronic polarizability resulted in a slight underestimation of diffusion constants, but the magnitude of error was no worse than the CHARMM additive model, and led to better agreement with NMR relaxation times, which arise from rotational correlation times of C−H bond vectors. Most importantly, the Drude polarizable alkane model produced significantly better agreement with experimental static dielectric constants. Whereas the additive model produced nearly uniform values of 1.0, irrespective of alkyl chain length (in the series of ethane, propane, butane, isobutene, heptane, and decane), the Drude polarizable model produced values in excellent agreement with experimental data, ranging from 1.71 to 2.13. Thus, the polarizable alkane model represented a significant improvement over the existing additive model, and explicit polarization was critical in obtaining good agreement with a wide range of condensed-phase behavior, while still retaining good agreement with QM conformational energetics.[149] These results have implications for the accurate modeling of hydrophobic environments, such as lipid bilayers and the interiors of folded proteins (see section 4.1.2).

*3.3.3.2. Aromatics.* Following the development of the Drude alkane model, parameters for aromatic compounds were developed using benzene (Figure 5C) as a model compound.[150] Transferability of parameters was evaluated by combining benzene and methyl parameters from the Drude alkane force field[149] to create a model for toluene. These compounds form the basis for aromatic amino acids such as phenylalanine and tyrosine. While many of the structural and dynamic features of the benzene and toluene liquids were comparable between the additive and polarizable force fields, the Drude model produced more accurate quadrupole moments and dielectric constants. The additive models of benzene and toluene showed similar behavior to the alkanes described above, producing nearly uniform $\varepsilon$ values of 1.0, considerably lower than the experimental values of 2.3 (benzene) and 2.4 (toluene); the Drude polarizable model

(applying an atomic polarizability scale factor of 0.724) yielded better agreement.[150] Of particular importance to biomolecular applications of these aromatic parameters is the interaction of water molecules with aromatic groups. Given the electronic structure of aromatic rings, the $\pi$ electron system gives rise to an accumulation of negative charge above and below the ring with partial positive charge accumulating at the edges. This electronic structure allows for hydrogen bonding with water.[151] The Drude polarizable model produced a better-defined population of water molecules in the first peak of the radial distribution function around the ring at 3.5 Å, indicative of more structured interactions that reflect this hydrogen bonding. Subtle differences in hydration shells around benzene were evident, suggestive of fundamental differences in the atomic interactions in additive versus polarizable systems. A proper description of these subtle effects is an important component of the polarizable model, as it produces a more realistic representation of hydration around these biologically important moieties. In this regard, the Drude oscillator model is particularly useful, as other methods, like fluctuating charge, will not capture out-of-plane polarization, as charge can only be redistributed among atoms in the planar ring system. More recently, the benzene model was extended to yield improved agreement with respect to QM data for cation−$\pi$ interactions by inclusion of a virtual particle in the center of the benzene ring that uses additional LJ interactions to optimize the interaction orientation and energy.[152]

*3.3.3.3. Linear and Cyclic Ethers.* The development of polarizable force fields for carbohydrates and nucleic acids rely, in part, upon the development of accurate parameters for cyclic ethers, which form the backbone of ribose and deoxyribose rings, as well as the backbones of the wider classes of furanoses and pyranoses. To this end, parametrization of a Drude polarizable model for linear and cyclic ethers was undertaken.[153] Ethers, while generally of low polarity, have the ability to participate in hydrogen bonding and ion coordination via their oxygen atoms, but their alkyl substituents are nonpolar. Thus, a central challenge in developing an accurate force field for ethers is the balance between electrostatic, dispersion, and repulsive forces.

The original electrostatic parametrization method of Anisimov et al.[125] was expanded to include a total of eight grids around each of the ethers with perturbing charges and an ESP map at $f = 1.3$, as well as another ESP map at $f = 2.2$ in addition to the five surfaces described above. These additional surfaces correspond to hydrogen-bonding geometries (hydrogen and heavy atoms, respectively) and were necessary to properly account for polarization response upon hydrogen-bond formation. Parametrization initially focused on two cyclic ethers, tetrahydrofuran and tetrahydropyran, which form the backbone of furanose and hexopyranose ring systems found in carbohydrates. In addition, cyclopentane and cyclohexane parameters were developed to extend the linear alkane series to these cyclic species. Following the goal of transferability, parameters for all $CH_x$ groups not directly bonded to the oxygen of the ether were transferred from the alkane force field,[149] and the parameters developed for cyclic ethers were subsequently transferred to a series of linear molecules. Polarizability scaling was applied to all species developed in this series of parameters, including cyclic alkanes. Whereas such scaling was not necessary in the development of the linear alkane series, the scaling here reflects the complex electronic

properties, likely due to correlation effects, which lead to cyclic species behaving more like polar compounds.[153]

A key outcome in the development of the Drude polarizable force field for ethers is the simultaneous agreement in relative energies of various conformations and their dipole moments. Several rotamers of linear species, as well as different conformers for the cyclic species, were assessed. While the additive model reasonably reproduced the relative conformational energies, the dipole moments of these conformers were generally too large. The polarizable model was capable of producing agreement with both quantities. This outcome reflects the ability of the polarizable model to accurately model the complex electronic structure of ring systems as well as the response of linear species to rotation about various covalent bonds. The implications for biomolecular force fields are clear; the conformational sampling of furanose and hexopyranose sugars is important to the stability and dynamics of nucleic acids and glycosylated proteins.[154,155] Other condensed-phase properties were reasonably reproduced by the additive model, indicating that the addition of polarizability in the context of ethers did not necessarily improve certain aspects of the physical model.[153] This limitation is likely due to the desire for transferability of parameters between molecules; the cyclic molecules represented a distinct improvement as a consequence of including explicit polarization, but linear species did not improve. As a result, the ether force field was subsequently revised in 2010,[156] using a parametrization scheme that allowed for independently tuning LJ parameters on linear and cyclic ether oxygen atoms, as well as carbon atoms bonded to ether oxygen atoms. Moreover, atom-specific Thole factors were employed. Previous parametrization efforts assumed that the value of $a$ in eq 7 should be held constant at a value of 2.6. In the refinement of the ethers, the parametrization protocol tuned the value of individual atomic Thole screening factors, an approach that was introduced on the basis of the need to improve the molecular polarizability tensor of the amides, as discussed below. The application of a 0.7 scaling factor for ether gas-phase polarizabilities led to a systematic underestimation of liquid dielectric constants; the updated ether parameter set applied a scaling factor of 0.85, which achieved significantly better agreement.[156] This updated ether parameter set, with rescaled atomic polarizabilities and atom-specific Thole factors, forms the basis of the carbohydrate force field described below.

*3.3.3.4. Alcohols.* Given the ubiquity of alcohol groups in proteins (amino acids like serine, threonine, and tyrosine), nucleic acids (2'- and 3'-hydroxyl groups), carbohydrates, and lipids, these functional groups are a major component of any force field. Properties of alcohols are interesting not only due to this prevalence, but because their hydration is dictated by competition between polar and nonpolar interactions. Thus, a proper description of alcohol interactions with various media will certainly require electronic polarization.

The first application study of a polarizable alcohol using the Drude oscillator model was conducted by Noskov et al.[106] Results from that study produced excellent agreement with solution properties (such as dielectric constant and diffusion of water and ethanol) and provided important insight into ethanol hydration properties. Supporting the hypothesis that water structure is affected by ethanol, the authors found that water–water hydrogen bonding was depleted in the vicinity of the ethanol hydroxyl group but that the dominant effect in water perturbation was due to restructuring of the water molecules in the second hydration shell around ethanol.[106]

Given the insights that arose from the work of Noskov et al., a more generalized parameter set for polarizable alcohols was subsequently produced to encompass a larger series of primary and secondary alcohols two years later,[157] including a more refined model of ethanol that included lone pairs on the hydroxyl oxygen atom (Figure 5A). Parametrization of electrostatic terms was carried out in the same manner as for the ethers described above. The parameters that were developed were further evaluated against QM water interaction energies and geometries. The inclusion of explicit polarization led to significantly better agreement with QM water interactions. In the final parametrization of the model, atomic polarizabilities were selectively scaled; to preserve the quality of the hydroxyl–water interactions (as well as other properties discussed below), the polarizability of the hydroxyl oxygen atom was not scaled, but $\alpha$ values of carbon atoms in the alcohols were scaled by 0.7. Off-diagonal terms (pair-specific LJ parameters that override normal Lorentz–Berthelot combination rules, called NBFIX in CHARMM) were introduced for the interaction of primary and secondary alcohol oxygen atoms with water oxygen atoms. While this extension of the normal combination rules may limit transferability between molecules, it was necessary to simultaneously achieve good agreement with experimental properties and QM target data. The final model was evaluated in training set molecules as well as longer-chain alcohols that were not part of the training set (2-butanol, 1-propanol, 1-butanol). The polarizable model was found to be a significant improvement over the additive CHARMM force field for all quantities. Most notably, the polarizable alcohol series improved $\Delta H_{vap}$ and $\Delta G_{hydr}$, among quantities, and reduced the average error in static dielectric constants from −35.9% in the additive model to −2.3% in the Drude model. Given the similarity in internal terms between these two force fields, the polarization response is clearly responsible for the improvement in agreement of these values, an outcome that reflects the improved treatment of nonbonded terms. While the additive force field was intrinsically overpolarized to mimic condensed-phase conditions, this approximation is ultimately inadequate, and the Drude model showed another critical feature, the response of the molecular dipole moment to environment. Whereas the additive CHARMM force field for alcohols produced dipole moments that were largely invariant between gas, neat alcohol, and aqueous systems, the Drude model responded dramatically to these different environments, shifting from low dipole moments in the gas phase (values in agreement with QM calculations) and, when solvated by benzene, to much higher values in aqueous solution, in agreement with previous theoretical calculations.[158] As a final point, the Drude alcohol parameter set influenced the dipole moment of water molecules hydrating the alcohols. Small variations in water dipole moment were noted as a function of distance around ethanol as well as hydrogen-bonding geometry (with water serving as either a donor or acceptor), suggesting that mutual polarization between ethanol and water is an important factor in dictating the interactions in aqueous systems. The ability to model these subtle effects in the polarizable alcohol series reflects a distinct advantage over additive force fields, in which rigid water models and nonpolarizable alcohols are incapable of responding to such effects, which may have implications for the dynamics of the many biomolecules that contain hydroxyl groups.

The Drude model of methanol was recently examined in the context of Kirkwood–Buff integrals of methanol in aqueous

solution[159] using the SWM4-NDP model[141] of water. By assessing solutions of varying methanol concentration (from pure methanol to pure water), the Kirkwood−Buff theory relates the microscopic details of the solution structure with thermodynamic (macroscopic) properties. While the Drude model of methanol resulted in better density in the neat liquid and increasing water self-association in aqueous methanol solution, ultimately there was little difference between the additive CHARMM and Drude polarizable models of methanol with respect to experimental Kirkwood−Buff integrals, excess coordination numbers, activity derivatives, and excess molar Gibbs energy. This outcome suggests that additional improvements can be made to the Drude methanol model (and perhaps alcohols in general) to even better reproduce the balance of forces between alcohols and water. Future work will also assess the quality of alcohol interactions with the SWM6 water model,[145] which, like the alcohols, includes lone pairs on the oxygen atom and consequently may achieve a better balance of intermolecular interactions.

*3.3.3.5. Amides.* Given that protein backbones are polyamides and two amino acid (asparagine and glutamine) side chains contain amides, an accurate model for amide groups is an essential part of a polarizable force field for proteins. The balance of forces between amides and water is an important determinant of protein conformational dynamics.[160] To this end, efforts to parametrize amide-containing model compounds were undertaken.[161] The parametrization was carried out as described above, and it was during the optimization of the amides that the use of atom-based Thole screening factor was introduced in the parameter optimization, as was used in the refinement of the ether parameters.[156] Using atom-specific Thole screening factors is justified due to the electronic resonance in chemical groups such as amides, for which electronic response along different axes may be different, and inclusion of this term leads to significant improvements in the treatment of molecular polarizability tensors. Notably, the final, optimized amide parameter set directly used the gas-phase atomic polarizabilities to reproduce the high dielectric of neat *N*-methylacetamide (NMA, 100 at 373 K).

Additive models of NMA (Figure 6B) with the CHARMM force field and formamide with OPLS-AA had previously been found to underestimate neat liquid $\varepsilon$ values by 70% and 50%, respectively,[162,163] indicating that the mean-field approximation inherent in additive models is an inadequate representation of these species. The Drude model for NMA reproduced a wide range of gas-phase QM and condensed-phase experimental data and was the first force field to yield dielectric values in near-quantitative agreement with experimental values, including temperature dependence. Cooperativity among dipole moments (head-to-tail alignment of dipoles such that their magnitudes are enhanced along the dipole vector) between hydrogen-bonded NMA molecules explains the physical behavior, a property that required accurate treatment of the molecular polarizability tensor by using atom-based Thole factors, as described above. In the Drude model, the Kirkwood $G_K$ factor was considerably larger (4.6 ± 0.6) than the CHARMM additive model (3.0 ± 0.2), indicating a greater degree of cooperative dipole alignment. The decrease in dielectric as a function of temperature is rationalized in the same way. Individual NMA molecules experience a minimal (∼0.1 D) decrease in dipole moment, but the value of $\varepsilon$ drops considerably. While the additive model experiences a small decrease in dielectric, its values are well below those observed

experimentally, but the Drude model reproduces experimental behavior nearly quantitatively. This result is a function of polarization response. The elevated temperature causes frequent breaking of hydrogen bonds, thus disrupting cooperativity among dipole moments and ultimately a decreased value of dielectric.

As a final note on dielectric values, the Drude polarizable model correctly reproduced the paradoxical trend among methylated amides. Rather than increasing as a function of hydrogen-bonding capacity, the $\varepsilon$ values at 373 K for acetamide, NMA, and *N,N*-dimethylacetamide are 66.3, 100, and 26, respectively. The $\varepsilon$ values of NMA and *N,N*-dimethylacetamide can be rationalized intuitively by differences in hydrogen-bonding capacity, but the value of acetamide is somewhat surprising, as it can participate in three hydrogen bonds per molecule, while NMA can participate in two. However, simulations with the Drude model revealed that hydrogen bonding to the *cis* amide proton results in anticorrelated dipole moments, opposing the dipole enhancement via the *trans* proton hydrogen bonding, ultimately leading to an intermediate dielectric value. This work on polarizable amides[161] represented an important advancement toward the construction of a polarizable protein force field, as NMA is a model of the polypeptide backbone. The ability of the Drude NMA parameters to reproduce QM water interactions, $\Delta G_{hydr}$, and dipole cooperativity has important implications for simulations of proteins, as will be discussed later.

Building on this work, which involved primarily simulations of neat acetamide liquid simulations, the properties of the acetamide series in aqueous solution using the Drude polarizable force field have recently been assessed in greater detail.[164] This work resulted in updated NMA parameters based on an initial reparametrization of internal terms to reflect polypeptide geometric data taken from crystal surveys, with subsequent reoptimization of nonbonded (electrostatic and LJ) terms. This newer model of NMA serves as the basis for the Drude polarizable protein force field that will be described below. Unlike the previous model of NMA that reproduced neat liquid properties well, gas-phase polarizabilities were scaled by a factor of 0.7. The resulting model reproduced many experimental quantities over a range of NMA and acetamide concentrations in water, including Kirkwood−Buff integrals,[164] water coordination numbers, solution density, partial molar volumes, and molar enthalpy of mixing. These outcomes indicate that the newer Drude polarizable model for amides accurately balances solute−solvent, solute−solute, and solvent−solvent interactions. Such balance is a crucial component of an accurate biomolecular force field, since protein conformational equilibria depend, in part, on the balance of intramolecular and protein−solvent interactions.[160]

*3.3.3.6. Nitrogen-Containing Heteroaromatics.* Several biologically important compounds are comprised of nitrogen-containing heterocyclic aromatic moieties, such as amino acids histidine and tryptophan, all of the bases in DNA and RNA, and several enzyme cofactors, and they occur in a wide range of medicinal compounds. Accordingly, a series of heterocyclic aromatic compounds (pyrrole, imidazole, pyridine, pyrimidine, indole, and purine) were parametrized to create the building blocks for these biologically relevant molecules.[165] The heteroaromatic series was developed much in the same manner as benzene and toluene described above and targeted many of the same experimental data, including neat liquid and crystal simulations. For heteroaromatics, the gas-phase polarizabilities

were scaled by 0.85 to achieve good agreement for pyridine and pyrrole dielectric constants. Whereas both scaled and unscaled polarizabilities yielded good agreement for pyridine, the unscaled parameters for pyrrole overestimated $\varepsilon$, necessitating the scaling. Achieving agreement with the liquid dielectric constant is the guiding principle for applying polarizability scaling, and the scaling factor can be an additional variable in parameter fitting.

Prior to the development of the nitrogen-containing heterocycles, lone pairs were only applied to carbonyl oxygen atoms that were treated anisotropically.[104] In the case of the heterocycles, virtual sites were added to the nitrogen atoms to represent the in-plane lone pairs of these atoms. The addition of these virtual sites improves agreement with the QM ESP maps and is chemically intuitive, given the nature of these compounds. Additional details of the nitrogen-containing heterocycles will be discussed below in the context of nucleic acid parametrization.

*3.3.3.7. Sulfur-Containing Compounds.* Thiols and sulfides feature prominently in many chemical and biological systems, particularly in proteins (methionine and cysteine) as well as in medicinal chemistry applications. The sulfur atom is polar and strongly polarizable; thus, an accurate representation of these chemical groups containing sulfurs is an important component of any force field. Parameters for sulfur-containing compounds using the Drude oscillator model were derived by Zhu and MacKerell in 2010.[166] Similar to the polarizable force field for alcohols,[157] the polarization of sulfur atoms was treated anisotropically, yielding good agreement with orientation-dependent QM water and ion interaction energies and a considerable improvement over the additive CHARMM22 parameters for these same compounds. Given the sensitivity of sulfur to local electric fields, polarizability scaling is a particularly important consideration. It was found that a scaling factor of 0.7 for thiols led to good agreement with experimental dielectric constants, while 0.85 was applied to dimethyl disulfide. Upon further investigation into ethyl methyl sulfide, it was determined that a scaling factor of 0.6 was more suitable for reproducing its condensed-phase properties and gas-phase dipole moment, and these are the parameters that are used in biological moieties such as those found in methionine. As in the case of alcohols, off-diagonal LJ terms (NBFIX) between sulfur and water oxygen atoms were necessary to simultaneously achieve agreement between the properties of the neat thiol and disulfide liquids and $\Delta G_{hydr}$.

*3.3.3.8. Summary of Parametrization for Organic Compounds.* The assembled body of model compounds that has been derived serves as the basis of the biomolecular Drude polarizable force field that will be described in the following sections. The judicious choice of target data, from both QM calculations and condensed-phase experimental data, and iterative refinement protocols that have been used for many years in the development of the additive CHARMM force field have shown that parameters are transferable across molecules. Freedom to tune Thole screening factors on a per-atom basis, the use of atom-pair-specific LJ parameters, and scaling of gas-phase polarizabilities based on the nature of the model compounds allow for robust models of a variety of molecules. The inclusion of electronic polarizability provides unique insight into microscopic properties at the small molecule level, a feature that extends to biological macromolecules, as will be shown below.

**3.3.4. Polarizable Ions.** Monatomic ions are prevalent throughout the aqueous cellular milieu and play important roles in neuronal excitation, folding and structural stability of proteins and nucleic acids, and catalysis.[167−172] Ions can be strongly polarizing, a property that can dramatically influence their solvation and interactions with biomolecules[173] but one that is impossible to model in additive force fields. The first set of ion parameters developed as part of the Drude-2013 force field was a series of monovalent alkali and halide ions, parametrized using the SWM4-DP water model. This ion parameter set became known as AH/SWM4-DP[174] ("AH" for "alkali and halides"). As monovalent ions have very few tunable parameters in the polarizable model (only atomic polarizability and LJ terms), the parametrization strategy employed in that work was to take gas-phase atomic polarizabilities (no scaling for cations, but scaled by a factor of 0.7 for anions as described in section 3.2) and optimize LJ parameters to obtain accurate $\Delta G_{hydr}$. Given that the biological function of ions also relies on correct coordination geometry in solution and when interacting with biomolecules, the optimized parameters were validated against structural properties of ion hydrates and bulk solution using available ab initio and experimental data. The ion parameters produce nearly perfect results for neutral salt $\Delta G_{hydr}$ (which are more reliable than the $\Delta G_{hydr}$ values for individual ions), and the structures of the hydrated alkali ions (which have highly ordered water configurations) agreed well with ab initio geometries and energies. The structural properties of halide hydrates, which have perturbed structures, showed some disagreement with ab initio calculations, but much of the disparity could be attributed to limitations in the water model, SWM4-DP. Further, both structural and thermodynamic halide properties in bulk solvent agreed well with experimental data, indicating that the ion parameters themselves were a reasonable representation of physical reality.

The same general parametrization strategy was employed in the newest version of Drude polarizable ions, which include monovalent and divalent ions.[107] The updated parameters were generated using the preferred SWM4-NDP water model. While the target data were the same as in the previous parametrization, and comparable accuracy was achieved, the newer ion parameters include several important advancements. Notably, divalent ions ($Ca^{2+}$, $Mg^{2+}$, $Zn^{2+}$, and $Sr^{2+}$) were parametrized for the first time. Divalent ions are prone to "polarization catastrophe" due to a Coulombic singularity that can arise when the Drude particle on the ion (which carries no LJ terms) interacts with the core charge on the water oxygen. This problem was circumvented by using a through-space Thole screening of the same form as that used for neighboring (bonded) dipoles described above for selected pairs (eq 7). This algorithm, called NBTHOLE (for "nonbonded Thole") in the CHARMM parameter file, was an optimal solution to the overpolarization problem, as additional restoring forces and/or assignment of LJ terms to Drude oscillators of the ions led to stable simulations, but it was not possible to achieve sufficient agreement with monohydrate energies. The use of the NBTHOLE formalism allowed for both stable bulk simulations and good agreement with ab initio target data. It should be noted that the agreement of monovalent ions in this parameter set is somewhat better than that of the divalent ions, due to charge-transfer effects that cannot be modeled by the Drude polarizable functional form. Despite this limitation, the divalent ion series still reproduces experimental and ab initio target data well, including the hydration structure in bulk solution.[107]

Correct coordination geometry is an important property for biomolecule—ion interactions, especially in $Mg^{2+}$ binding to RNA, which is known to facilitate folding and stabilize complex tertiary structures.[169,170] Finally, the diffusion constants of the monovalent and divalent ion series were assessed and found to be in good agreement with experimental values. The ion parameters thus represent a good physical model for these species in terms of thermodynamic, transport, and gas-phase hydrate properties.

The existing Drude polarizable ion series has been parametrized at infinite dilution; thus, the behavior of the parameters in concentrated ion solutions requires additional exploration. At higher concentration, significant ion—ion interactions can affect dynamics and bulk solution properties, though it is clear that there are interesting physical insights to be obtained, as concentrated salt solutions lead to augmented water dipole moments (Figure 5A). One approach for assessing the quality of ion parameters in solution is the calculation of osmotic pressure, for which a straightforward computational approach has been developed.[175] This method has subsequently been applied in an initial examination of concentrated solutions of Drude polarizable ions[176] and also in the context of balancing interactions among ions, water, and nucleic acid functional groups.[177] These osmotic pressure calculations represent an important advancement in force field parametrization and validation, as the target data are well-defined and are good indicators of the balance between ion—ion, ion—neutral solute, and ion—water interactions, critical components of an accurate force field.

**3.3.5. Peptides and Proteins.** Building upon the small molecule parametrization efforts described above, the first protein force field based on the Drude oscillator was completed in 2013.[178] Emphasis was placed on transferability between compounds during the development of the Drude-2013 force field for small molecules, with the end goal being the ability to combine these parameter sets into a full model for biomolecules. The initial protein model (called "Drude-1") was produced using the simplest means, by combining NMA[161,164] and ethane into a model of the alanine dipeptide, which is a central model in the force field parametrization of proteins. Additional parameter refinement is necessary when creating new connectivity between discrete molecules; to this end, internal terms were optimized, and 2-dimensional cross-term energy correction maps (CMAP)[42] were generated on the basis of high-level gas-phase QM data. The simple Drude-1 model did not produce good agreement with $J$-coupling data for the model $(Ala)_5$ peptide, as extended ($C_5$, Figure 7) conformations were overly favored. This behavior arose because the two neighboring peptide bonds are oriented in opposite directions, such that their local dipole moments cooperatively enhanced each other, thereby maximizing hydrogen bonding with water. These cooperative dipole—dipole interactions also led to peptide conformations that understabilized helical conformations, in favor of elongated conformations. Analysis of dipole moments of $(Ala)_5$ indicated that this behavior was due to improper treatment of the polarization response associated with direct application of electrostatic parameters based on NMA to the polypeptide. As a result, further refinement of this simple model was necessary.

To address the proper polarization response associated with peptide bond—peptide bond cooperative dipole interactions as a function of $\phi$ and $\psi$ values, a model ("Drude-2") was developed by individually targeting ESP maps of the alanine
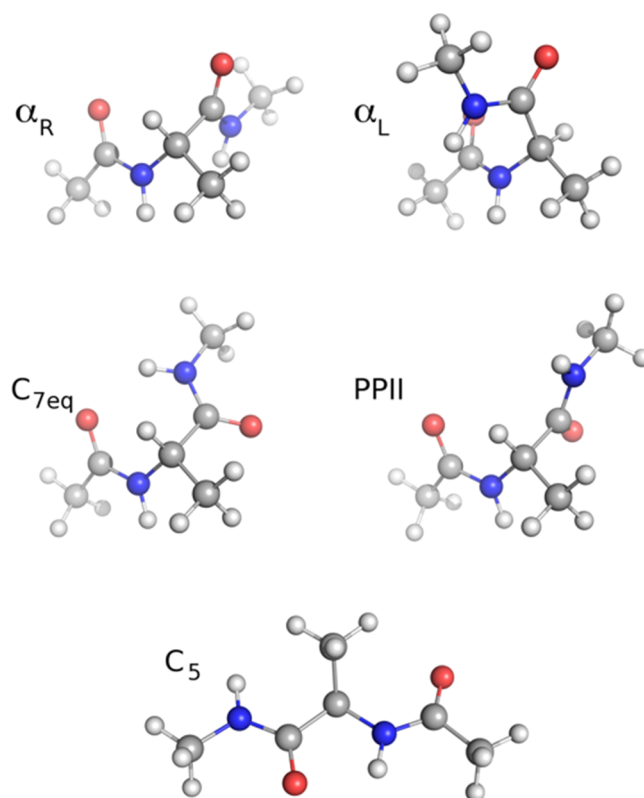


**Figure 7.** Alanine dipeptide conformations used in the electrostatic parameter fitting. The $\alpha_R$, $\alpha_L$, $C_{7eq}$, and PPII conformations are aligned along the $C\alpha$—$C\beta$ bond vector to illustrate the relative positioning of the backbone as a function of $\phi$ and $\psi$.

dipeptide in five different conformations ($\alpha_R$, $\alpha_L$, $C_5$, PPII, and $C_{7eq}$, Figure 7) and taking the average of these parameter sets. The rationale for this approach was to capture the changes in ESP as a result of dipole reorientation concomitant with changes in the relative orientation of the two peptide bonds. While the Drude-2 model yielded better conformational properties relative to Drude-1, it suffered from defects in representing gas-phase QM dipole moments, and the conformational properties of a model $(Ala)_5$ polypeptide in aqueous solution were inadequate; extended conformations were still overpopulated despite a reduction in dipole enhancement relative to Drude-1.

The final model, called "Drude-3" by Lopes et al. and ultimately "Drude-2013" (thus the basis of the name of the force field described in this review), was produced by refining the Drude-2 model with a Monte Carlo/simulated annealing (MC/SA) protocol that targeted a wide variety of target data, including the polarizability and dipole moments of the alanine dipeptide, relative energetics of the $(Ala)_5$ peptide, and QM interaction energies between water and the alanine dipeptide with different backbone conformations. This strategy was motivated to obtain an improved balance among intramolecular electrostatic interactions (from gas-phase target data) and condensed phase properties (water interactions). The MC/SA protocol generates random changes in an $N$-dimensional vector (where $N$ is the number of parameters being fit) to move through parameter space by minimizing an error function calculated as the root-mean-square difference between all QM target data and MM values. During MC/SA, the temperature is periodically scaled down such that it approaches zero. The

temperature is an important component of the algorithm, as increases in the error function are accepted according to the Metropolis criterion to promote exhaustive sampling of parameter space. Targeting this wide range of data simultaneously, while also including the longer (Ala)$_5$, an extension of the typical additive CHARMM parametrization protocol using only the alanine dipeptide, was a critical component in creating a viable polarizable model for the protein backbone. The longer (Ala)$_5$ peptide was utilized to account for the conformational properties that result from communication between neighboring peptide-bond dipoles, including $i$ to $i + 4$ interactions occurring in $\alpha$-helices, a property that was poorly represented by directly using NMA.

The Drude-2013 protein force field was able to successfully model relative energies between different peptide backbone conformations, while also suitably reproducing water interaction energies. The balance between intrapeptide and solvent interactions plays an important role in determining the conformational ensemble of a protein in folded and unfolded states, and the ability of the Drude-2013 force field to capture these effects is an important aspect of its parametrization. To achieve this agreement, gas-phase polarizabilities were scaled by a factor of 0.85, similar to the scaling applied to small molecules described above.

A few additional details of the Drude-2013 protein force field merit discussion here. By convention, the peptide bond group (including the C$\alpha$ and H$\alpha$ atoms) has a net charge of zero. Side-chain moieties (from C$\beta$ onward) also have a unit charge ($-1$, 0, or 1 depending on the amino acid). This convention facilitated combining the small molecule electrostatic parameters into the polypeptide and, notably, allows for straightforward assessment of dipole moments for these groups. These dipole moments have important implications for the properties of the Drude-2013 force field and the physical insights it provides. Compared to the most recent additive CHARMM force field, CHARMM36,[27] the Drude-2013 force field has larger backbone peptide-bond dipole moments. Despite the inherent overestimation of dipole moments in the additive force field, the magnitude of the dipoles in the Drude-2013 force field is in better agreement with QM data.[179] These dipoles also respond to peptide conformation, including longer-range intramolecular interactions, such as the $i$ to $i + 4$ hydrogen bonds in $\alpha$-helices (see sections 4.1.1 and 4.1.2). Additionally, it was shown that the side-chain dipole moments of tryptophan residues respond very strongly to conformation (by rotation about $\chi_1$ torsions and changes in the local environment), varying over a significantly wider range than CHARMM36. This observation indicates that the polarizable model exhibits sensitivity to variations in the local electric field in response to small conformational changes, a feature that is essentially absent in additive force fields. Similarly, the dipole moments of water molecules around charged amino acids vary as a function of distance from these groups, as shown in Figure 5B,C. While the nearest water molecules have slightly decreased dipole moments, at the maximum of the radial distribution function around the charged groups, the water dipole moments increase, before converging to the bulk value ($\sim$2.46 D for SWM4-NDP) at longer distance. The mutual polarization response between the charged residues of the protein and solvating waters shows a small but nontrivial response to the properties of microenvironments on the protein surface.

The Drude-2013 protein force field exhibits greater flexibility than CHARMM36, such that the RMSD values of a wide range of proteins are typically larger than in the additive simulations. This outcome suggests that the polarizable model allows for sampling of a broader conformational ensemble, while still remaining in reasonable agreement with available NMR data. The Drude-2013 model represents the first generation in protein force fields with the Drude oscillator model, and future refinements of the model will target better agreement with solution NMR and other data to even more accurately represent conformational dynamics of proteins.

Recently, the interactions between the Drude-2013 protein force field and several biologically important ions (Na$^+$, K$^+$, Ca$^{2+}$, and Cl$^-$) have been refined.[180] Off-diagonal atom-pair-specific LJ terms (NBFIX) and through-space Thole screening (NBTHOLE) parameters were derived by targeting QM geometries and energies of ion–small molecule complexes. These small molecules included NMA (as a representation of the protein backbone, and side-chain amides of glutamine and asparagine), ethanol (a model of serine and threonine), and propionic acid (to model aspartic and glutamic acids). The newly parametrized interactions were evaluated in terms of solvation free energies ($\Delta G_{solv}$) in water, NMA, and ethanol; both single ion and salt $\Delta G_{solv}$ values were in near-quantitative agreement with experimental data. Even more importantly, the pair-specific interactions were able to reproduce transfer free energies between solvents, indicating that the polarizable ion models can respond to changes in dielectric of the surrounding solvent, a key property of the polarizable model. The Drude polarizable model also manifested cooperativity in geometry changes as a function of ligands. For instance, the oxygen–ion distances in NMA complexes increased as a function of the number of NMA molecules in the complex (from one to four). This behavior reflects the outcome of the QM calculations, suggesting that the empirical model can recover the cooperative dipole behavior on the QM level. The final, and most significant, outcome of this parameter refinement was shown in the behavior of protein–ion complexes. By assessing 30 enzyme structures that have defined ion binding sites, it was found that the polarizable model correctly reproduced structural and thermodynamic data. Subsets of each ion binding site were extracted from structures following short MD simulations for QM and MM evaluation. Whereas the additive CHARMM36 force field systematically overestimated binding affinity, from 37 kcal mol$^{-1}$ in the case of K$^+$ to as much as 80 kcal mol$^{-1}$ in the case of Ca$^{2+}$, the Drude polarizable model was in near-quantitative agreement for monovalent ions and within 10% of QM values even for Ca$^{2+}$, which strongly polarizes its coordinating ligands. This investigation also allowed for the first quantitative assessment of many-body effects in protein–ion complexes. The Drude-2013 force field was compared with CHARMM36 and the AMOEBA polarizable force field[57,58] and was found to best reproduce the energy associated with nonadditivity. AMOEBA overestimated this phenomenon, and by definition the nonadditive contribution in CHARMM36 was zero. These effects of polarization may account for as much as 30% of the binding energy (in the case of Ca$^{2+}$) in protein–ion complexes, highlighting how traditional fixed-charge force fields cannot account for ion binding in strongly polarizable microenvironments such as those found in proteins; even the most carefully parametrized models are unlikely to ever correctly account for such phenomena. In this regard, the development of parameters that accurately describe such interactions in the Drude polarizable model represents a

breakthrough in the pursuit of more realistic simulations of protein–ion complexes.

**3.3.6. DNA.** A rudimentary nucleic acid force field was proposed in the early work by Anisimov et al. when describing the method for deriving electrostatic parameters for the Drude polarizable force field.[125] A short (5 ns) simulation of a small DNA duplex was carried out using a topology produced by combining parameters of nucleic acid bases derived in that work with sugar and dimethyl phosphate (DMP) parameters. The DNA duplex remained intact over this short time period, suggesting that the proposed parametrization approach would be suitable for complex (and highly charged) systems such as DNA, but the progressive increase in RMSD of the structure indicated that further refinement would be needed. Building upon the work of Lopes et al. in deriving parameters for nitrogen-containing heterocycles,[165] parameters for the five common nucleic acid bases (adenine, guanine, cytosine, thymine, and uracil) were produced in 2011.[181] These nucleic acid base parameters became the basis of the full DNA force field, which was published in 2014.[177,182]

Development of the Drude-2013 DNA force field simultaneously targeted the intrinsic conformational energetics of several model compounds, as well as the condensed-phase behavior of DNA. The performance of the force field was assessed in terms of the energetics of these model compounds, crystal survey data on A- and B-form DNA, and NMR order parameters, which provide crucial insight into the structure and dynamics of DNA. The most important of these local conformational changes include sugar repuckering between north (N) and south (S) conformations and the BI/BII equilibrium in B-DNA that is a function of backbone $\varepsilon$ and $\zeta$ torsions (BI, $\varepsilon - \zeta < 0°$; BII, $\varepsilon - \zeta > 0°$). The BI/BII equilibrium modulates the major and minor groove widths in DNA, so proper balance of these states is important in establishing a physically realistic model of DNA, especially with the goal of a comprehensive force field capable of simulating heterogeneous systems, such as protein–DNA complexes.

As with the Drude-2013 protein force field, combining small-molecule building blocks required additional optimization of bonded terms to yield a stable force field, and suitable model compounds had to be identified that could account, at least in part, for the correlation that is inherent in DNA backbone torsions, sugar repuckering, and base orientation. Several noncanonical geometries of sugar puckering and base orientation relative to the sugar were assessed to act as target data in the development of the Drude model; such species are important in capturing the energetics during conformational transitions between A- and B-form DNA helices. In addition to model compounds shown in Figure 8, the correlation in DNA dynamics had to be assessed by conducting simulations of DNA in solution.

The Drude-2013 DNA force field better reproduces the two-dimensional QM potential energy surface for the rotations of the $\zeta$ torsion as a function of $\varepsilon$ than the additive CHARMM36 model. This agreement was achieved through specific tuning of torsional terms on the small molecule level using model T3PM in Figure 8[182] and also empirical optimization of electrostatic and LJ terms in the phosphodiester backbone.[177] While CHARMM36 has a pronounced minimum in the BII state and a low barrier to the BI state, the Drude model was better able to capture the subtle features of the QM energy surface, including better relative energetics between the BI and BII states. Similarly, the Drude model reproduced the conforma-
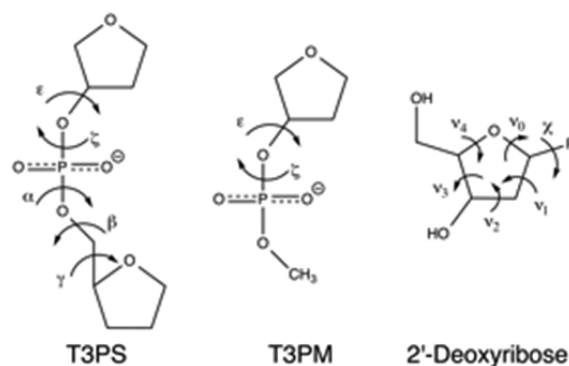


**Figure 8.** Model compounds used in the parametrization of the Drude-2013 DNA force field. For 2′-deoxyribose, "R" indicates one of the four standard nucleic acid bases (adenine, guanine, cytosine, and thymine). Backbone dihedral angles ($\alpha$, $\beta$, $\gamma$, $\varepsilon$, $\zeta$) are shown for T3PS and T3PM, and the torsions defining sugar puckering ($\nu_0$, $\nu_1$, $\nu_2$, $\nu_3$, $\nu_4$) and base orientation relative to the sugar ($\chi$) are labeled on 2′-deoxyribose.

tional energies of noncanonical nucleoside geometries, while the additive model generally overestimated the energies. These outcomes suggest that the inclusion of explicit polarization allows for a better representation of subtle conformational change that is critical to the proper representation of DNA dynamics. As a result, the Drude-2013 force field was able to reasonably reproduce experimentally observed, sequence-specific BII content and $S^2$ order parameters from NMR for the *Eco*RI dodecamer. Sugar repuckering was shown to occur on a time scale of ~200 ps, in agreement with experimental data.[183] Taken together, these data indicate that the polarizable model more accurately describes the kinetics of local conformational change, an attribute that derives from the greater ability of the Drude model to reproduce the QM conformational energies, though further testing is required.

Another important test of any DNA force field is the response of the DNA to its environment. In solutions with low water activity, DNA helices generally favor the A-form. Nucleotide content can modulate this effect, with GC-rich regions stabilizing the A-form to a greater extent than AT-rich sequences. The Drude-2013 DNA force field was assessed by simulating a GC-rich A-form DNA sequence in water and a solution of 75% ethanol. The A-form was stable in ethanol and quickly converted to the B-form in water. The results indicate that the Drude model is capable of correctly responding to changes in solvent polarity. However, we note that the Drude-2013 DNA force field does not yield stable Z-DNA conformations; ongoing efforts are being made to address this and other limitations of the model as well as to extend the model to RNA.

Given that DNA is a polyanion, with charge density much larger than that of most proteins, the electronic response of nucleic acid bases to hydrating water and ions is of particular importance. Over the course of the full-length DNA simulations, the dipole moments of the nucleic acid bases were systematically larger and more variable than those calculated from the additive CHARMM36 simulations, despite the inherent overestimation of gas-phase dipole moments in the parametrization of the additive force field. This outcome suggests that the bases in the Drude model are more sensitive to the electronic environment within the DNA helix and that the additive assumption of overestimating dipole moments may

be an inadequate representation of the condensed-phase environment. Given the fact that gas-phase polarizabilities were scaled by 0.85 in the parametrization of the bases,[181] the Drude model is unlikely to be overpolarized. The functional implications of nucleic acid base dipoles will be discussed later in this review (see section 4.2).

As was the case for water molecules solvating charged residues in the Drude-2013 protein force field,[178] water molecules in the major and minor grooves of DNA also exhibited distance-dependent variations in their dipole moments. Differences in dipole moments can also be seen as a function of geometry; that is, waters respond slightly differently in the major and minor grooves. At larger distances (beyond the first hydration shell), the water dipole moments relax toward the bulk value (∼2.46 D). In contrast, the dipole moments of the rigid TIP3P water molecules in the additive CHARMM36 simulations were invariant, by definition. The inclusion of explicit polarization suggests that there is interesting behavior in hydrating waters that is unaccounted for in simulations with additive force fields. The balance between intramolecular and biomolecule−water interactions is an important factor in determining structure and functional dynamics. As a consequence, the polarizable model represents a significant advancement in our understanding of these fundamental forces.

In addition to interactions with water, the interactions between polyanionic DNA and mobile ions in aqueous solution are also very important to the structure and dynamics of DNA, and nucleic acids in general. To tune the parameters for ion−DNA interactions, osmotic pressure calculations[175] were conducted on solutions of Na+−DMP and Na+−acetate over a range of concentrations. QM interaction energies and geometries of DMP−water complexes and the $\Delta G_{solv}$ of DMP were also targeted in the refinement of the LJ parameters of DMP and, thus, the phosphodiester backbone of DNA. These attributes are important in characterizing the hydrated structure of the DNA backbone and the overall balance of interactions within DNA and between DNA and the surrounding solvent. Another critical refinement in the Drude-2013 DNA force field was balancing interactions between Na+ ions and nucleic acid bases. Although the parameters for the ions[107,176] and bases[181] were developed using similar strategies, their interplay had not been previously validated. It was observed that the Na+ ions were too strongly attracted to nitrogen atoms in the purine and pyrimidine ring systems. To better match QM interaction energies and distances, off-diagonal NBFIX terms had to be applied to preserve the quality of the base parameters and simultaneously produce reasonable interactions with ions. The completed Drude-2013 DNA force field was shown to quantitatively reproduce the extent of charge neutralization by ions around DNA, in accordance with counterion condensation (CC) theory.[184] While both the CHARMM36 and Drude-2013 force fields yielded counterion shells around the DNA, the level of condensation in the Drude model was improved, indicating that the nonbonded interactions in the polarizable model provide a better physical description of the balance of interactions among ions, water, and DNA.

**3.3.7. Carbohydrates.** Carbohydrates are ubiquitous in biological systems, playing a central role in metabolism as well as serving as covalent adducts for proteins and lipids.[154,155] The force field for linear polyols was introduced in 2013,[185] and

parameter sets for hexopyranose[186] and aldopentofuranose[187] monosaccharides have recently been completed.

The linear polyol series was parametrized by targeting the conformational energetics and molecular dipole moments of nearly 2000 unique conformers of compounds with two to six alcohol groups (ethylene glycol, glycerol, and a series of tetritols, pentitols, and hexitols). The additive CHARMM parameters for these molecules overestimated energies and dipoles in many of these conformers due to the inherent overpolarization of the alcohol groups. This convention produces considerable repulsion between neighboring oxygen atoms due to the large partial charges. In contrast, the Drude model better reproduced the target QM data, demonstrating the role of explicit polarization in recovering gas-phase conformational energetics and dipole response. This behavior is a critical component of the carbohydrate force field, as the balance of intra- and intermolecular hydrogen bonding is crucial to the physicochemical properties of this class of molecules. Correctly accounting for the polarization effects allowed the Drude-2013 force field to improve upon the additive model in all quantities considered in the validation, including dielectric constants, $V_m$, $\Delta H_{vap}$, $\Delta G_{hydr}$, crystal volumes and associated lattice parameters, and aqueous solutions of varying polyol concentration. As a notable example, the $\Delta H_{vap}$ for glycerol with the additive CHARMM force field was underestimated by 14.2% because the intrinsic over-polarization of the alcohol groups that led to reasonable interactions with water caused very strong intramolecular hydrogen bonding between vicinal hydroxyls in the gas phase, depressing the energy barrier relative to the neat liquid. In contrast, $\Delta H_{vap}$ with the Drude model of glycerol agreed with the experimental value within 1.0%. The cooperative dipole enhancement of the three vicinal hydroxyls is an important factor in achieving proper balance between intra- and intermolecular nonbonded interactions and is unique to the Drude polarizable model.

Given the similarities in the parametrization protocol and validation of the hexopyranose[186] and aldopentofuranose[187] monosaccharide force fields, the remainder of this section will be devoted to a unified discussion of both. Electrostatic parameters were derived using the MC/SA fitting algorithm described above, starting with parameters transferred from the linear polyol[185] and cyclic ether[153,156] force fields. The target data for the fitting consisted of QM dipole moments to tune partial charges, atomic polarizabilities, and atom-based Thole screening factors. For the hexopyranoses, 237 unique con-formers (encompassing 16 diastereomers) were used as target data, while 60 furanose conformers were used for the aldopentofuranoses. Fitting was carried out using three strategies, the first based on per-isomer fitting (termed "Isomerfit"), yielding 16 parameter sets for the hexopyranoses and four sets for the furanoses. An alternative, per-anomer fitting protocol (called "Anomerfit") employed a constraint such that parameters were equivalent between the two anomers of each diastereomer, thus halving the total number of possible parameter combinations. The final strategy, targeting all conformers in one fitting run (called "Globalfit") and enforcing equivalency among alcohol groups in the ring to avoid overfitting, yielded a single parameter set that was chosen as the final parameter set for both of the carbohydrate series. While fitting on a per-isomer or per-anomer basis typically yielded better overall RMSD between QM and MM dipoles, these parameters were not transferable between monosacchar-
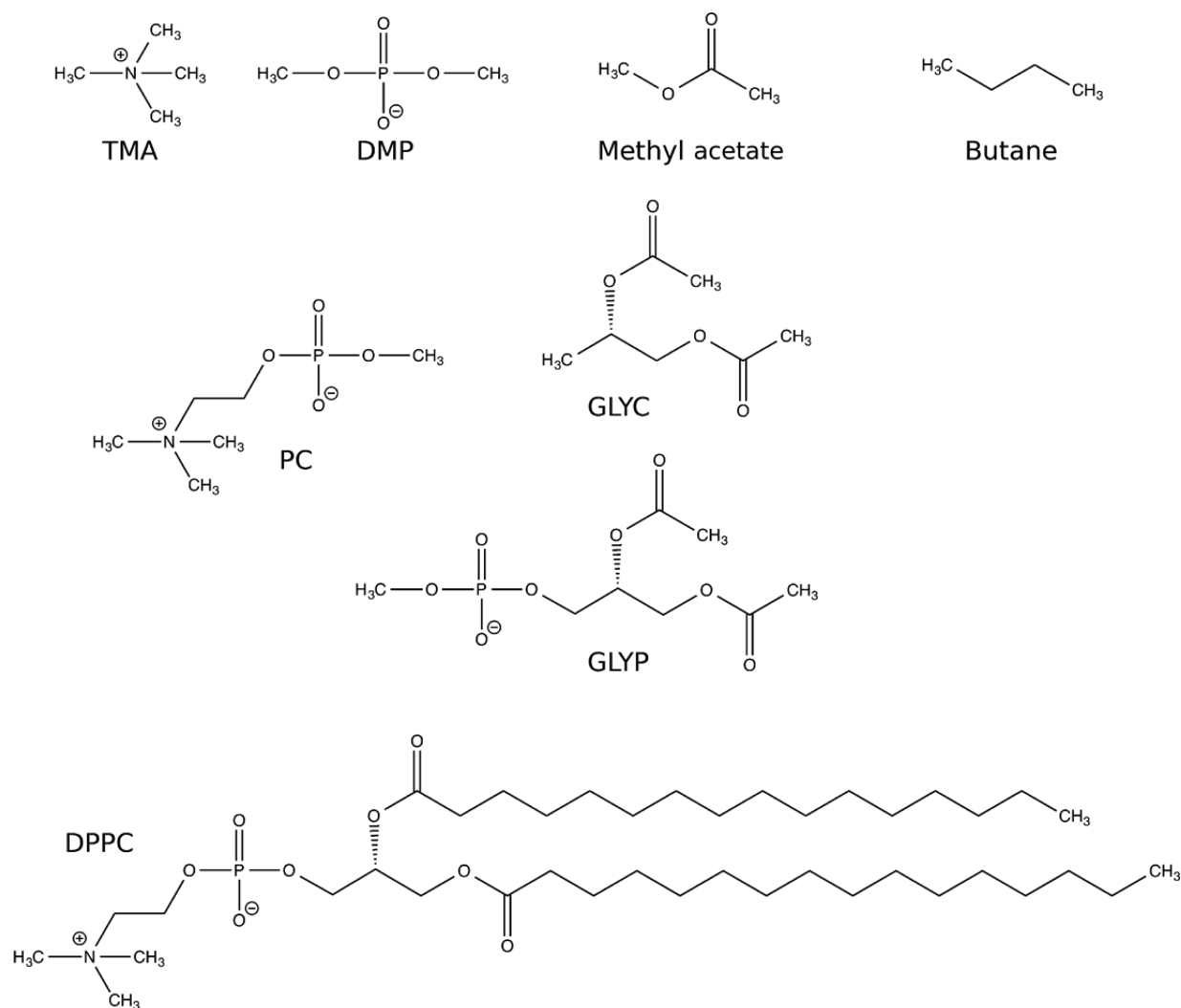
**Figure 9.** Hierarchical approach to DPPC parametrization. Model compounds used for constituent functional groups (trimethylammonium, TMA; dimethyl phosphate, DMP; methyl acetate; and butane) and for torsional fitting (phosphatidylcholine, PC; esterified glycerol, GLYC; phosphorylated glycerol, GLYP) and the full structure of the DPPC phospholipid.

ides. The global fitting protocol produced a unified, transferable parameter set the quality of which was comparable to that of the other methods and represented a significant improvement over the additive CHARMM force field, which utilizes overestimated gas-phase dipole moments. Torsional fitting was carried out using the same global fitting strategy as applied to the additive model, targeting the gas-phase energetics of an extensive conformer series (1846 hexopyranose conformers and 1096 aldopentofuranose conformers). Again, significant improvements in the reproduction of the QM data over the additive CHARMM force field were obtained. Achieving correct conformational energetics is a critical component of the carbohydrate force field, as only low energy barriers may separate different ring and hydroxyl conformers; relative energetics dictate the quality of the conformational sampling in the force field.

The polarizable carbohydrate parameter sets were validated by examining a number of condensed-phase properties, including crystal volumes, NMR $J$-coupling constants, solution densities of varying monosaccharide concentration, and sugar puckering during MD simulations. Crystal simulations of hexopyranoses and O-methylated furanosides showed an improvement over the additive CHARMM force field; despite

the fact that both force fields slightly overestimate crystal volumes, the Drude-2013 force field was in decidedly better agreement with experimental values. Aqueous simulations of varying monosaccharide concentration were in better agreement with available experimental data using the Drude model, differing generally by <1% and manifesting correct temperature-dependent behavior. However, underestimation of the diffusion constants of monosaccharides at higher concentrations was observed, a problem also present in the additive model. Cooperativity in neighboring dipoles of the monosaccharide alcohol groups was observed in aqueous solution, a unique feature of the Drude polarizable model that cannot be obtained with an additive model. Such behavior is responsible for the environment- and conformation-dependent behavior of the force field.

The additive and polarizable models performed similarly in terms of the ability to reproduce sugar-puckering behavior, but the Drude model is notable in its ability to qualitatively reproduce the conformational tendencies of exocyclic groups on the monosaccharide rings. In the development of the hexopyranose force field,[186] Hamiltonian replica-exchange (HREX) simulations were used to comprehensively sample conformational space of the $\omega$ torsion angle (O5−C5−C6−

O6), while long, standard MD simulations were used in the case of the aldopentofuranoses to sample the O4−C4−C5−O5 exocyclic torsion angle.[187] The high quality of the monosaccharide force field derives from strong transferability between the parameters, indicating that development of higher-order polysaccharide models will be possible in the context of a comprehensive biomolecular force field. As of the writing of this review, work on parameters for a range of disaccharides was ongoing, along with further refinement of the hexapyranose monosaccharide parameters.

**3.3.8. Lipid Membranes.** Lipid membranes serve as barriers between intra- and extracellular environments, around organelles, and to regulate the transport of materials throughout the cell. As such, the proper thermodynamic and transport properties of lipids represent an important component of a force field. The dielectric environment of a lipid membrane is complex, with strongly polarizable head groups with large dipole moments at the membrane−water interface, glycerol linkages, and a low-polarity (hydrophobic) alkane interior. Such properties represent a considerable challenge for additive force fields, and though progress has been made toward accurate models of membranes,[28] multibody effects at interfaces are significant in determining the membrane dipole potential, which additive force fields overestimate.[188] By conducting simulations of lipid monolayers with a first-generation dipalmitoylphosphatidylcholine (DPPC, Figure 9) using the Drude polarizable model at air/water interfaces, Harder et al. quantified the effects of induced polarization of lipid headgroups, alkyl chains, and water molecules. The polarization response was found to be dependent upon both distance and orientation, thus illustrating the synergy between membrane structure and electronic response. Interfacial water molecules were strongly polarized by the lipid headgroups, which themselves gave a large contribution to the membrane potential. Water dipole moments relaxed back toward a gas-phase value at the air/water interface, demonstrating the sensitivity of the polarizable model. Additive simulations with the TIP3P water model did not, by definition, show such a response, and the contribution to the membrane potential from water was significantly lower. As a result, the net difference in membrane potential between the two interfaces ($\Delta V$) was approximately 2 times larger than experimental values using the additive model but in quantitative agreement with the polarizable model. The polarization response in the alkyl chains of the lipids was a significant factor in this outcome, as it buffered the dipole potential.

On the basis of this foundational study,[188] and the fact that the membrane potential modulates the permeability of ions[189] and the association of proteins with membranes and their resulting structures and dynamics,[190] a polarizable description of lipids is desirable for simulating complex, heterogeneous biological systems. To this end, a refined Drude polarizable model for DPPC was developed by Chowdhary et al. in 2013.[119] The DPPC model was initially constructed from several model compounds that represent the various chemical moieties in the full lipid (Figure 9), including tetramethylammonium (TMA) and DMP for the phosphatidylcholine (PC) headgroup, esterified glycerol and its phosphorylated form to represent the linkage between the headgroup and glycerol backbone, and n-alkanes for the long alkyl chains. The parameters for n-heptane were transferred directly from the Drude alkane force field,[149] with dihedral parameters refined to accurately account for the delicate balance between gauche and

trans forms of the torsions, which are critical to a proper representation of the structural features in the hydrophobic interior of the membrane. The linkages between these alkyl chains and the glycerol backbone required the parametrization of the ester linkage, which was not previously available in the force field. Methyl acetate was chosen as a model compound for this functional group, and the accuracy of its parameters was crucial to the success of the final DPPC model, as its hydration properties and conformational energetics give rise to important features of the membrane/water interface. Good agreement was achieved over a wide range of QM and thermodynamic quantities of methyl acetate, including molecular polarizability, dipole moment, water interaction energies, $\Delta H_{\text{vap}}$, neat liquid $V_{\text{m}}$, $\Delta G_{\text{hydr}}$, and dielectric permittivity. Transfer of the methyl acetate parameters to larger esters yielded good agreement with available $\Delta H_{\text{vap}}$ and $V_{\text{m}}$ data, indicating that the derived parameters were suitable for use with larger compounds such as lipids.

The fully constructed Drude-2013 DPPC model produced good agreement with experimental data and the CHARMM36 force field,[191] which also reproduces many structural properties of DPPC membranes. Among the most extensively studied properties of DPPC are X-ray scattering form factors, deuterium order parameters, and area and volume per lipid. The Drude-2013 DPPC model produced form factors in slightly better agreement than the CHARMM36 simulations, though both force fields perform well in this regard. The X-ray scattering form factors are produced from an analytic fit of heavy-atom electron densities, which are better reproduced by CHARMM36, though the differences with the Drude model are small. The discrepancy between these two observables may be related to the quality of the analytic fit and the considerable scatter in experimental data for long wavelengths. The balance of forces between the hydrophobic effect in the membrane core and the hydration of the lipid headgroups and glycerol region gives rise to surface tension, and thus, this quantity is a measure of the quality of the balance between nonbonded terms in the force field. The balance of these interactions also manifests itself in the area and volume per lipid in the membrane, as well as the membrane thickness. Together, these quantities are indicators of the overall membrane structure and the cohesive forces within it. The Drude model slightly underestimated area per lipid and slightly overestimated the volume per lipid and membrane thickness, but all results were in reasonable agreement with the range of experimental data and were of comparable quality to the CHARMM36 force field. Refinement would likely improve the overall agreement, but the present model reflects a viable first-generation lipid force field for DPPC.

Deuterium order parameters along the DPPC structure were used as specific target data in the development of the CHARMM36 force field,[191] and for this reason, good agreement is obtained in simulations with the additive model. Despite not being directly targeted, comparable agreement was observed with the Drude-2013 force field, but several notable deviations in the headgroup carbon atoms were observed, indicating that improvements are necessary. Torsional scans of several rotatable bonds revealed that the ability of the Drude model to reproduce gas-phase QM conformational energetics in model compounds was lacking, though the large number of conformers that needs to be assessed remains a considerable challenge. The locations of energetic minima were in agreement, indicating a reasonable model for conformational

sampling, but future improvements will be needed to yield better agreement with deuterium order parameters as well as chemical shift anisotropy.

The interactions of the Drude-2013 DPPC force field with water also reveal important insights into the viability and application of the polarizable membrane model. The Drude model produced better surface tension values than CHARMM36 as a function of area per lipid, though this improvement is due in large part to the better surface tension in SWM4-NDP relative to TIP3P. Overall, the consistency achieved indicates reasonable balance between the lipid and water parameter sets, indicating that future refinements targeting other experimental data should be straightforward. The interaction of DPPC with SWM4-NDP also resulted in slower lateral diffusion of the lipids, again as a consequence of the slower diffusion of SWM4-NDP relative to TIP3P; the former water model is in better quantitative agreement with experimental data, while the latter diffuses nearly 3 times too fast. The final assessment of the Drude-2013 DPPC model was to simulate DPPC monolayers at air/water interfaces, as was done in work by Harder et al.[188] As in that study, the membrane electrostatic potential of the Drude model was in decidedly better agreement than the CHARMM36 force field, which overestimates it by a factor of 2. The polarizability of the alkane moieties in the hydrophobic core is crucial to the agreement with experimental data.

As a final note, it is important to highlight the fact that the Drude-2013 DPPC force field (and the entire Drude-2013 force field described in the present review) was developed for use with a potential-switching function for LJ interactions, unlike the CHARMM36 lipid force field,[191] which was developed using force-switching.[192] Proper treatment of LJ forces is critical to obtain reproducible values for membrane properties,[193] and all investigators should take great care to correctly employ these methods.

## 4. BIOMOLECULAR SIMULATIONS WITH THE DRUDE-2013 POLARIZABLE FORCE FIELD

To date, the Drude-2013 force field has been applied to a number of systems, with results from our own laboratory presented below. Inherent to the utility of a force field for condensed-phase simulations of large molecular systems is its computational efficiency. Being able to access a time scale that is biologically relevant and comparable to additive simulations allows for the acquisition of novel insights concerning protein dynamics and conformational ensemble properties. Notably, it also allows for extensive testing of the force field to identify limitations in the model, allowing for improvements going forward.

### 4.1. Proteins

The Drude-2013 protein force field has been applied to a variety of peptides and proteins, with typical simulation times ranging from hundreds of nanoseconds to the microsecond scale. Some specific examples of recent efforts follow.

**4.1.1. Cooperative Helix Formation.** Cooperativity is a central feature in protein folding.[194] However, the driving forces behind this cooperativity are still poorly understood. While force field-based simulations have led to numerous insights into the nature of protein folding, they have failed to capture the level of cooperativity observed in the experimental folding of polypeptides and proteins. One specific example of this failure is the temperature dependence of secondary structure formation. Recently, we showed that the Drude-2013 protein force field can reproduce the experimental temperature dependence of helix formation in the acetyl-(AAQAA)$_3$-NH$_2$ peptide, and the electronic induction of dipole moments of peptide bonds drives such a cooperative folding process.[195]

The (AAQAA)$_3$ peptide serves as a model system in the validation and development of protein force fields. The average helical content at 300 K, which was determined experimentally to be around 20%,[196,197] was used as the target data in developing several of the most recent additive force fields, including AMBER03*, AMBER99SB*, and CHARMM36.[27,198] These data were not included in the parametrization of Drude-2013 protein force field. HREX simulations using the Drude-2013 force field led to a helical content of 25% at 300 K, indicating good balance between the helix and coil conformations that arises from the quality of the underlying parametrization and electrostatic model, rather than as a function of having been directly targeted during force field development. Most importantly, good agreement with experimental helical content was achieved with the Drude-2013 force field at three other temperatures considered (280, 320, and 340 K, with the 320 K data not reported in the original study), reproducing the strong dependence of the helical content on temperature observed in experimental measurements (Figure 10A).

Having shown that the thermodynamic aspect of cooperativity was captured by the Drude model, as reflected in the small temperature interval over the transition from the unfolded to folded state, the microscopic picture of folding cooperativity was investigated, which involves the presence of partially folded states at very low populations along with fully, or nearly fully, folded states. As illustrated in Figure 10B, the Drude-2013 polarizable force field produced a bimodal distribution in which both unfolded coil and folded, "long helix" conformations are highly populated. In contrast, the additive CHARMM36 protein force field generated a more compact, unimodal conformational ensemble with the more disordered or partially folded helices. Analysis of the Lifson–Roig parameters also showed the agreement with the polarizable model to be improved relative to the CHARMM36 force field, consistent with the improved treatment of cooperative helix formation.

The enhanced dipole cooperativity associated with the Drude-2013 force field was shown to be due to the explicit inclusion of electronic polarizability instead of any particular aspect of the force field parametrization. Analysis of peptide bond dipole moments, with emphasis on their variation as a function of peptide backbone conformation, showed the peptide bond dipole moments to be enhanced by ∼0.9 D upon formation of a helical conformation. MD trajectories of Drude simulations can be postprocessed by reminimizing the Drude particles frame-by-frame after any system modifications, such as removing all the water molecules, with the difference in dipole moment reflecting the impact of changing the molecular environment. Such analysis showed that roughly one-third of this enhancement comes from the hydration environment, one-third from the directly adjacent residues, and the remaining one-third from the $i$ to $i + 4$ hydrogen bonding, which is the unique feature of helical structure. The contributions from these terms reflect the multibody nature of polarization response.
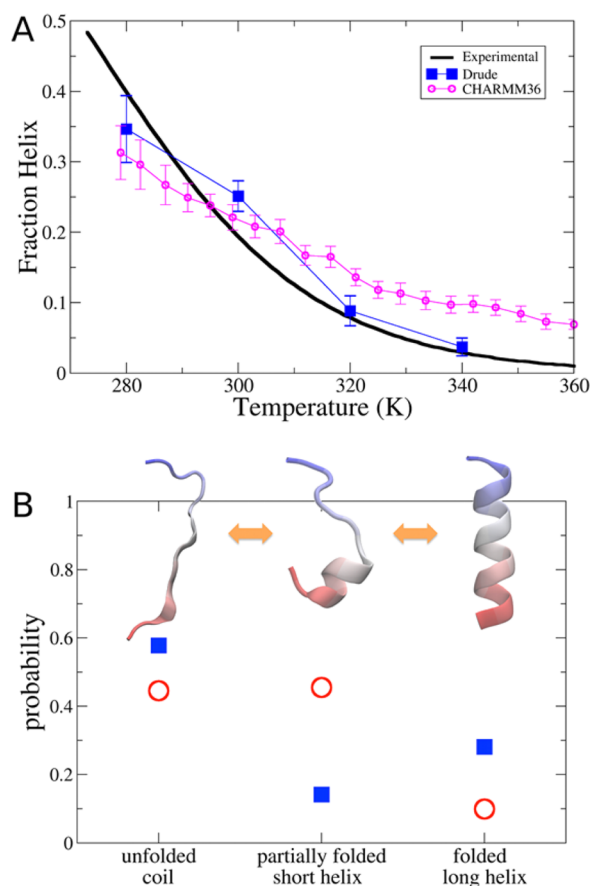
**Figure 11.** Dipole moment distributions for different conformations of the model $(AAQAA)_3$ peptide conducted at 300 K using the CHARMM36 and Drude-2013 force fields. $\alpha$-Helical residues were defined as those found in a sequence of at least three consecutive residues with $(\phi,\psi)$ values characteristic of an $\alpha$-helix ($-100° < \phi < -30°$, $-67° < \psi < -7°$). The peak centered at $\mu = 4.6$ D for the CHARMM36 nonhelix represents sampling of the $\alpha_L$ conformation (which are oversampled by the force field) and are associated with a change in geometry rather than any change in dipole moment associated with electronic structure. $\alpha_L$ conformations were not sampled in the Drude simulations.

**Table 2. Mean Value and Root-Mean-Square Fluctuation of Peptide Bond Dipole Moments (in D) in Different Secondary Structures from MD Simulations of Six Proteins Using the Drude-2013 and CHARMM36 Force Fields**

| | peptide bonds in helices | | peptide bonds in sheets | |
|---|---|---|---|---|
| | mean | fluctuation | mean | fluctuation |
| Drude-2013 | 4.74 | 0.31 | 5.14 | 0.30 |
| CHARMM36 | 3.82 | 0.11 | 3.71 | 0.09 |

helix or $\beta$-hairpin. Instead, the explicit treatment of induced polarization in the Drude-2013 force field can quickly respond to local variations in the environment.

QM calculations have indicated that the peptide bond dipole moment is approximately 5 D,[179] consistent with the values obtained with the Drude-2013 force field. It is worth noting that most theoretical arguments regarding the helix macrodipole (arising from the alignment of individual peptide bonds along the $\alpha$-helix axis) assume a peptide bond dipole of 3.5 D.[199] However, taking into account a mutual induction effect, it should be 4.7 D, which means the macrodipole should be 35% stronger and modeled with a charge separation of ±0.7 $e$, instead of ±0.5 $e$, at the two ends of the helix that is typically assumed.

The larger magnitude of dipoles observed with the Drude model is notable in itself. The parametrization of additive force fields typically involves systematic overestimation of dipole moments relative to the gas phase, which in the case of the CHARMM force fields is about 20%, to capture the induction effect of the environment according to a mean-field approximation.[44] In a polarizable model, the alignment of peptide dipole moments during secondary structure formation leads to favorable interactions enhanced with the larger dipoles, which is counterbalanced by the positive self-energy for polarizing the atoms (eq 5). The Drude results indicate that amino acids carry larger dipole moments that are also more responsive to the complex protein environment.

**Figure 10.** Folding cooperativity of $(AAQAA)_3$ obtained with the Drude-2013 protein force field. (A) Fraction helix as a function of temperature determined from NMR experiments and replica exchange simulations with the Drude polarizable (blue) and the CHARMM36 additive (purple) force fields. (B) Probabilities of the peptide being in the unfolded coil, partially folded, or folded states from 300 K simulations with Drude (filled blue squares) and CHARMM36 (open red circles) force fields. "Long helix" is defined as at least eight consecutive residues being in the helical segment, forming two or more helical turns.

### 4.1.2. Peptide Bond Dipole Moments.

The average peptide bond dipole moment of helical residues in the $(AAQAA)_3$ peptide was 4.91 D in simulations with the Drude-2013 force field, much larger than the value of 3.87 D obtained in simulations with CHARMM36. The CHARMM36 results also show an insensitivity to peptide conformation, with dipole moment distributions for $\alpha$-helix and random coil that are overlapping; in contrast, the Drude-2013 force field showed conformation-dependent variations in dipole moments (Figure 11). Similar analysis was carried out for the microsecond MD trajectories of several globular proteins,[120] and the results are summarized in Table 2. With the Drude model, residues in different secondary structures carry significantly different dipole moments, while with the additive CHARMM36 there is little differentiation. The CHARMM36 simulations also yield much smaller fluctuations. For a fixed-charge force field such as CHARMM36, the variation of dipole moment is only associated with the change in geometry (for example, in the case of backbone dipole, only the elongation of the C=O bond). The nearly constant values of peptide backbone dipole moment in additive force fields might limit their ability to balance secondary structure elements without bias toward $\alpha$-
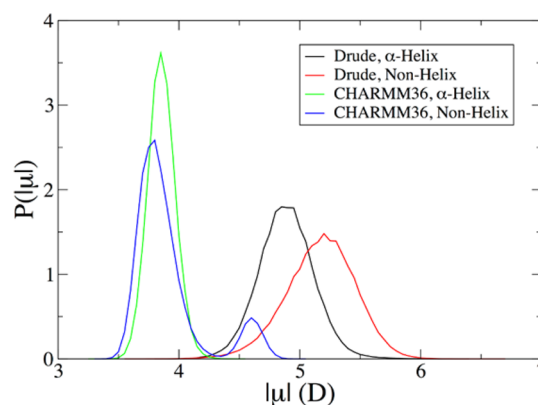
Using the Drude-2013 force field, the role of peptide-bond dipole moments in the stability of $\alpha$-helices has recently been explored in the context of the amyloid $\beta$-peptide (A$\beta$),[200] the central pathological protein in Alzheimer's disease. By simulating a fragment of A$\beta$ that harbors several familial mutations that alter the electrostatic nature of positions 22 and 23 in the amino acid sequence, polarizable simulations revealed that side-chain dynamics have important implications for the stability of the initial $\alpha$-helical structure due to alterations of the dipole moments of peptide bonds. For example, in the wild-type peptide, electrostatic repulsion between the Glu22 and Asp23 side chains led to a perturbation of peptide-bond dipole moments up to four residues away, through the hydrogen-bonding network of the $\alpha$-helix. This perturbation directly led to unfolding. Similarly, the formation of side chain–side chain or side chain–backbone hydrogen bonds in the D23N and E22Q mutants led to destabilization, suggesting that charge–charge interactions are not unique in their ability to destabilize the A$\beta$ helical structure; different behaviors exhibited by the various mutants could perturb the peptide-bond dipoles and lead to unfolding. Taken together, these results suggest subtle but important details that are revealed by the use of a polarizable force field, as these same behaviors were not observed when using the additive CHARMM36 force field, which suggested only nonspecific unfolding phenomena.[200]

**4.1.3. Protein Dielectric Properties.** The fluctuation of dipole moments in biomolecules is related to their dielectric properties.[201] The specific dielectric constants computed from Kirkwood $G$-factors[202] were recently compared between the Drude polarizable model and the additive CHARMM36 force field.[120] The static dielectric constants are similar for whole proteins, but for their hydrophobic cores, the Drude-2013 dielectrics are systematically larger than those from CHARMM36. This observation indicates that charged residues on protein surfaces dominate the dielectric properties of entire proteins. For the buried residues, additional electronic degrees of freedom of the charge distribution are important, an outcome related to the better description of alkane dielectric properties described above.[149]

The Drude model allows the calculation of ensemble-averaged optical dielectric constants ($\varepsilon_{inf}$), based on molecular polarizability along simulation trajectories. Computed $\varepsilon_{inf}$ for six proteins range from 1.76 to 2.10 with an average value of 2.0,[120] which equals the commonly assumed value.[202,203] The variation across proteins suggests that a polarizable model is necessary for studying properties dominated by the protein dielectric relaxation, such as reorganization energy.[204,205]

## 4.2. DNA

Due to the polyanionic nature of nucleic acids, modeling their structure and dynamics using empirical force fields requires an accurate description of their interactions with water and ions. Considerable effort has gone into achieving such a delicate balance,[177] and recent studies using the Drude-2013 DNA force field have focused on analyzing DNA–ion interactions, including the impact of ions on DNA conformation,[206−208] and base flipping in DNA.[209] We summarize these recent efforts below.

**4.2.1. Ionic Structure around DNA.** The microscopic details of the interactions between DNA and monovalent ions have been examined in recent studies, including a description of the ionic structure in solution[177] and ion competition.[207] Structural studies often suffer from an inability to distinguish

between biologically relevant ions; thus, simulations provide a useful tool in rationalizing ion-specific behavior, even providing evidence that can resolve conflicting experimental outcomes, such as those related to groove-specific ion binding.[210,211] Experimental techniques such as buffer equilibration–atomic emission spectroscopy (BE−AES) can be used to study cation competition in the ionic atmosphere around DNA using mixtures of salts with a common anion.[212] In BE−AES experiments, the excess of a given cation and the depletion of the anion relative to their bulk concentrations quantitatively describe the ionic atmosphere around DNA. These data provide a very powerful metric for evaluating the performance of empirical force fields in describing the electrostatic environment of nucleic acids in solution.

Ion competition simulations were carried out using the Drude-2013 DNA force field for the monovalent cations Li$^+$, Na$^+$, K$^+$, and Rb$^+$,[207] following refinement of nonbonded parameters against target gas-phase QM energetic data and osmotic pressure calculations, as described previously.[177] The first assessment of the ionic environment around DNA was a series of volume-Jacobian-normalized radial distribution functions (RDF, Figure 12) for each of the salts, establishing
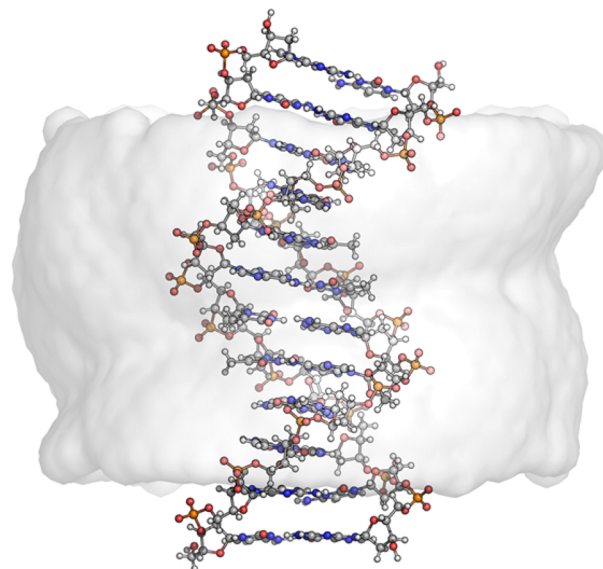


**Figure 12.** A representation of the volume Jacobian defined by an equidistant shell around the central base pairs of a canonical B-form DNA.

a baseline for the ionic structure around DNA. In this regard, both the Drude polarizable model and the additive CHARMM36 force field produced qualitatively similar results, in that smaller ions (Li$^+$ and Na$^+$) retained their hydration shells and participated in water-mediated interactions with DNA, whereas the larger ions (K$^+$ and Rb$^+$) could participate in short-range (direct) interactions with DNA. These results reflect the dehydration penalties for this series of ions. In comparison with CC theory, however, the Drude-2013 force field produced near quantitative agreement with the predicted 76% charge neutralization by the monovalent salts; the CHARMM36 force field generally underestimated the amount of charge neutralization by 5−10%. Charge neutralization is largely a nonspecific and macroscopic phenomenon; that is, DNA of any sequence accumulates positively charged ions in the nearby ionic atmosphere.

Competition between cations was determined relative to Na$^+$ as a background cation; the ion/ion ratio within the conventional 9 Å Manning radius reflects the relative propensity of a given ion to associate with DNA compared to Na$^+$. In agreement with both BE−AES[212] and DNA mobility experiments,[213] the propensity of ions to occupy the ionic atmosphere around DNA follows the trend of Li$^+$ > Na$^+$ > K$^+$ > Rb$^+$ in the simulations, though the use of simulation systems of traditional sizes (e.g., solvated 10 to 12 Å beyond the DNA) allows the results to only be interpreted qualitatively, since the anion RDF does not converge to unity in small simulation systems. To expand upon these findings and obtain results that are quantitatively comparable to BE−AES experiments, large simulation systems are required to avoid ambiguities arising from the distinction between neutralizing cations and those constituting additional salt. By increasing the size of the simulation cell (e.g., box lengths of ∼120 Å), this ambiguity is mitigated, in that the fraction of specific neutralizing counterions is considerably reduced at the expense of increased computational cost. Such larger simulation systems reveal quantitative improvement of the Drude-2013 force field over CHARMM36. Whereas the smaller systems indicated that ion/ion ratios were comparable between the Drude and additive models, the larger systems revealed that, in fact, the Drude model was better able to capture ion competition effects that are in near-quantitative agreement with BE−AES data. This was demonstrated by performing simulations at the appropriate competing ion (Li$^+$, K$^+$, and Rb$^+$) concentrations that BE−AES experiments show yield a ratio of 2.0 for the excess background cations (Na$^+$) in the absence and presence of competing cations. The Drude-2013 DNA force field reproduced this value almost exactly across all of these ions, whereas the CHARMM36 force field produced considerable variation as a function of competing cation identity. These findings reinforce the physical relevance of the specific ion binding described above;[206] that is, the improvement in the microscopic details of DNA−ion interactions manifests into a better macroscopic description of the ionic atmosphere around DNA.

**4.2.2. Impact of Ions on DNA Structure and Dynamics.** Simulations are an essential technique for studying the dynamics of DNA in aqueous solution, due to crystal packing effects that may adversely affect structures determined using X-ray diffraction[214,215] and potentially conflicting outcomes from the mathematical framework provided by NMR that limit is applicability in describing the global conformational properties of DNA.[216] Small-angle X-ray scattering (SAXS) provides a powerful means to understand the conformational properties of DNA in solution, as the dominant properties are highlighted in a one-dimensional trace that accounts for the ensemble of conformations. SAXS fingerprints can be further decomposed into contributions from all the moieties in a biomolecule, for instance, separating the contributions from the nucleic acid bases and the sugar−phosphate backbone in DNA.

To assess the ability of the Drude-2013 DNA force field to describe the solution structure of DNA, the SAXS profiles of two B-form DNA sequences, d(CGCGAATTCGCG)$_2$ (EcoRI)[217] and d(CCGCTAGCGG)$_2$ (PDB ID 1DCV),[218] were computed from MD simulations in aqueous solution containing 100 mM NaCl. These calculated profiles were compared to two additive force fields that are widely used in the simulation of nucleic acids, CHARMM36[25,26] and AMBER parmbsc0.[18] The Drude-2013 DNA force field best captured the features of the SAXS profile, including the positions of spectral peaks that correspond to structural features on different length scales. Both of the additive force fields were in poorer agreement, indicating deviation from the experimental conformations even for global features of both DNA sequences.

Building upon these simulations, the cation identity was varied to compare the effects of Li$^+$, Na$^+$, K$^+$, and Rb$^+$ on the structural properties of both DNA sequences, using SAXS profiles to quantify these effects. The different ions caused distinct differences in these SAXS profiles under the Drude polarizable model, producing shifts in SAXS peaks and small variations in peak amplitude. No such response was observed in the additive CHARMM36 force field, showing that the polarizable model is more sensitive to the effects of different ions. Specifically, while the different ions did not produce differences in the geometric properties of the DNA major groove, the width of the minor groove was modulated as a function of cation size, with spectral shifts related to the minor groove following the order of cation size: Li$^+$ > Na$^+$ > K$^+$ > Rb$^+$. This effect was further found to be sequence-specific, with the EcoRI sequence (which contains an AATT tract) being more sensitive. As no experimental data yet exist to corroborate these findings, they represent a true prediction of the Drude polarizable force field. Thus, the use of the Drude-2013 DNA force field is clearly capable of producing testable hypotheses that arise due to the many-body effects of polarization between the polyanionic DNA and biologically relevant cations.

A mechanism for this ion-mediated perturbation of DNA minor groove widths was recently proposed in a study of the microscopic effects of ion-mediated hydrogen bonding.[208] Specifically, hydrated ions in the minor groove of DNA were able to form what were called "strand(1)−ion−strand(2) hydrogen-bond bridges" (SIS-HBB), which connected electronegative atoms in the paired DNA strands, leading to ion-dependent reduction in minor groove width. That is, smaller ions like Li$^+$ and Na$^+$ had a more profound effect than larger ions like K$^+$ and Rb$^+$, which are also more likely to have hydrating waters dissociate due to the lower affinity of these ions for water. Interestingly, this behavior was only observed with the Drude-2013 DNA force field; ions in the additive CHARMM36 force field were more likely to partially dehydrate, leading to direct interactions of the monovalent ions instead of water-mediated interactions. Direct interactions did not alter minor groove widths. The Drude-2013 DNA force field showed that the effects of SIS-HBB on minor groove width were also sequence-dependent, with the shorter 1DCV sequence being more susceptible to SIS-HBB modulation than the 2L8Q and EcoRI sequences.

In addition to the dynamics of partial dehydration, the Drude-2013 force field produced interesting insights into the variability of water dipole moments. Dipole moments of waters bound to Li$^+$ and Na$^+$, which have a large charge density, increased from their bulk value (∼2.45 D) to ∼2.8 and ∼2.57 D, respectively. Water molecules bound to these ions that also participated in SIS-HBB had dipole moments that were further enhanced by ∼0.1 D. This observation shows the interplay between changes in the electrostatic environment, from diffusion through bulk medium to binding with an ion, and ultimately interacting with DNA.

Finally, by subdividing the MD trajectories between frames in which one or more SIS-HBB were formed and those in which none were present, it was possible to analyze DNA structural properties as a function of SIS-HBB formation. As described above (see section 3.3.6), B-DNA populates both BI

and BII substates, dictated by the values of backbone $\varepsilon$ and $\zeta$ dihedrals. The formation of SIS-HBB in the minor groove of the DNA sequences examined produced an increase in the BII substate, concomitant with an increase in sampling of the "north" sugar pucker. Taken together, these findings indicate that the inclusion of explicit polarization is of particular importance in characterizing the structure and dynamics of DNA in solution, especially in the presence of biologically relevant ions.

**4.2.3. DNA Base Flipping.** In the initial validation simulation of the Drude-2013 DNA force field,[177,182] considerable variation in nucleic acid base dipole moments was observed, reflecting sensitivity to the surrounding environment. To assess whether these effects had implications on high-energy, extrahelical states of the bases, simulations were carried out to determine the free energy change associated with base flipping under the additive CHARMM36 and Drude-2013 DNA force fields.[209] Simulations were carried out using the umbrella sampling technique, in which a harmonic biasing potential is applied along a reaction coordinate to sample high-energy states along a given pathway. In the case of base flipping, the reaction coordinate was a pseudodihedral defined by the center of mass of groups of atoms in consecutive nucleotides in the DNA structure.[219,220] The CHARMM36 force field yielded large energy barriers and thus very low equilibrium constants for base-open versus base-closed states. The Drude-2013 DNA force field produced larger equilibrium constants that were in near-quantitative agreement with NMR data, an improvement of 1−2 orders of magnitude over the additive force field. This outcome indicates that induced polarization is an important contribution to base flipping. Indeed the local electric field that a base experiences along the reaction coordinate is highly variable, and the stabilization of the open state leading to the larger equilibrium constants was shown to a be a result of mutual polarization between water molecules in the first hydration shell around the flipped base and the base itself. The dipole moments of the flipping base increased by up to 1−2 D upon transfer of a base from its Watson−Crick base-paired conformation into the aqueous environment, with hydrating waters subsequently responding on the order of ~0.1 D. While this increase is small on a per-molecule basis, the cumulative effect of dipole enhancement over several water molecules, especially those participating in hydrogen bonds to the base, led to more favorable solvation of the base in the open states in the polarizable force field versus the additive model. These observations further support the notion that induced polarization is a critically important component of a proper physical model and is likely relevant in many biological processes.

## 5. CONCLUSIONS AND FUTURE PERSPECTIVES

In the present review, we have summarized the theory, history, parametrization strategy, and applications of an empirical force field based on the classical Drude oscillator model, called "Drude-2013". The force field has its origins in the additive CHARMM force field and follows much of the same parametrization approach, an iterative refinement of internal and nonbonded terms to match experimental and theoretical (QM) target data. Across a broad class of chemically diverse small molecules, the inclusion of explicit polarization is clearly an improved representation of electrostatic properties relative to traditional fixed-charge (additive) models. Electronic response as a function of environment allows for broad applicability to varying chemical environments, including gas

and condensed phases, as well as differential behavior between solvents of different polarity.

Currently, as of 2016, work is ongoing to further refine parameters for proteins, carbohydrates, lipids, and DNA to achieve agreement with a broader array of experimental data, and a complete nucleic acid force field is nearing completion that will include RNA (Lemkul and MacKerell, work in progress). Parametrization of linkages between monosaccharides will give rise to a fully functional polysaccharide force field, which can then be combined with protein and lipid parameters for representation of glycoproteins and glycolipids. All of these efforts are simultaneously working toward a generalized Drude force field that will encompass druglike molecules, analogous to the CGenFF force field[221] that complements the additive CHARMM force field.[34] The implementation of algorithms for carrying out simulations with the Drude-2013 force field in CHARMM,[111] NAMD,[115] ChemShell QM/MM,[116] OpenMM,[117] and GROMACS,[118] combined with available input generation servers such as the "Drude Prepper" in the CHARMM-GUI[222] and automated parametrization in GAAMP[223] facilitates broad use throughout the theoretical chemistry community.

Finally, we note that period of time required for the development of the model to its current, 2015, state. Work was initiated in March 2000, with the initial SWM4-DP water model published in 2003.[13] Work on the small model compounds and atomic ions[107,180] occurred over the next 8 years up to 2011. It took an additional 2−3 years to produce the first-generation Drude-2013 polarizable force field for proteins,[178] DNA,[177,182] monosaccharides,[186,187] and the membrane phospholipid DPPC.[119] This lag was associated with unexpected challenges inherently associated with non-additive effects encountered upon going from small molecules to the larger, polymeric macromolecules. For example, small imbalances in the interactions between positively charged basic residues and negatively charged acidic residues that were not treated at the small molecule level can lead to unphysical interactions, including polarization catastrophe. Dealing with these issues, which were much less problematic with the CHARMM additive force field in our own experience, required additional time, delaying the release of the Drude protein force field until 2013. However, as presented above, the availability of a computationally accessible, fully polarizable force field for a range of systems including biomolecules has already yielded quantitative improvements over the additive force fields that are associated with the presence of polarizability, allowing for novel insights on the contribution of polarizability to the structure and dynamics of macromolecules to be obtained; we anticipate a large number of further interesting findings as the Drude-2013 force field is applied more widely in the simulation community.

## AUTHOR INFORMATION

### Corresponding Author

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

## Biographies

Justin A. Lemkul received a B.S. in biochemistry In Honors, summa cum laude in 2007 and a Ph.D. in 2012 in biochemistry from Virginia Tech. His dissertation work focused on the conformational dynamics and aggregation pathway of the Alzheimer's amyloid β-peptide in aqueous and membrane environments. He was named the Outstanding Doctoral Student in the College of Agriculture and Life Sciences in 2012 and received the Outstanding Dissertation Award in STEM fields by the Graduate School at Virginia Tech the following year. In 2013, he was awarded a Ruth L. Kirschstein National Research Service Award (F32) from the National Institutes of Health and is now a postdoctoral fellow with Prof. Alex MacKerell in the Computer-Aided Drug Design Center at the University of Maryland, Baltimore. His current research focuses on polarizable force field development and applications related to nucleic acids.

Jing Huang received a B.Sc. in 2005 and a M.S. in 2007 in physics from Tsinghua University, and a Ph.D. in chemistry with honors (summa cum laude) from the University of Basel in 2011. He moved to the United States in 2012 with a fellowship for prospective researchers from the Swiss National Science Foundation, and has worked with Prof. Alex MacKerell in the Computer-Aided Drug Design Center at University of Maryland, Baltimore since then. His general research interest lies in developing better models for computer simulations.

Benoît Roux was born in the city of Montreal, Canada, in 1958. In 1981, he received a B.Sc. in physics from the University of Montreal, followed by a M.Sc. in biophysics in 1985 under the supervision of Remy Sauve. In 1990, he obtained a Ph.D. in biophysics from Harvard University under the direction of Martin Karplus. In the past decade, he has held positions at the University of Montreal and the Weill Medical College of Cornell University. Since 2005, he has been at the University of Chicago, where he is the Amgen Professor of Biochemistry and Molecular Biology and Professor in the Chemistry Department. He also has a joint appointment at Argonne National Laboratory, where he is Senior Computational Biologist.

Alex MacKerell received an A.S. in biology in 1979 from Gloucester County College, Sewell, NJ, followed by a B.S. in chemistry in 1981 from the University of Hawaii, Honolulu, HI, and a Ph.D. in biochemistry in 1985 from Rutgers University, New Brunswick, NJ. Subsequent training involved postdoctoral fellowships in the Department of Medical Biophysics, Karolinska Intitutet, Stockholm, Sweden, in the area of experimental and theoretical biophysics and in the Department of Chemistry, Harvard University in theoretical chemistry. Following one year as a visiting professor at Swarthmore College, Swarthmore, PA, he assumed his faculty position in the School of Pharmacy, University of Maryland, Baltimore in 1993. MacKerell is currently the Grollman-Glick Professor of Pharmaceutical Sciences in the School of Pharmacy and the Director of the University of Maryland Computer-Aided Drug Design Center. MacKerell is also cofounder and Chief Scientific Officer of SilcsBio LLC. Research interests include the development of theoretical chemistry methods, with emphasis on empirical force field development; structure–function studies of proteins, carbohydrates, and nucleic acids; and the application of theoretical methods to drug discovery.

## ABBREVIATIONS USED

| | |
|---|---|
| ahp | atomic hybrid polarizability |
| BE−AES | buffer equilibration−atomic emission spectroscopy |
| CC | counterion condensation |
| DNA | DNA |
| DMP | dimethyl phosphate |
| DPPC | dipalmitoylphosphatidylcholine |
| ESP | electrostatic potential |
| GPU | graphical processing unit |
| HREX | Hamiltonian replica exchange molecular dynamics |
| LJ | Lennard-Jones |
| MD | molecular dynamics |
| NMA | N-methylacetamide |
| PC | phosphatidylcholine |
| QM | quantum mechanics |
| QM/MM | hybrid quantum mechanics/molecular mechanics |
| RDF | radial distribution function |
| RESP | restrained electrostatic potential |
| RNA | ribonucleic acid |
| SAXS | small-angle X-ray scattering |
| SCF | self-consistent field |
| SIS-HBB | strand(1)−ion−strand(2) hydrogen-bond bridge |
| TMA | tetramethylammonium ion |

## REFERENCES

(1) McCammon, J. A.; Gelin, B. R.; Karplus, M. Dynamics of Folded Proteins. *Nature* **1977**, *267*, 585−590.
(2) Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D. E. Pathway and Mechanism of Drug Binding to G-Protein-Coupled Receptors. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 13118−13123.
(3) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; et al. Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **2010**, *330*, 341−346.
(4) Freddolino, P. L.; Arkhipov, A. S.; Larson, S. B.; McPherson, A.; Schulten, K. Molecular Dynamics Simulations of the Complete Satellite Tobacco Mosaic Virus. *Structure* **2006**, *14*, 437−449.
(5) Whitford, P. C.; Blanchard, S. C.; Cate, J. H. D.; Sanbonmatsu, K. Y. Connecting the Kinetics and Energy Landscape of tRNA Translocation on the Ribosome. *PLoS Comput. Biol.* **2013**, *9*, e1003003.
(6) Head-Gordon, M. Quantum Chemistry and Molecular Processes. *J. Phys. Chem.* **1996**, *100*, 13213−13225.
(7) Friesner, R. A. *Ab Initio* Quantum Chemistry: Methodology and Applications. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6648−6653.
(8) Raghavachari, K.; Saha, A. Accurate Composite and Fragment-Based Quantum Chemical Models for Large Molecules. *Chem. Rev.* **2015**, *115*, 5643−5677.
(9) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227−249.
(10) Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11*, 700−733.
(11) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.

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

(13) Lamoureux, G.; MacKerell, A. D., Jr.; Roux, B. A Simple Polarizable Model of Water Based on Classical Drude Oscillators. *J. Chem. Phys.* **2003**, *119*, 5185−5197.

(14) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr; 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.

(15) Wang, J.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.* **2000**, *21*, 1049−1074.

(16) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; et al. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.

(17) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950−1958.

(18) Pérez, A.; Marchán, I.; Svozil, D.; Šponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of $\alpha/\gamma$ Conformers. *Biophys. J.* **2007**, *92*, 3817−3829.

(19) Banáš, P.; Hollas, D.; Zgarbová, M.; Jurečka, P.; Orozco, M.; Cheatham, T. E., III; Šponer, J.; Otyepka, M. Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins. *J. Chem. Theory Comput.* **2010**, *6*, 3836−3849.

(20) Krepl, M.; Zgarbová, M.; Stadlbauer, P.; Otyepka, M.; Banáš, P.; Koča, J.; Cheatham, T. E., III; Jurečka, P.; Šponer, J. Reference Simulations of Noncanonical Nucleic Acids with Difference X Variants of the AMBER Force Field: Quadruplex DNA, Quadruplex RNA, and Z-DNA. *J. Chem. Theory Comput.* **2012**, *8*, 2506−2520.

(21) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. Molecular Mechanical and Molecular Dynamical Simulations of Glycoproteins and Oligosaccharides. 1. GLYCAM_93 Parameter Development. *J. Phys. Chem.* **1995**, *99*, 3832−3846.

(22) Basma, M.; Sundara, S.; Çalgan, D.; Vernali, T.; Woods, R. J. Solvated Ensemble Averaging in the Calculation of Partial Atomic Charges. *J. Comput. Chem.* **2001**, *22*, 1125−1137.

(23) 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.

(24) MacKerell, A. D., Jr.; Banavali, N. K. All-Atom Empirical Force Field for Nucleic Acids: II. Application to Solution MD Simulations of DNA. *J. Comput. Chem.* **2000**, *21*, 105−120.

(25) 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.

(26) Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; MacKerell, A. D., Jr. Impact of 2′-Hydroxyl Sampling on the Conformational Properties of RNA: Update of the CHARMM All-Atom Additive Force Field for RNA. *J. Comput. Chem.* **2011**, *32*, 1929−1943.

(27) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D., Jr. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone Φ,Ψ and Side-Chain $X_1$ and $X_2$ Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257−3273.

(28) Pastor, R. W.; MacKerell, A. D., Jr. Development of the CHARMM Force Field for Lipids. *J. Phys. Chem. Lett.* **2011**, *2*, 1526−1532.

(29) Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D., Jr. Additive Empirical Force Field for Hexopyranose Monosaccharides. *J. Comput. Chem.* **2008**, *29*, 2543−2564.

(30) Hatcher, E. R.; Guvench, O.; MacKerell, A. D., Jr. CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates, and Inositol. *J. Chem. Theory Comput.* **2009**, *5*, 1315−1327.

(31) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D., Jr. CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J. Chem. Theory Comput.* **2009**, *5*, 2353−2370.

(32) Hatcher, E.; Guvench, O.; MacKerell, A. D., Jr. CHARMM Additive All-Atom Force Field for Aldopentofuranoses, Methyl-Aldopentofuranosides, and Fructofuranose. *J. Phys. Chem. B* **2009**, *113*, 12466−12476.

(33) Raman, E. P.; Guvench, O.; MacKerell, A. D., Jr. CHARMM Additive All-Atom Force Field for Glycosidic Linkages in Carbohydrates Involving Furanoses. *J. Phys. Chem. B* **2010**, *114*, 12981−12994.

(34) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(35) Soares, T. A.; Hünenberger, P. H.; Kastenholz, M. A.; Kräutler, V.; Lenz, T.; Lins, R. D.; Oostenbrink, C.; van Gunsteren, W. F. An Improved Nucleic Acid Parameter Set for the GROMOS Force Field. *J. Comput. Chem.* **2005**, *26*, 725−737.

(36) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and Testing of the GROMOS Force-Field Versions 54A7 and 54B7. *Eur. Biophys. J.* **2011**, *40*, 843−856.

(37) Horta, B. A. C.; Fuchs, P. F. J.; van Gunsteren, W. F.; Hünenberger, P. H. New Interaction Parameters for Oxygen Compounds in the GROMOS Force Field: Improved Pure-Liquid and Solvation Properties for Alcohols, Ethers, Aldehydes, Ketones, Carboxylic Acids, and Esters. *J. Chem. Theory Comput.* **2011**, *7*, 1016−1031.

(38) Pol-Fachin, L.; Rusu, V. H.; Verli, H.; Lins, R. D. GROMOS 53A6$_{GLYC}$, an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates. *J. Chem. Theory Comput.* **2012**, *8*, 4681−4690.

(39) Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W. L. OPLS All-Atom Force Field for Carbohydrates. *J. Comput. Chem.* **1997**, *18*, 1955−1970.

(40) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins Via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−7487.

(41) Kony, D.; Damm, W.; Stoll, S.; van Gunsteren, W. F. An Improved OPLS-AA Force Field for Carbohydrates. *J. Comput. Chem.* **2002**, *23*, 1416−1429.

(42) MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., III Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* **2004**, *25*, 1400−1415.

(43) Saiz, L.; Klein, M. L. Computer Simulation Studies of Model Biological Membranes. *Acc. Chem. Res.* **2002**, *35*, 482−489.

(44) MacKerell, A. D., Jr. Empirical Force Fields for Biological Macromolecules: Overview and Issues. *J. Comput. Chem.* **2004**, *25*, 1584−1604.

(45) Halgren, T. A.; Damm, W. Polarizable Force Fields. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236−242.

(46) Caldwell, J. W.; Kollman, P. A. Cation-$\pi$ Interactions: Nonadditive Effects Are Critical in Their Accurate Representations. *J. Am. Chem. Soc.* **1995**, *117*, 4177−4178.

(47) Zhang, J.; Tuguldur, B.; van der Spoel, D. Force Field Benchmark of Organic Liquids. 2. Gibbs Energy of Solvation. *J. Chem. Inf. Model.* **2015**, *55*, 1192−1201.

(48) Lopes, P. E. M.; Roux, B.; MacKerell, A. D., Jr. Molecular Modeling and Dynamics Studies with Explicit Inclusion of Electronic

Polarization: Theory and Applications. *Theor. Chem. Acc.* **2009**, *124*, 11−28.

(49) Rick, S. W.; Stuart, S. J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, 2002; Vol. *18*.

(50) Cieplak, P.; Dupradeau, F.-Y.; Duan, Y.; Wang, J. Polarization Effects in Molecular Mechanical Force Fields. *J. Phys.: Condens. Matter* **2009**, *21*, 333102.

(51) Ponder, J. W.; Case, D. A. Force Fields for Protein Simulations. *Adv. Protein Chem.* **2003**, *66*, 27−85.

(52) Xie, W.; Gao, J. Design of a Next Generation Force Field: The X-POL Potential. *J. Chem. Theory Comput.* **2007**, *3*, 1890−1900.

(53) Liu, Y.-P.; Kim, K.; Berne, B. J.; Friesner, R. A.; Rick, S. W. Constructing *Ab Initio* Force Fields for Molecular Dynamics Simulations. *J. Chem. Phys.* **1998**, *108*, 4739−4755.

(54) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X.; Murphy, R. B.; Zhou, R.; Halgren, T. A. Development of a Polarizable Force Field for Proteins Via *Ab Initio* Quantum Chemistry: First Generation Model for Gas Phase Tests. *J. Comput. Chem.* **2002**, *23*, 1515−1531.

(55) Ren, P.; Ponder, J. W. Consistent Treatment of Inter- and Intramolecular Polarization in Molecular Mechanics Calculations. *J. Comput. Chem.* **2002**, *23*, 1497−1506.

(56) Jorgensen, W. L.; Jensen, K. P.; Alexandrova, A. N. Polarization Effects for Hydrogen-Bonded Complexes of Substituted Phenols with Water and Chloride Ion. *J. Chem. Theory Comput.* **2007**, *3*, 1987−1992.

(57) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins. *J. Chem. Theory Comput.* **2013**, *9*, 4046−4063.

(58) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Jr.; et al. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 2549−2564.

(59) Rick, S. W.; Stuart, S. J.; Berne, B. J. Dynamical Fluctuating Charge Force Fields: Application to Liquid Water. *J. Chem. Phys.* **1994**, *101*, 6141−6156.

(60) Rick, S. W.; Berne, B. J. Dynamical Fluctuating Charge Force Fields: The Aqueous Solvation of Amides. *J. Am. Chem. Soc.* **1996**, *118*, 672−679.

(61) Stern, H. A.; Rittner, F.; Berne, B. J.; Friesner, R. A. Combined Fluctuating Charge and Polarizable Dipole Models: Application to a Five-Site Water Potential Function. *J. Chem. Phys.* **2001**, *115*, 2237−2251.

(62) Patel, S.; Brooks, C. L., III CHARMM Fluctuating Charge Force Field for Proteins: I Parameterization and Application to Bulk Organic Liquid Simulations. *J. Comput. Chem.* **2004**, *25*, 1−16.

(63) Patel, S.; MacKerell, A. D., Jr.; Brooks, C. L., III CHARMM Fluctuating Charge Force Field for Proteins: II Protein/Solvent Properties from Molecular Dynamics Simulations Using a Nonadditive Electrostatic Model. *J. Comput. Chem.* **2004**, *25*, 1504−1514.

(64) 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.

(65) Kunz, A.-P. E.; van Gunsteren, W. F. Development of a Nonlinear Classical Polarization Model for Liquid Water and Aqueous Solutions: COS/D. *J. Phys. Chem. A* **2009**, *113*, 11570−11579.

(66) Donchev, A. G.; Ozrin, V. D.; Subbotin, M. V.; Tarasov, O. V.; Tarasov, V. I. A Quantum Mechanical Polarizable Force Field for Biomolecular Interactions. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 7829−7834.

(67) Donchev, A. G.; Galkin, N. G.; Illarionov, A. A.; Khoruzhii, O. V.; Olevanov, M. A.; Ozrin, V. D.; Subbotin, M. V.; Tarasov, V. I. Water Properties from First Principles: Simulations by a General-Purpose Quantum Mechanical Polarizable Force Field. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 8613−8617.

(68) Li, X.; Ponomarev, S. Y.; Sigalovsky, D. L.; Cvitkovic, J. P.; Kaminski, G. A. POSSIM: Parameterizing Complete Second-Order Polarizable Force Field for Proteins. *J. Chem. Theory Comput.* **2014**, *10*, 4896−4910.

(69) Wang, J.; Cieplak, P.; Li, J.; Hou, T.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical Calculations I: Parameterization of Atomic Polarizability. *J. Phys. Chem. B* **2011**, *115*, 3091−3099.

(70) Wang, J.; Cieplak, P.; Li, J.; Wang, J.; Cai, Q.; Hsieh, M.-J.; Lei, H.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical Calculations II: Induced Dipole Models Significantly Improve Accuracy of Intermolecular Interaction Energies. *J. Phys. Chem. B* **2011**, *115*, 3100−3111.

(71) Wang, J.; Cieplak, P.; Cai, Q.; Hsieh, M.-J.; Wang, J.; Duan, Y.; Luo, R. Development of Polarizable Models for Molecular Mechanical Calculations. 3. Polarizable Water Models Conforming to Thole Polarization Screening Schemes. *J. Phys. Chem. B* **2012**, *116*, 7999−8008.

(72) Wang, J.; Cieplak, P.; Li, J.; Cai, Q.; Hsieh, M.-J.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical Calculations. 4. Van Der Waals Parameterization. *J. Phys. Chem. B* **2012**, *116*, 7088−7101.

(73) Dang, L. X.; Chang, T.-M. Many-Body Interactions in Liquid Methanol and Its Liquid/Vapor Interface: A Molecular Dynamics Study. *J. Chem. Phys.* **2003**, *119*, 9851−9857.

(74) Wick, C. D.; Dang, L. X. Investigating Pressure Effects on Structural and Dynamical Properties of Liquid Methanol with Many-Body Interactions. *J. Chem. Phys.* **2005**, *123*, 184503.

(75) Chang, T.-M.; Dang, L. X. Computational Studies of [bmim][PF$_6$]/$n$-Alcohol Interfaces with Many-Body Potentials. *J. Phys. Chem. A* **2014**, *118*, 7186−7193.

(76) Dang, L. X. Molecular Dynamics Study of Benzene-Benzene and Benzene-Potassium Ion Interactions Using Polarizable Potential Models. *J. Chem. Phys.* **2000**, *113*, 266−273.

(77) Koneshan, S.; Rasaiah, J.; Dang, L. X. Computer Simulation Studies of Aqueous Solutions at Ambient and Supercritical Conditions Using Effective Pair Potential and Polarizable Potential Models for Water. *J. Chem. Phys.* **2001**, *114*, 7544−7555.

(78) Dang, L. X. A Mechanism for Ion Transport across the Water/Dichloromethane Interface: A Molecular Dynamics Study Using Polarizable Potential Models. *J. Phys. Chem. B* **2001**, *105*, 804−809.

(79) Dang, L. X.; Chang, T.-M.; Panagiotopoulos, A. Z. Gibbs Ensemble Monte Carlo Simulations of Coexistence Properties of a Polarizable Potential Model of Water. *J. Chem. Phys.* **2002**, *117*, 3522−3523.

(80) Wick, C. D.; Kuo, I. F. W.; Mundy, C. J.; Dang, L. X. The Effect of Polarizability for Understanding the Molecular Structure of Aqueous Interfaces. *J. Chem. Theory Comput.* **2007**, *3*, 2002−2010.

(81) Wick, C. D.; Dang, L. X. Molecular Mechanism of Transporting a Polarizable Iodide Anion across the Water-CCl$_4$ Liquid/Liquid Interface. *J. Chem. Phys.* **2007**, *126*, 134702.

(82) Chang, T.-M.; Dang, L. X. Computational Studies of Structures and Dynamics of 1,3-Dimethylimidazolium Salt Liquid and Their Interfaces Using Polarizable Potential Models. *J. Phys. Chem. A* **2009**, *113*, 2127−2135.

(83) Nguyen, P. T.; Nguyen, V. T.; Annapureddy, H. V.; Dang, L. X.; Do, D. D. Thermodynamics and Kinetics of Na$^+$/K$^+$-Formate Ion Pairs Assocation in Polarizable Water: A Molecular Dynamics Study. *Chem. Phys. Lett.* **2012**, *554*, 90−95.

(84) Dang, L. X.; Truong, B. T.; Ginovska-Pangovska, B. Note: Interionic Potentials of Mean Force for Ca$^{2+}$-Cl$^-$ in Polarizable Water. *J. Chem. Phys.* **2012**, *136*, 126101.

(85) Dang, L. X.; Annapureddy, H. V. Computational Studies of Water Exchange around Aqueous Li$^+$ with Polarizable Potential Models. *J. Chem. Phys.* **2013**, *139*, 084506.

(86) Baranyai, A.; Kiss, P. T. A Transferable Classical Potential for the Water Molecule. *J. Chem. Phys.* **2010**, *133*, 144109.

(87) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J.-P. Anisotropic, Polarizable Molecular Mechanics Studies of Inter- and Intramolecular Interactions and Ligand-Macromolecular Complexes. A Bottom-up Strategy. *J. Chem. Theory Comput.* **2007**, *3*, 1960−1986.

(88) Lucas, R. T.; Bauer, B. A.; Patel, S. Charge Equilibration Force Fields for Molecular Dynamics Simulations of Lipids, Bilayers, and Integral Membrane Protein Systems. *Biochim. Biophys. Acta, Biomembr.* **2012**, *1818*, 318−329.

(89) Zhong, Y.; Bauer, B. A.; Patel, S. Solvation Properties of N-Acetyl-$\beta$-Glucosamine: Molecular Dynamics Study Incorporating Electrostatic Polarization. *J. Comput. Chem.* **2011**, *32*, 3339−3353.

(90) Banks, J. L.; Kaminski, G. A.; Zhou, R.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. Parametrizing a Polarizable Force Field from Ab Initio Data. I. The Fluctuating Point Charge Model. *J. Chem. Phys.* **1999**, *110*, 741−754.

(91) Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B. J.; Friesner, R. A. Fluctuating Charge, Polarizable Dipole, and Combined Models: Parameterization from Ab Initio Quantum Chemistry. *J. Phys. Chem. B* **1999**, *103*, 4730−4737.

(92) Zhong, Y.; Patel, S. Binding Structures of Tri-N-Acetyl-$\beta$-Glucosamine in Hen Egg White Lysozyme Using Molecular Dynamics with a Polarizable Force Field. *J. Comput. Chem.* **2013**, *34*, 163−174.

(93) Ou, S.; Patel, S. Temperature Dependence and Energetics of Single Ions at the Aqueous Liquid-Vapor Interface. *J. Phys. Chem. B* **2013**, *117*, 6512−6523.

(94) Hu, Y.; Ou, S.; Patel, S. Free Energetics of Arginine Permeation into Model DMPC Lipid Bilayers: Coupling of Effective Counterion Concentration and Lateral Bilayer Dimensions. *J. Phys. Chem. B* **2013**, *117*, 11641−11653.

(95) Giese, T. J.; York, D. M. Improvement of Semiempirical Response Properties with Charge-Dependent Response Density. *J. Chem. Phys.* **2005**, *123*, 164108.

(96) Stuart, S. J.; Berne, B. J. Effects of Polarizability on the Hydration of the Chloride Ion. *J. Phys. Chem.* **1996**, *100*, 11934−11943.

(97) Schmollngruber, M.; Schröder, C.; Steinhauser, O. Polarization Effects on the Solvation Dynamics of Coumarin C153 in Ionic Liquids: Components and Their Cross-Correlations. *J. Chem. Phys.* **2013**, *138*, 204504.

(98) Schröder, C.; Steinhauser, O. Simulating Polarizable Molecular Ionic Liquids with Drude Oscillators. *J. Chem. Phys.* **2010**, *133*, 154511.

(99) Car, R.; Parrinello, M. Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **1985**, *55*, 2471−2474.

(100) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *J. Phys. Chem. B* **2013**, *117*, 9956−9972.

(101) Straatsma, T. P.; McCammon, J. A. Molecular Dynamics Simulations with Interaction Potentials Including Polarization. Development of a Noniterative Method and Application to Water. *Mol. Simul.* **1990**, *5*, 181−192.

(102) Simmonett, A. C.; Pickard, F. C., IV; Shao, Y.; Cheatham, T. E., III; Brooks, B. R. Efficient Treatment of Induced Dipoles. *J. Chem. Phys.* **2015**, *143*, 074115.

(103) Schmollngruber, M.; Lesch, V.; Schröder, C.; Heuer, A.; Steinhauser, O. Comparing Induced Point-Dipoles and Drude Oscillators. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14297−14306.

(104) 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.

(105) Thole, B. T. Molecular Polarizabilities Calculated with a Modified Dipole Interaction. *Chem. Phys.* **1981**, *59*, 341−350.

(106) Noskov, S. Y.; Lamoureux, G.; Roux, B. Molecular Dynamics of Hydration in Ethanol-Water Mixtures Using a Polarizable Force Field. *J. Phys. Chem. B* **2005**, *109*, 6705−6713.

(107) 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.

(108) Miller, K. J. Additivity Methods in Molecular Polarizability. *J. Am. Chem. Soc.* **1990**, *112*, 8533−8542.

(109) Lipparini, F.; Lagardère, L.; Stamm, B.; Cancès, E.; Schnieders, M.; Ren, P.; Maday, Y.; Piquemal, J.-P. Scalable Evaluation of Polarization Energy and Associated Forces in Polarizable Molecular Dynamics: I. Toward Massively Parallel Direct Space Computations. *J. Chem. Theory Comput.* **2014**, *10*, 1638−1651.

(110) Van Belle, D.; Couplet, I.; Prevost, M.; Wodak, S. J. Calculations of Electrostatic Properties in Proteins: Analysis of Contributions from Induced Protein Dipoles. *J. Mol. Biol.* **1987**, *198*, 721−735.

(111) Lamoureux, G.; Roux, B. Modeling Induced Polarization with Classical Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm. *J. Chem. Phys.* **2003**, *119*, 3025−3039.

(112) Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* **1984**, *81*, 511−519.

(113) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695−1697.

(114) Martyna, G. J.; Tuckerman, M.; Tobias, D. J.; Klein, M. L. Explicit Reversible Integrators for Extended System Dynamics. *Mol. Phys.* **1996**, *87*, 1117−1157.

(115) 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.

(116) 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.; et al. 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.

(117) Eastman, P.; Friedrichs, M. S.; Chodera, J. D.; Radmer, R. J.; Bruns, C. M.; Ku, J. P.; Beauchamp, K. A.; Lane, T. J.; Wang, L.-P.; Shuka, D.; et al. OpenMM 4: A Reusable, Extensible, Hardware Independent Library for High Performance Molecular Simulation. *J. Chem. Theory Comput.* **2013**, *9*, 461−469.

(118) Lemkul, J. A.; Roux, B.; van der Spoel, D.; MacKerell, A. D., Jr. Implementation of Extended Lagrangian Dynamics in GROMACS for Polarizable Simulations Using the Classical Drude Oscillator Model. *J. Comput. Chem.* **2015**, *36*, 1473−1479.

(119) Chowdhary, J.; Harder, E.; Lopes, P. E. M.; 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.

(120) Huang, J.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D., Jr. Recent Advances in Polarizable Force Fields for Macromolecules: Microsecond Simulations of Proteins Using the Classical Drude Oscillator Model. *J. Phys. Chem. Lett.* **2014**, *5*, 3144−3150.

(121) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618−622.

(122) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 Energy Evaluation by Direct Methods. *Chem. Phys. Lett.* **1988**, *153*, 503−506.

(123) Hariharan, P. C.; Pople, J. A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theoret. Chim. Acta (Berl.)* **1973**, *28*, 213−222.

(124) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient Diffuse Function-Augmented Basis Sets for Anion Calculations. III. The 3-21+G Basis Set for First-Row Elements, Li-F. *J. Comput. Chem.* **1983**, *4*, 294−301.

(125) Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D., Jr. Determination of Electrostatic Parameters for a Polarizable Force Field Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2005**, *1*, 153−168.

(126) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098−3100.

(127) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(128) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.

(129) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(130) Giese, T. J.; York, D. M. Many-Body Force Field Models Based Solely on Pairwise Coulomb Screening Do Not Simultaneously Reproduce Correct Gas-Phase and Condensed-Phase Polarizability Limits. *J. Chem. Phys.* **2004**, *120*, 9903.

(131) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A. Development of an Accurate and Robust Polarizable Molecular Mechanics Force Field from Ab Initio Quantum Chemistry. *J. Phys. Chem. A* **2004**, *108*, 621–627.

(132) Morita, A. Water Polarizability in Condensed Phase: *Ab Initio* Evaluation by Cluster Approach. *J. Comput. Chem.* **2002**, *23*, 1466–1471.

(133) Morita, A.; Kato, S. An *Ab Initio* Analysis of Medium Perturbation on Molecular Polarizabilities. *J. Chem. Phys.* **1999**, *110*, 11987.

(134) in het in het Panhuis, M.; Popelier, P. L. A.; Munn, R. W.; Ángyán, J. G. Distributed Polarizability of the Water Dimer: Field-Induced Charge Transfer Along the Hydrogen Bond. *J. Chem. Phys.* **2001**, *114*, 7951.

(135) Tu, Y.; Laaksonen, A. The Electronic Properties of Water Molecules in Water Clusters and Liquid Water. *Chem. Phys. Lett.* **2000**, *329*, 283–288.

(136) Schropp, B.; Tavan, P. The Polarizability of Point-Polarizable Water Models: Density Functional Theory/Molecular Mechanics Results. *J. Phys. Chem. B* **2008**, *112*, 6233–6240.

(137) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(138) Connolly, M. L. Solvent-Accessible Surfaces of Proteins and Nucleic Acids. *Science* **1983**, *221*, 709–713.

(139) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-Set Convergence in Correlated Calculations on Ne, $N_2$, and $H_2O$. *Chem. Phys. Lett.* **1998**, *286*, 243–252.

(140) Neria, E.; Fischer, S.; Karplus, M. Simulation of Activation Free Energies in Molecular Systems. *J. Chem. Phys.* **1996**, *105*, 1902.

(141) 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.

(142) Gregory, J. K.; Clary, D. C.; Liu, K.; Brown, M. G.; Saykally, R. J. The Water Dipole Moment in Water Clusters. *Science* **1997**, *275*, 814–817.

(143) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(144) Yeh, I.-C.; Hummer, G. System-Size Dependence of Diffusion Coefficients and Viscosities from Molecular Dynamics Simulations with Periodic Boundary Conditions. *J. Phys. Chem. B* **2004**, *108*, 15873–15879.

(145) Yu, W.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D., Jr. Six-Site Polarizable Model of Water Based on the Classical Drude Oscillator. *J. Chem. Phys.* **2013**, *138*, 034508.

(146) Mahoney, M. W.; Jorgensen, W. L. A Five-Site Model for Liquid Water and the Reproduction of the Density Anomaly by Rigid, Nonpolarizable Potential Functions. *J. Chem. Phys.* **2000**, *112*, 8910–8922.

(147) DeBolt, S. E.; Kollman, P. A. Investigation of Structure, Dynamics, and Solvation in 1-Octanol and Its Water-Saturated Solution: Molecular Dynamics and Free-Energy Perturbation Studies. *J. Am. Chem. Soc.* **1995**, *117*, 5316–5340.

(148) Koehl, P. Electrostatics Calculations: Latest Methodological Advances. *Curr. Opin. Struct. Biol.* **2006**, *16*, 142–151.

(149) Vorobyov, I. V.; Anisimov, V. M.; MacKerell, A. D., Jr. Polarizable Empirical Force Field for Alkanes Based on the Classical Drude Oscillator Model. *J. Phys. Chem. B* **2005**, *109*, 18988–18999.

(150) 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.

(151) Suzuki, S.; Green, P. G.; Bumgarner, R. E.; Dasgupta, S.; Goddard, W. A., III; Blake, G. A. Benzene Forms Hydrogen Bonds with Water. *Science* **1992**, *257*, 942–945.

(152) Orabi, E. A.; Lamoureux, G. Cation-$\pi$ and $\pi$-$\pi$ Interactions in Aqueous Solution Studied Using Polarizable Potential Models. *J. Chem. Theory Comput.* **2012**, *8*, 182–193.

(153) 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.

(154) Capon, B. Mechanism in Carbohydrate Chemistry. *Chem. Rev.* **1969**, *69*, 407–498.

(155) Dwek, R. A.; Butters, T. D. Introduction: Glycobiology - Understanding the Language and Meaning of Carbohydrates. *Chem. Rev.* **2002**, *102*, 283–284.

(156) Baker, C. M.; MacKerell, A. D., Jr. Polarizability Rescaling and Atom-Base Thole Scaling in the CHARMM Drude Polarizable Force Field for Ethers. *J. Mol. Model.* **2010**, *16*, 567.

(157) 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.

(158) van Erp, T. S.; Meijer, E. J. Ab Initio Molecular Dynamics Study of Aqueous Solvation of Ethanol and Ethylene. *J. Chem. Phys.* **2003**, *118*, 8831–8840.

(159) Lin, B.; He, X.; MacKerell, A. D., Jr. A Comparative Kirkwood-Buff Study of Aqueous Methanol Solutions Modeled by the CHARMM Additive and Drude Polarizable Force Fields. *J. Phys. Chem. B* **2013**, *117*, 10572–10580.

(160) Tran, H. T.; Mao, A.; Pappu, R. V. Role of Backbone-Solvent Interactions in Determining Conformational Equilibria of Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **2008**, *130*, 7380–7392.

(161) Harder, E.; Anisimov, V. M.; Whitfield, T.; 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.

(162) Whitfield, T. W.; Martyna, G. J.; Allison, S.; Bates, S. P.; Vass, H.; Crain, J. Structure and Hydrogen Bonding in Neat *N*-Methylacetamide: Classical Molecular Dynamics and Raman Spectroscopy Studies of a Liquid of Peptidic Fragments. *J. Phys. Chem. B* **2006**, *110*, 3624–3637.

(163) Essex, J. W.; Jorgensen, W. L. Dielectric Constants of Formamide and Dimethylformamide Via Computer Simulation. *J. Phys. Chem.* **1995**, *99*, 17956–17962.

(164) Lin, B.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D., Jr. Kirkwood-Buff Analysis of Aqueous *N*-Methylacetamide and Acetamide Solutions Modeled by the CHARMM Additive and Drude Polarizable Force Fields. *J. Chem. Phys.* **2013**, *139*, 084509.

(165) 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.

(166) 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*, 2330–2341.

(167) Record, M. T.; Anderson, C. F.; Lohman, T. M. Thermodynamic Analysis of Ion Effects on the Binding and Conformational Equilibria of Proteins and Nucleic Acids: The Roles of Ion Association or Release, Screening, and Ion Effects on Water Activity. *Q. Rev. Biophys.* **1978**, *11*, 103–178.

(168) Baldwin, R. L. How Hofmeister Ion Interactions Affect Protein Stability. *Biophys. J.* **1996**, *71*, 2056–2063.

(169) Draper, D. E. RNA Folding: Thermodynamic and Molecular Descriptions of the Roles of Ions. *Biophys. J.* **2008**, *95*, 5489–5495.

(170) Draper, D. E. Folding of RNA Tertiary Structure: Linkages between Backbone Phosphates, Ions, and Water. *Biopolymers* **2013**, *99*, 1105–1113.

(171) Assche, F. V.; Clijsters, H. Effects of Metals on Enzyme Activity in Plants. *Plant, Cell Environ.* **1990**, *13*, 195–206.

(172) Rembert, K. B.; Paterová, J.; Heyda, J.; Hilty, C.; Jungwirth, P.; Cremer, P. S. Molecular Mechanisms of Ion-Specific Effects on Proteins. *J. Am. Chem. Soc.* **2012**, *134*, 10039–10046.

(173) Riahi, S.; Roux, B.; Rowley, C. N. QM/MM Molecular Dynamics Simulations of Hydration of Mg(II) and Zn(II) Ions. *Can. J. Chem.* **2013**, *91*, 552–558.

(174) Lamoureux, G.; Roux, B. Absolute Hydration Free Energy Scale for Alkali and Halide Ions Established from Simulations with a Polarizable Force Field. *J. Phys. Chem. B* **2006**, *110*, 3308–3322.

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

(176) Luo, Y.; Jiang, W.; Yu, H.; 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.

(177) Savelyev, A.; MacKerell, A. D., Jr. Balancing the Interactions of Ions, Water, and DNA in the Drude Polarizable Force Field. *J. Phys. Chem. B* **2014**, *118*, 6742–6757.

(178) 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.

(179) Wada, A. The Alpha-Helix as an Electric Macro-Dipole. *Adv. Biophys.* **1976**, *9*, 1–63.

(180) Li, H.; Ngo, V.; Da Silva, M. C.; Salahub, D. R.; Callahan, K.; Roux, B.; Noskov, S. Y. Representation of Ion-Protein Interactions Using the Drude Polarizable Force Field. *J. Phys. Chem. B* **2015**, *119*, 9401–9416.

(181) 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*, 580–596.

(182) 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.

(183) Duchardt, E.; Nilsson, L.; Schleucher, J. Cytosine Ribose Flexibility in DNA: A Combined NMR $^{13}$C Spin Relaxation and Molecular Dynamics Simulation Study. *Nucleic Acids Res.* **2008**, *36*, 4211–4219.

(184) Manning, G. S. The Molecular Theory of Polyelectrolyte Solutions with Applications to the Electrostatic Properties of Polynucleotides. *Q. Rev. Biophys.* **1978**, *11*, 179–246.

(185) 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.

(186) 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* **2015**, *119*, 637–652.

(187) Jana, M.; MacKerell, A. D., Jr. CHARMM Drude Polarizable Force Field for Aldopentofuranoses and Methyl-Aldopentofuranosides. *J. Phys. Chem. B* **2015**, *119*, 7846–7859.

(188) Harder, E.; MacKerell, A. D., Jr.; Roux, B. Many-Body Polarization Effects and the Membrane Dipole Potential. *J. Am. Chem. Soc.* **2009**, *131*, 2760–2761.

(189) Pickar, A. D.; Benz, R. Transport of Oppositely Charged Lipophilic Probe Ions in Lipid Bilayer Membranes Having Various Structures. *J. Membr. Biol.* **1978**, *44*, 353–376.

(190) Brockman, H. Dipole Potential of Lipid Membranes. *Chem. Phys. Lipids* **1994**, *73*, 57–79.

(191) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D., Jr.; Pastor, R. W. Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* **2010**, *114*, 7830–7843.

(192) Steinbach, P. J.; Brooks, B. R. New Spherical-Cutoff Methods for Long-Range Forces in Macromolecular Simulation. *J. Comput. Chem.* **1994**, *15*, 667–683.

(193) Piggot, T.; Piñeiro, Á.; Khalid, S. Molecular Dynamics Simulations of Phosphatidylcholine Membranes: A Comparative Force Field Study. *J. Chem. Theory Comput.* **2012**, *8*, 4593–4609.

(194) Dobson, C. M. Protein Folding and Misfolding. *Nature* **2003**, *426*, 884–890.

(195) Huang, J.; MacKerell, A. D., Jr. Induction of Peptide Bond Dipoles Drives Cooperative Helix Formation in the (AAQAA)$_3$ Peptide. *Biophys. J.* **2014**, *107*, 991–997.

(196) Shalongo, W.; Dugad, L.; Stellwagen, E. Distribution of Helicity within the Model Peptide Acetyl (AAQAA)$_3$ Amide. *J. Am. Chem. Soc.* **1994**, *116*, 8288–8293.

(197) Rohl, C. A.; Baldwin, R. L. Comparison of NH Exchange and Circular Dichroism as Techniques for Measuring the Parameters of the Helix-Coil Transition in Peptides. *Biochemistry* **1997**, *36*, 8435–8442.

(198) Best, R. B.; Hummer, G. Optimized Molecular Dynamics Force Fields Applied to the Helix-Coil Transition of Polypeptides. *J. Phys. Chem. B* **2009**, *113*, 9004–9015.

(199) Hol, W. G. J.; van Duijnen, P. T.; Berendsen, H. J. C. The $\alpha$-Helix Dipole and the Properties of Proteins. *Nature* **1978**, *273*, 443–446.

(200) Lemkul, J. A.; Huang, J.; MacKerell, A. D. Induced Dipole-Dipole Interactions Influence the Unfolding Pathways of Wild-Type and Mutant Amyloid $\beta$-Peptides. *J. Phys. Chem. B* **2015**, *119*, 15574–15582.

(201) Kirkwood, J. G.; Shumaker, J. B. The Influence of Dipole Moment Fluctuations on the Dielectric Increment of Proteins in Solution. *Proc. Natl. Acad. Sci. U. S. A.* **1952**, *38*, 855–862.

(202) Simonson, T.; Brooks, C. L., III Charge Screening and the Dielectric Constant of Proteins: Insights from Molecular Dynamics. *J. Am. Chem. Soc.* **1996**, *118*, 8452–8458.

(203) Simonson, T.; Perahia, D.; Bricogne, G. Intramolecular Dielectric Screening in Proteins. *J. Mol. Biol.* **1991**, *218*, 859–886.

(204) Krishtalik, L. I. The Medium Reorganization Energy for the Charge Transfer Reactions in Proteins. *Biochim. Biophys. Acta, Bioenerg.* **2011**, *1807*, 1444–1456.

(205) Blumberger, J.; Lamoureux, G. Reorganization Free Energies and Quantum Corrections for a Model Electron Self-Exchange Reaction: Comparison of Polarizable and Nonpolarizable Solvent Models. *Mol. Phys.* **2008**, *106*, 1597–1611.

(206) Savelyev, A.; MacKerell, A. D., Jr. Differential Impact of the Monovalent Ions Li$^+$, Na$^+$, K$^+$, and Rb$^+$ on DNA Conformational Properties. *J. Phys. Chem. Lett.* **2015**, *6*, 212–216.

(207) Savelyev, A.; MacKerell, A. D., Jr. Competition among Li$^+$, Na$^+$, K$^+$, and Rb$^+$ Monovalent Ions for DNA in Molecular Dynamics Simulations Using the Additive CHARMM36 and Drude Polarizable Force Fields. *J. Phys. Chem. B* **2015**, *119*, 4428–4440.

(208) Savelyev, A.; MacKerell, A. D., Jr. Differential Deformability of the DNA Minor Groove and Altered BI/BII Backbone Conformational Equilibrium by the Monovalent Ions Li$^+$, Na$^+$, K$^+$, and Rb$^+$ Via Water-Mediated Hydrogen Bonding. *J. Chem. Theory Comput.* **2015**, *11*, 4473–4485.

(209) Lemkul, J. A.; Savelyev, A.; MacKerell, A. D., Jr. Induced Polarization Influences the Fundamental Forces in DNA Base Flipping. *J. Phys. Chem. Lett.* **2014**, *5*, 2077–2083.

(210) Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. The B-DNA Dodecamer at High Resolution Reveals a Spine of Water on Sodium. *Biochemistry* **1998**, *37*, 8341–8355.

(211) Chiu, T. K.; Kaczor-Grzeskowiak, M.; Dickerson, R. E. Absence of Minor Groove Monovalent Cations in the Crosslinked Dodecamer C-G-C-G-A-A-T-T-C-G-C-G. *J. Mol. Biol.* **1999**, *292*, 589–608.

(212) Bai, Y.; Greenfeld, M.; Travers, K. J.; Chu, V. B.; Lipfert, J.; Doniach, S.; Herschlag, D. Quantitative and Comprehensive Decomposition of the Ion Atmosphere around Nucleic Acids. *J. Am. Chem. Soc.* **2007**, *129*, 14981−14988.

(213) Stellwagen, E.; Dong, Q.; Stellwagen, N. C. Monovalent Cations Affect the Free Solution Mobility of DNA by Perturbing the Hydrogen-Bonded Structure of Water. *Biopolymers* **2005**, *78*, 62−68.

(214) Abrescia, N. G. A.; Thompson, A.; Huynh-Dinh, T.; Subirana, J. A. Crystal Structure of an Antiparallel DNA Fragment with Hoogsteen Base Pairing. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 2806−2811.

(215) Abrescia, N. G. A.; González, C.; Gouyette, C.; Subirana, J. A. X-Ray and NMR Studies of the DNA Oligomer d(ATATAT): Hoogsteen Base Pairing in Duplex DNA. *Biochemistry* **2004**, *43*, 4092−4100.

(216) Kuszewski, J.; Schwieters, C.; Clore, G. M. Improving the Accuracy of NMR Structures of DNA by Means of a Database Potential of Mean Force Describing Base-Base Positional Interactions. *J. Am. Chem. Soc.* **2001**, *123*, 3903−3918.

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

(218) Eichman, B. F.; Vargason, J. M.; Mooers, B. H.; Ho, P. S. The Holliday Junction in an Inverged Repeat DNA Sequence: Sequence Effects on the Structure of Four-Way Junctions. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 3971−3976.

(219) Priyakumar, U. D.; MacKerell, A. D., Jr. NMR Imino Proton Exchange Experiments on Duplex DNA Primarily Monitor the Opening of Purine Bases. *J. Am. Chem. Soc.* **2006**, *128*, 678−679.

(220) Priyakumar, U. D.; MacKerell, A. D., Jr. Base Flipping in a GCGC Containing DNA Dodecamer: A Comparative Study of the Performance of the Nucleic Acid Force Fields, CHARMM, AMBER, and BMS. *J. Chem. Theory Comput.* **2006**, *2*, 187−200.

(221) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31*, 671−690.

(222) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859−1865.

(223) Huang, L.; Roux, B. Automated Force Field Parameterization for Nonpolarizable and Polarizable Atomic Models Based on Ab Initio Target Data. *J. Chem. Theory Comput.* **2013**, *9*, 3543−3556.