

Recent Advances in Polarizable Force Fields for Macromolecules: Microsecond Simulations of Proteins Using the Classical Drude Oscillator Model

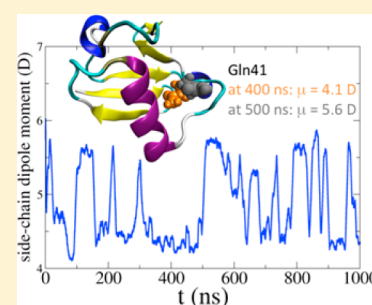
Jing Huang,[†] Pedro E. M. Lopes,[†] Benoît Roux,[‡] and Alexander D. MacKerell, Jr.*[†]

[†]Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, 20 Penn Street, Room 629, Baltimore, Maryland 21201, United States

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

S Supporting Information

ABSTRACT: In this Perspective, we summarize recent efforts to include the explicit treatment of induced electronic polarization in biomolecular force fields. Methods used to treat polarizability, including the induced dipole, fluctuating charge, and classical Drude oscillator models, are presented, including recent advances in force fields using those methods. This is followed by recent results obtained with the Drude model, including microsecond molecular dynamics (MD) simulations of multiple proteins in explicit solvent. Results show significant variability of backbone and side-chain dipole moments as a function of environment, including significant changes during individual simulations. Dipole moments of water in the vicinity of the proteins reveal small but systematic changes, with the direction of the changes dependent on the environment. Analyses of the full proteins show that the polarizable Drude model leads to larger values of the dielectric constant of the protein interior, especially in the case of hydrophobic regions. These results indicate that the inclusion of explicit electronic polarizability leads to significant differences in the physical forces affecting the structure and dynamics of proteins, which can be investigated in a computationally tractable fashion in the context of the Drude model.



These results indicate that the inclusion of explicit electronic polarizability leads to significant differences in the physical forces affecting the structure and dynamics of proteins, which can be investigated in a computationally tractable fashion in the context of the Drude model.

Molecular modeling and simulations are important tools for exploring the structure, dynamics, and function of complex chemical and biochemical systems in condensed phases. Underlying these methods are force fields, which represent a computationally tractable approximation to the Born–Oppenheimer potential energy surface.¹ The functional form of the force fields (FFs) used to simulate macromolecular systems has remained largely unchanged during past decades, though the parameters in those FFs are continually being refined. In the widely used macromolecular FFs, the electrostatics term is described with Coulombic interactions between fixed partial atomic point charges, referred to as an additive FF. This simplification is used largely to reduce the computational cost as well as to facilitate parametrization. In such additive fixed-charge FFs, electronic polarization effects are considered in a mean-field, average manner by empirically optimizing the partial atomic charges to overestimate dipole moments during the FF parametrization, thereby being representative of the condensed, typically aqueous, phase. While additive FFs have seen considerable use, the additive approximation significantly limits the accuracy of the method, for example, in treating highly polar versus hydrophobic environments. To achieve a more accurate description of the response of the charge distribution to variations in the surrounding electrostatic field, the explicit inclusion of electronic polarizability in the model is essential.

There are a variety of methods to include explicit electronic polarizability in FFs. An obvious approach is to treat the

To achieve a more accurate description of the response of the charge distribution to variations in the surrounding electrostatic field, the explicit inclusion of electronic polarizability in the model is essential.

electronic response quantum mechanically. While this has been performed to treat local regions of macromolecular systems in the context of QM/MM methods^{2–4} and preliminary FFs that include the treatment of electrostatics using semiempirical methods have been presented (e.g., XPOL),⁵ these are computationally demanding approaches. The alternative is a more simple, computationally tractable extension of the potential energy function to mimic the quantum mechanical response of the electronic distribution to changes in the surrounding electric field. One way to perform this is to assign an induced point dipole at each atomic site.^{2,6–11} Typically, the induced dipoles are determined by a self-consistent field (SCF) iterative procedure using isotropic atomic polarizabilities,

Received: June 26, 2014

Accepted: August 27, 2014

Published: August 27, 2014

following which the charge–charge, charge–dipole, and dipole–dipole interactions are computed to yield the electrostatic energy of the system. The computational bottleneck of induced dipole models is usually the SCF evaluation of mutual polarization. Recently, Wang et al. proposed an iAMEOBA model for water in which a single SCF step is performed with induced dipoles initially set to zero (ie. the induced dipoles are only determined by the electric field associated with the fixed charges or multipoles).¹² However, this approach neglects mutual polarization between the induced dipoles in the system, though it has been indicated that this contribution may be modeled in an average manner during the parametrization, an approach analogous to that performed with current additive FFs. Using this approach, the iAMEOBA water model was shown to outperform the original AME OBA model in reproducing experimental measurements, although this is likely due to a parametrization algorithm that directly targets condensed-phase properties.^{12,13} While the iAMEOBA approach is likely satisfactory for homogeneous systems, there are legitimate concerns that such an approximation will represent a significant compromise in accuracy in more polar, heterogeneous systems such as ion solvation.

Another approach to treat polarizability is the fluctuating charge model,^{14–18} which treats partial atomic charges as dynamical variables based on the electronegativity at each atomic site. Charges are propagated according to the principle of electronegativity equalization,¹⁹ and charge conservation constraints are assumed. Usually, charges are only allowed to redistribute within molecules to avoid unphysical intermolecular charge transfer, and an extended Lagrangian approach is adopted to propagate the charge variables during molecular dynamics (MD) simulations. One limitation of the model is the inability to describe out-of-plane polarization for a planar system such as benzene, although it is possible to add auxiliary out-of-plane charge sites to overcome this limitation. The CHEQ FF is based on fluctuating charges, with parameters presented for proteins,^{17,18} lipids,²⁰ and select carbohydrates.²¹ In addition, CHEQ has been applied to ligand binding to lysozyme,²² ion solvation,²³ and lipid bilayer permeability.²⁴

A third approach for the explicit treatment of polarization, which has been developed in our groups, is the classical Drude oscillator model.^{25,26} The Drude model has also been referred to as the Shell²⁷ or the Charge-On-Spring model.²⁸ Briefly, a charged auxiliary particle (the Drude oscillator or particle) is attached to the atomic core of its parent atom via a harmonic spring with force constant k_D . The displacement of the Drude particle relative to its parent atom in the presence of an electric field E is given by

$$\mathbf{d} = \frac{q_D \mathbf{E}}{k_D} \quad (1)$$

where q_D is the charge of the Drude particle. The induced dipole is

$$\boldsymbol{\mu} = q_D \mathbf{d} = \left(\frac{q_D^2}{k_D} \right) \mathbf{E} \quad (2)$$

which is equivalent to an atomic polarizability of $\alpha = q_D^2/k_D$.

In our Drude FF, the Drude particles are only associated with non-hydrogen atoms, which has been shown to reproduce molecular polarizabilities while maintaining computational efficiency.²⁹ For the sake of simplicity, a fixed value of 1000

kcal/mol/Å² is used for the restoring force constant, k_D , such that the charge q_D is the parameter that governs the magnitude of α for a given non-hydrogen atom. While eq 2 describe isotropic atomic polarizability, anisotropy can be incorporated in the Drude model by considering $(1/k_D)$ as a tensor, that is, attributing an inhomogeneous spring constant based on a local reference frame for the virtual spring between the atom and its Drude particle. Such a treatment is implemented in the current Drude FF but only applied to hydrogen bond acceptors such as the peptide backbone carbonyl oxygen atoms. This and virtual particles representative of lone pairs are included to improve the treatment of nonbonded interactions as a function of orientation involving hydrogen bond acceptors.

An essential aspect of the Drude FF is that the dipole–dipole interactions between atoms involved in bond or valence angles (1–2 or 1–3 interactions) are explicitly included. However, for 1–2 and 1–3 interactions, Coulomb's Law fails due to the short spatial separation. This problem may be overcome by applying an electrostatic shielding factor S_{ij} to the electrostatic interactions as proposed by Thole³⁰

$$S_{ij}(r_{ij}) = 1 - \left(1 + \frac{(a_i + a_j)r_{ij}}{2(\alpha_i\alpha_j)^{1/6}} \right) e^{-(\alpha_i + \alpha_j)r_{ij}/2(\alpha_i\alpha_j)^{1/6}} \quad (3)$$

where r_{ij} is the distance between the charge site i and j and α_i , α_j and a_i , a_j are the polarizabilities and Thole factors, respectively. The atom-based Thole factors introduced in the Drude model provide fine-tuning of near-field electrostatics, yielding improvements in the treatment of the orientation of molecular polarizabilities. Thus, atomic polarizabilities and Thole factors are FF parameters that, in addition to partial charges, need be optimized during the parametrization of the Drude polarizable FF. While the remainder of the potential energy function, including the bonded terms and the van der Waals (vdW) interaction, is identical to the CHARMM additive FF, the associated parameters also need to be optimized. This is due to the coupling of all terms in the energy function, such that changes in the electrostatic model require reoptimization of the vdW and internal bonded parameters to produce a carefully balanced empirical FF.

Parametrization started with the water model, initially the SWM4-DP model²⁵ followed by the SWM4-NDP model,³¹ with the latter becoming hereafter the default model for all subsequent developments of the Drude FF. Parameters were then derived for model compounds that represent components of biomolecules. While the parametrization of the Drude FF shares similar target data with the CHARMM additive FF, additional target data such as the molecular polarizability were included to facilitate optimization of the electrostatic parameters. In general, the more sophisticated physical model and associated additional FF parameters in the Drude FF allow better reproduction of these target data including both gas-phase QM data and condensed-phase experimental data, an example being the treatment of polyalcohols.³² Also, during the optimization, the atomic polarizabilities were scaled to reproduce pure solvent experimental dielectric constants,^{33–35} yielding scaling factors ranging from 0.6 to 1.0, such that the Drude polarizabilities are equal to or lower than gas-phase experimental or QM values. Due to the nonadditive nature of the Drude FF, particular care was taken when transferring the parameters from model compounds to biomolecules, and often, target data for relatively large model compounds need to be included for further optimization. For example, the gas-phase

relative energies between α R, PPII, and C5 conformations of the (Ala)₅ peptide and interactions of water with the alanine dipeptide are included as additional target data for optimizing the parameters for the polypeptide backbone.

An essential feature of an empirical FF is computational efficiency allowing for simulations of macromolecules with an explicit solvent representation on time scales of 100s of ns through microseconds. With the Drude model, this has been attained by maintaining the simplicity of the potential energy function described above as well as the implementation of an extended Lagrangian integrator with a dual thermostat, allowing for computationally efficient MD simulations while maintaining approximate SCF treatment of the polarizable degrees of freedom.²⁶ Unlike extended Lagrangians used in Carr–Parrinello (CP) MD or fluctuating charge model simulations, where fictitious masses are attributed to the electronic degrees of freedom, the Drude particle may be treated as a real particle that moves in space, allowing it to be assigned a mass, typically 0.4 amu, taken from its parent atom. In the dual-thermostat algorithm, the relative motion of each Drude–nucleus oscillator pair is coupled to a low-temperature thermostat (typically 1 K), such that the electronic degrees of freedom approach the adiabatic SCF limit during the MD simulation. The remainder of the system, including the center-of-mass motion of the Drude–nucleus pairs, is thermostated to the target temperature of the simulation (e.g., room temperature). To avoid polarization catastrophe, the distance d between the Drude particle and its parent atom is limited to typically 0.2 Å by imposing a hard wall constraint to d during the MD simulation.³⁶ This combination of a simple form of the potential energy function and an extended Lagrangian integrator implemented in an efficient parallelizable code, such as CHARMM, NAMD,³⁷ and ChemShell QM, has allowed for MD simulations on the order of 100s of ns on proteins, lipids, and DNA.^{38,39} With NAMD, as compared to the additive FF, the computational overhead with the Drude model is approximately two-fold, which, when using a 1 versus 2 fs integration time step, yields an overall four-fold computational increase with the Drude model. In the remainder of this Perspective, we will present results from fully polarizable microsecond MD simulations of proteins using the Drude FF along with comparisons with the CHARMM36 (C36) additive protein FF.

Presented in Figure 1 are protein $C\alpha$ (root-mean square deviations) RMSDs from 1 μ s MD simulations of ubiquitin (1UBQ) and cold shock protein A (CspA, 1MJC) using the Drude and the C36 FFs. MD simulations were carried out in the *NPT* ensemble with NAMD, with simulation details provided in the Supporting Information. For the Drude simulations, a shorter time step of 1 fs was used versus 2 fs for the additive FF, which is due to the high-frequency motion of the Drude particles related to their small masses. Drude simulations were also performed on crambin (1EJG), the tight junction regulatory protein (3VQF), circular permutant of ribosomal protein S6 (3ZZP), DNA methyltransferase associated protein (4IEJ), and protein G B3 domain (1P7E), with RMSD plotted in Figure S1 of the Supporting Information. The RMSDs of both ubiquitin and CspA were stable, and those from the Drude and C36 simulations were comparable. No systematic drift was observed, indicating the ability of Drude FF to maintain the protein folded structures on the microsecond time scale.

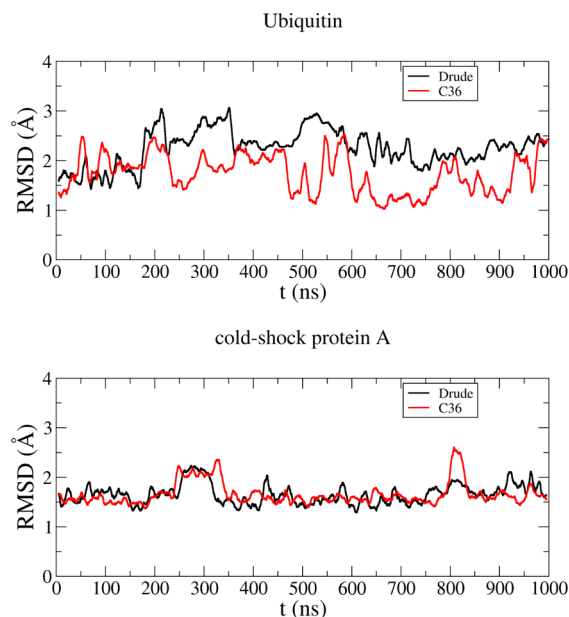


Figure 1. RMSD plots of 1 μ s simulations of ubiquitin and cold-shock protein A. RMSDs were computed for $C\alpha$ atoms in all residues, and results are presented as running 10 ns averages.

The implication is that an explicit treatment of induced polarization captures local variations in the environment that are not realistically accounted for with an additive mean-field approximation.

With polarizable FFs, the charge distribution of functional groups (i.e., dipole moments) is dominated by the electrostatic environment, which is fundamentally different from additive FFs where the dipoles only vary due to changes in the intramolecular geometry of the individual functional groups. This can be illustrated by examining dipole moments of the peptide backbone and protein side chains during the protein simulations. The backbone peptide bond group is defined as the C and O atoms of residue i and the N, H, $C\alpha$, $H\alpha$ atoms of residue $i + 1$, yielding neutral charge in both FFs. The Drude simulations yield a mean value of 4.74 ± 0.31 D for peptide bonds in the helices and 5.14 ± 0.30 D for those in sheets, where the errors are the RMS fluctuations. The C36 simulations yield much narrower distributions, with mean values and RMS fluctuations of 3.82 ± 0.11 D for helical peptide bonds and 3.71 ± 0.09 D for sheet peptide bonds. The larger magnitude dipoles observed with the Drude model is notable because the parametrization of additive FFs typically involves systematic overestimation of dipole moments relative to that of the gas phase in an attempt to capture the effect of the environment in an average mean-field way.⁴⁰ While the enhanced favorable interactions associated with the larger dipoles are partly canceled by the positive self-energy for polarizing the atoms, the implication is that an explicit treatment of induced polarization captures local variations in the environment that are not realistically accounted for with an additive mean-field approximation. Similar results have been

obtained for the dipole moments of nucleic acid bases in duplex DNA.³⁹ The ability of peptide bond dipoles to dynamically respond to the external electric field is likely to be important for the dynamics of protein and peptides. Recently, the Drude FF was found to reproduce the cooperativity of helix formation in the acetyl-(AAQAA)₃-NH₂ peptide, and such folding cooperativity was shown to be associated with enhanced dipole moments of the peptide backbone upon helix formation.⁴¹

The dipole moments of amino acid side chains are also typically larger in the Drude model as compared to those for the additive FF, although this varies based on residue type, as illustrated in Figures S2 and S3 of the Supporting Information. Exceptions occur with nonpolar residues such as Phe, Val, and Ile, with their side-chain dipole magnitudes being similar between the two models. Notably, the polarizable FF shows wide variability between residues of a given type during the MD simulations as well as in individual residues themselves, as previously observed for Trp residues in lysozyme.³⁸ An example is shown in Figure 2, where dipole moments for four Gln residues in ubiquitin are plotted as a function of time. Transitions in dipole moments on a time scale of tens of nanoseconds are observed for the buried Gln41 residue, correlated with the side chain rotating about its χ_1 dihedral angle (Figure 2C). Similar rotation also occurs in the additive

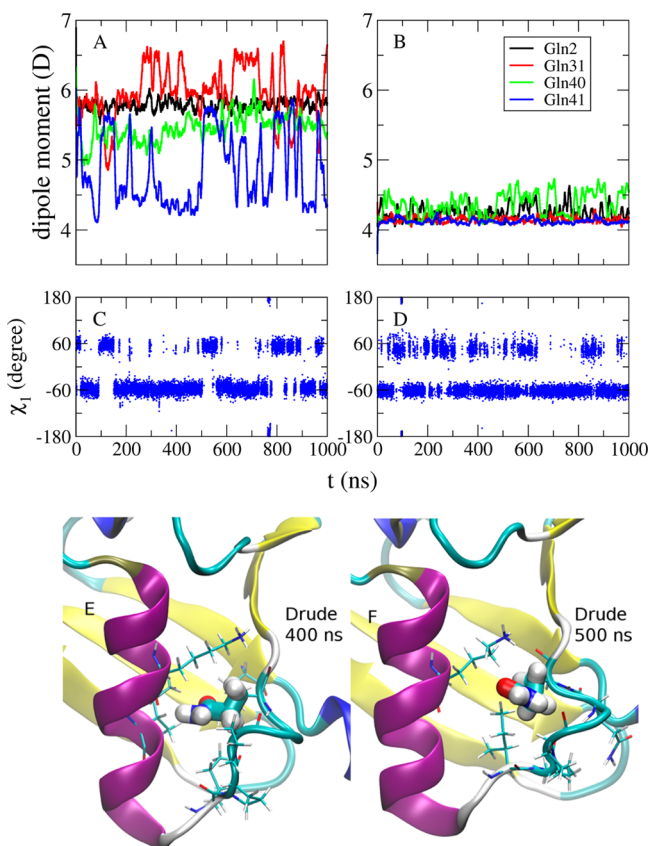


Figure 2. Side-chain dipole moments of glutamine residues in ubiquitin during 1 μ s MD simulations with the Drude (A) and the C36 (B) FFs. The χ_1 dihedral angles for Gln41 are plotted for the Drude (C) and the C36 (D) simulations, and images of Gln41 (Licorice) in ubiquitin (cartoon) are shown from the 400 (E) and 500 ns (F) time points. Residues containing atoms within 3 Å of the Gln41 side chain are shown in lines. Dipole moments are running averages over 10 ns.

C36 simulation (Figure 2D); however, due to the fixed charge nature, no significant changes in the dipole moment are observed. Images of the two environments of Gln41 sampled in the Drude simulation (Figure 2E and F) show that its side chain carries a smaller dipole moment when pointing toward a helix and larger dipole moment when it forms a hydrogen bond with the carbonyl oxygen of Pro38 in a loop region. In contrast, the dipole moment of surface residue Gln2 is stable throughout both the Drude and C36 simulations. However, the Drude model yields a much larger absolute value (5.8 D) than C36 (4.2 D).

The distribution of water dipole moments during the Drude simulations was also analyzed. For the Drude simulation, the SWM4-NDP water model has a dipole moment of 1.85 D in the gas phase and a much larger average value of 2.46 D in the bulk phase (black dashed lines in Figure 3) due to the mutual

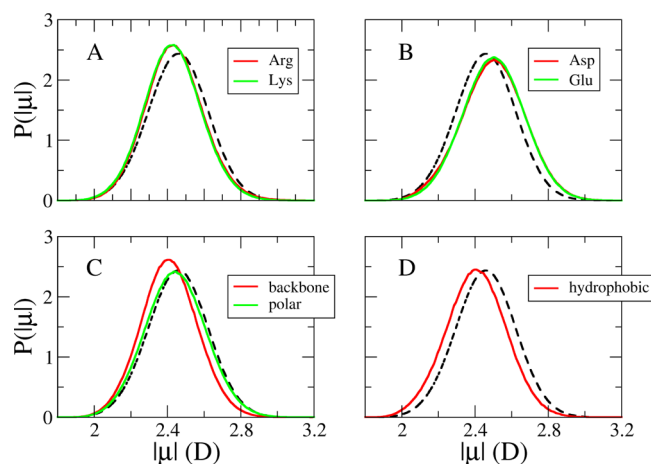


Figure 3. Probability density distribution of water dipole moments in bulk (black lines) and in the first solvation shell in the vicinity of (A) arginine and lysine side chain groups, (B) aspartic acid and glutamic acid side chain groups, (C) peptide backbone groups and polar (Gln, Thr, Asn, and Ser) side-chain groups, and (D) hydrophobic (Val, Ala, Phe, Trp, Tyr, Met, Ile, and Leu) side-chain groups.

polarization. The additive TIP3P water model has a fixed dipole moment of 2.35 D. RMS fluctuation of the bulk water dipole moments in the polarizable FF is 0.16 D, similar to the value of 0.22 D⁴² obtained from CPMD simulations using the Bader approach.⁴³ The distribution of water dipole moments in bulk is compared with those in the first solvation shell of selected residues in Figure 3, as defined as any water within 2.4 Å of any non-hydrogen atom in the protein. Larger dipole moments are observed for water in the proximity of negatively charged Asp and Glu side chains, while smaller values are obtained near positively charged Arg and Lys. Water close to the peptide backbone groups and the hydrophobic side chains also carries slightly smaller dipole moments compared to those of bulk water. Such a difference in the dipole distributions, though small as compared to water in the bulk phase, indicates that water can probe and respond to the complex electric environment of the protein in the polarizable FF. Ab initio MD simulations and QM calculations of ion solvation in water found that the average dipole moment of water in the first hydration shell of K⁺ is 0.22 D smaller than the average bulk value, and those around Cl⁻ are 0.11 D smaller.^{44,45} The importance of variations in water dipoles were shown in a recent study on DNA base flipping using the Drude DNA FF,

Table 1. Protein Dielectric Constants Computed from MD Simulations^a

	ϵ^{inf}			ϵ			
	Drude	entire protein		protein interior		hydrophobic core	
		Drude	C36	Drude	C36	Drude	C36
1UBQ	2.10 ± 0.00	10.3 ± 1.1	11.3 ± 1.1	2.6 ± 0.1	2.6 ± 0.0	2.1 ± 0.1	1.4 ± 0.0
1MJC	2.00 ± 0.00	10.8 ± 0.8	9.0 ± 0.5	3.6 ± 0.4	2.5 ± 0.1	4.6 ± 0.5	2.8 ± 0.2
1EJG	1.76 ± 0.00	2.5 ± 0.1	2.1 ± 0.0	2.7 ± 0.0	1.8 ± 0.1	2.2 ± 0.0	1.6 ± 0.1
3VQF	2.10 ± 0.00	7.0 ± 0.5	5.7 ± 0.4	3.1 ± 0.2	3.2 ± 0.2	3.7 ± 0.3	2.1 ± 0.1
3ZZP	1.91 ± 0.00	8.3 ± 0.3	5.8 ± 0.3	5.0 ± 0.2	3.1 ± 0.2	2.4 ± 0.2	1.6 ± 0.1
1P7E	1.91 ± 0.00	8.7 ± 0.4	5.2 ± 0.4	4.3 ± 0.3	1.8 ± 0.0	1.2 ± 0.0	1.0 ± 0.0

^aOptical dielectric constants ϵ^{inf} were computed from the protein polarizability according to the Clausius–Mosotti equation, and the molecular polarizability tensor for a given conformation was computed by applying an external electric field along the x , y , and z directions, relaxing the Drude particles and then evaluating the changes in dipole moments. Static dielectric constants of entire proteins, the protein interior (residues with a relative SASA less than 50%), and the hydrophobic core (residues with a relative SASA less than 20%) were computed from the Drude and the C36 simulations. Proteins were modeled as spheres whose radius r is given by the radius of gyration R_g multiplied by $(5/3)^{1/2}$.⁴⁷

where changes in water dipole moments in the first solvation shell around the flipping bases occur during the transition from the Watson–Crick base paired to flipped conformations.⁴⁶

We also examined the dipole moments and the molecular polarizabilities of the entire proteins, which are closely related to their dielectric properties (Table 1). The traces of protein polarizability tensors during the 1 μ s Drude simulations for ubiquitin and CspA are shown in Figure S4 of the Supporting Information, showing them to be stable along the 1 μ s MD trajectories, with fluctuation occurring on the nanosecond time scale. Dielectric constants obtained by modeling the proteins as spheres with radii based on the radius of gyration⁴⁷ are summarized in Table 1. The average ϵ^{inf} over the six proteins considered is 2.0 in the Drude model, which equals the commonly assumed value.^{47,48} In contrast, some studies have suggested that a higher value should be used due to the aromatic groups in proteins or the higher density of proteins.^{49,50} However, the present results indicate that this is not necessary. While ϵ^{inf} among the different proteins is similar, their variation will be important when studying properties dominated by the protein dielectric relaxation, for example, as when computing the reorganization energy for electron transfer in proteins using Marcus theory.

Additional analysis involved estimates of the static dielectric constant ϵ of different regions of the proteins from Kirkwood g factors based on the fluctuation of protein dipole moments as presented by Simonson and Brooks.^{47,51} The protein interior is defined using the relative solvent-accessible surface area (SASA) as a criterion, and the computation of ϵ 's is detailed in the Supporting Information. As shown in Table 1, dielectric constants vary across the different proteins. The small value of 2.5 for crambin is consistent with the fact that this protein is not soluble in water. For the other proteins, ϵ ranges from 7 to 12, much larger than the value of 2–5 from the measurements on dry protein powders⁵² and consistent with previous calculations from MD simulations using additive FFs.⁴⁷ Notably, ϵ values of the protein interiors are much smaller than the ϵ of the entire protein, and the Drude model yields dielectric constants that are larger or similar to the C36 model. For the protein hydrophobic cores, defined as residues with a relative SASA less than 20%, the Drude dielectrics are systematically larger than those from C36. These results are consistent with ϵ values for the entire proteins being dominated by charged surface residues,¹⁵ while for the buried residues in the core, additional electronic degrees of freedom of the charge distribution become important. Finally, we note that the

dielectric constant is a macroscopic quantity, and estimating its value in microscopic regions of proteins, which are anisotropic and inhomogeneous, remains challenging.^{53,54} Furthermore, ϵ is dependent on the actual protein site being considered and how the site is defined.^{53,55} This analysis highlights the importance of using an explicit, fully polarizable model to obtain a detailed and accurate description of the response of proteins to their own electric fields as well as those of the surrounding environment.

The Drude model incorporates a more physically realistic treatment of electronic polarizability versus additive force fields and yet maintains a computational efficiency that is crucial for a force field, as illustrated here by microsecond simulations of globular proteins.

In this Perspective, we summarized results on protein simulations using the Drude-2013 polarizable FF. The Drude model incorporates a more physically realistic treatment of electronic polarizability versus additive FFs and yet maintains a computational efficiency that is crucial for a FF, as illustrated here by microsecond simulations of globular proteins. The Drude polarizable protein model, along with currently available parameters for DNA,^{39,56} dipalmitoylphosphatidylcholine,⁵⁷ and hexapyranoses,⁵⁸ as well as ongoing efforts to parametrize RNA and more lipid and carbohydrate molecules, will allow for studies of heterogeneous biomolecular systems using a fully polarizable FF.

■ ASSOCIATED CONTENT

📄 Supporting Information

Simulation and analysis details, time series of protein RMSDs, and time series of side-chain dipole moments. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: alex@outerbanks.umaryland.edu. Phone: (410) 706-7442. Fax: (410) 706-5017.

Notes

The authors declare no competing financial interest.

Biographies

Jing Huang received a B.Sc. in 2005 and a M.Sc. in 2007 from the Department of Physics at Tsinghua University and a Ph.D. in Chemistry from the University of Basel under the supervision of Markus Meuwly in 2011. He is now a postdoctoral fellow with Alex MacKerell in the Computer-Aided Drug Design Center at the University of Maryland, Baltimore.

Pedro E. M. Lopes received a degree in chemical engineering from the Instituto Superior Técnico in 1992 and Ph.D. in Computational Chemistry from Instituto de Tecnologia Química e Biológica at Oeiras, Portugal in 2000. He was a postdoctoral researcher at the University of Manchester from 2000 to 2003 and a Center of Disease Control Regular Fellow in Morgantown, WV and Baltimore, MD from 2003 to 2004. He was subsequently a postdoctoral fellow and research specialist at the University of Maryland, Baltimore and became a research assistant professor in the UMB School of Pharmacy in 2011.

Benoit Roux received a B.Sc. in Physics in 1981 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 Professor in the Department of Biochemistry and Molecular Biology at the University of Chicago with joint appointments as Professor in the Department of Chemistry and as Senior Computational Biologist at Argonne National Laboratory.

Website: <http://thallium.bsd.uchicago.edu/RouxLab/>

Alex MacKerell received a Ph.D. in Biochemistry in 1985 from Rutgers University, followed by postdoctoral fellowships in Medical Biophysics, Karolinska Institutet, Stockholm from 1986 to 1988 and in Chemistry, Harvard University from 1988 to 1992. He is currently in the School of Pharmacy, University of Maryland where he is the Grollman–Glick Professor of Pharmaceutical Sciences and Director of the Computer-Aided Drug Design Center.

Website: http://mackerell.umaryland.edu/MacKerell_Lab.html

ACKNOWLEDGMENTS

Financial support from the NIH (GM051501 and GM072558) and computational support from the University of Maryland Computer-Aided Drug Design Center and the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation Grant Number OCI-1053575, are acknowledged. J.H. acknowledges support by a SNSF Fellowship (PBBSP2_144301).

REFERENCES

- (1) MacKerell, A. D., Jr. Empirical Force Fields for Biological Macromolecules: Overview and Issues. *J. Comput. Chem.* **2004**, *25*, 1584–1604.
- (2) 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.
- (3) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
- (4) Lin, H.; Truhlar, D. G. QM/MM: What Have We Learned, Where Are We, and Where Do We Go from Here? *Theor. Chem. Acc.* **2007**, *117*, 185–199.
- (5) Xie, W.; Gao, J. Design of a Next Generation Force Field: The X-Pol Potential. *J. Chem. Theory Comput.* **2007**, *3*, 1890–1900.

- (6) 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.

- (7) 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 and Gas Phase Tests. *J. Comput. Chem.* **2002**, *23*, 1515–1531.

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

- (9) Xie, W.; Pu, J.; MacKerell, A. D.; Gao, J. Development of a Polarizable Intermolecular Potential Function (PIPF) for Liquid Amides and Alkanes. *J. Chem. Theory Comput.* **2007**, *3*, 1878–1889.

- (10) 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.

- (11) 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.

- (12) 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.

- (13) Wang, L.-P.; Martinez, T. J.; Pande, V. S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 1885–1891.

- (14) 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.

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

- (16) 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.

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

- (18) Patel, S.; MacKerell, A. D., Jr.; Brooks, C. L. 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.

- (19) Pauling, L.; Yost, D. M. The Additivity of the Energies of Normal Covalent Bonds. *Proc. Natl. Acad. Sci. U.S.A.* **1932**, *18*, 414.

- (20) Lucas, T. R.; 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.

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

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

- (23) 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.

- (24) 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.

- (25) 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.

- (26) Lamoureux, G.; Roux, B. Modeling Induced Polarization with Classical Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm. *J. Chem. Phys.* **2003**, *119*, 3025–3039.
- (27) 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.
- (28) 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.
- (29) 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.
- (30) Thole, B. T. Molecular Polarizabilities Calculated with a Modified Dipole Interaction. *Chem. Phys.* **1981**, *59*, 341–350.
- (31) 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.
- (32) He, X.; Lopes, P. E.; Mackerell, A. D., Jr. Polarizable Empirical Force Field for Acyclic Polyalcohols Based on the Classical Drude Oscillator. *Biopolymers* **2013**, *99*, 724–38.
- (33) 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.
- (34) 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.
- (35) Baker, C. M.; MacKerell, A. D., Jr. Polarizability Rescaling and Atom-Based Thole Scaling in the CHARMM Drude Polarizable Force Field for Ethers. *J. Mol. Model.* **2010**, *16*, 567–576.
- (36) 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.
- (37) 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.* **2010**, *2*, 87–92.
- (38) 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.
- (39) 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.
- (40) Cerutti, D. S.; Rice, J. E.; Swope, W. C.; Case, D. A. Derivation of Fixed Partial Charges for Amino Acids Accommodating a Specific Water Model and Implicit Polarization. *J. Phys. Chem. B* **2013**, *117*, 2328–2338.
- (41) Huang, J.; MacKerell, A. D. Induction of Peptide Bond Dipoles Drives Cooperative Helix Formation in the (AAQAA)₃ Peptide. *Biophys. J.* **2014**, *107*, 991–997.
- (42) Dyer, P. J.; Cummings, P. T. Hydrogen Bonding and Induced Dipole Moments in Water: Predictions from the Gaussian Charge Polarizable Model and Car–Parrinello Molecular Dynamics. *J. Chem. Phys.* **2006**, *125*, 144519.
- (43) Bader, R. F. W. *Atoms in Molecules*; Wiley Online Library: New York, 1990.
- (44) Bucher, D.; Kuyucak, S. Polarization of Water in the First Hydration Shell of K⁺ and Ca²⁺ Ions. *J. Phys. Chem. B* **2008**, *112*, 10786–10790.
- (45) Zhao, Z.; Rogers, D. M.; Beck, T. L. Polarization and Charge Transfer in the Hydration of Chloride Ions. *J. Chem. Phys.* **2010**, *132*, 014502.
- (46) 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.
- (47) Simonson, T.; Brooks, C. L. Charge Screening and the Dielectric Constant of Proteins: Insights from Molecular Dynamics. *J. Am. Chem. Soc.* **1996**, *118*, 8452–8458.
- (48) Simonson, T.; Perahia, D.; Bricogne, G. Intramolecular Dielectric Screening in Proteins. *J. Mol. Biol.* **1991**, *218*, 859–886.
- (49) Krishtalik, L. I.; Kuznetsov, A. M.; Mertz, E. L. Electrostatics of Proteins: Description in Terms of Two Dielectric Constants Simultaneously. *Proteins: Struct., Funct., Bioinf.* **1997**, *28*, 174–182.
- (50) Krishtalik, L. I. The Medium Reorganization Energy for the Charge Transfer Reactions in Proteins. *Biochim. Biophys. Acta, Bioenerg.* **2011**, *1807*, 1444–1456.
- (51) Goh, G. B.; García-Moreno E, B.; Brooks, C. L. The High Dielectric Constant of Staphylococcal Nuclease Is Encoded in Its Structural Architecture. *J. Am. Chem. Soc.* **2011**, *133*, 20072–20075.
- (52) Rosen, D. Dielectric Properties of Protein Powders with Adsorbed Water. *Trans. Faraday Soc.* **1963**, *59*, 2178–2191.
- (53) King, G.; Lee, F. S.; Warshel, A. Microscopic Simulations of Macroscopic Dielectric Constants of Solvated Proteins. *J. Chem. Phys.* **1991**, *95*, 4366–4377.
- (54) Schutz, C. N.; Warshel, A. What Are the Dielectric “Constants” of Proteins and How to Validate Electrostatic Models? *Proteins: Struct., Funct., Bioinf.* **2001**, *44*, 400–417.
- (55) Guest, W. C.; Cashman, N. R.; Plotkin, S. S. A Theory for the Anisotropic and Inhomogeneous Dielectric Properties of Proteins. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6286–6295.
- (56) Savelyev, A.; MacKerell, A. D. Balancing the Interactions of Ions, Water, and DNA in the Drude Polarizable Force Field. *J. Phys. Chem. B* **2014**, *118*, 6742–6757.
- (57) Chowdhary, J.; Harder, E.; Lopes, P. E.; Huang, L.; MacKerell, A. D., Jr.; Roux, B. A Polarizable Force Field of Dipalmitoylphosphatidylcholine Based on the Classical Drude Model for Molecular Dynamics Simulations of Lipids. *J. Phys. Chem. B* **2013**, *117*, 9142–60.
- (58) Patel, D. S.; He, X.; MacKerell, A. D., Jr. Polarizable Empirical Force Field for Hexopyranose Monosaccharides Based on the Classical Drude Oscillator. *J. Phys. Chem. B* **2014**, DOI: 10.1021/jp412696m.