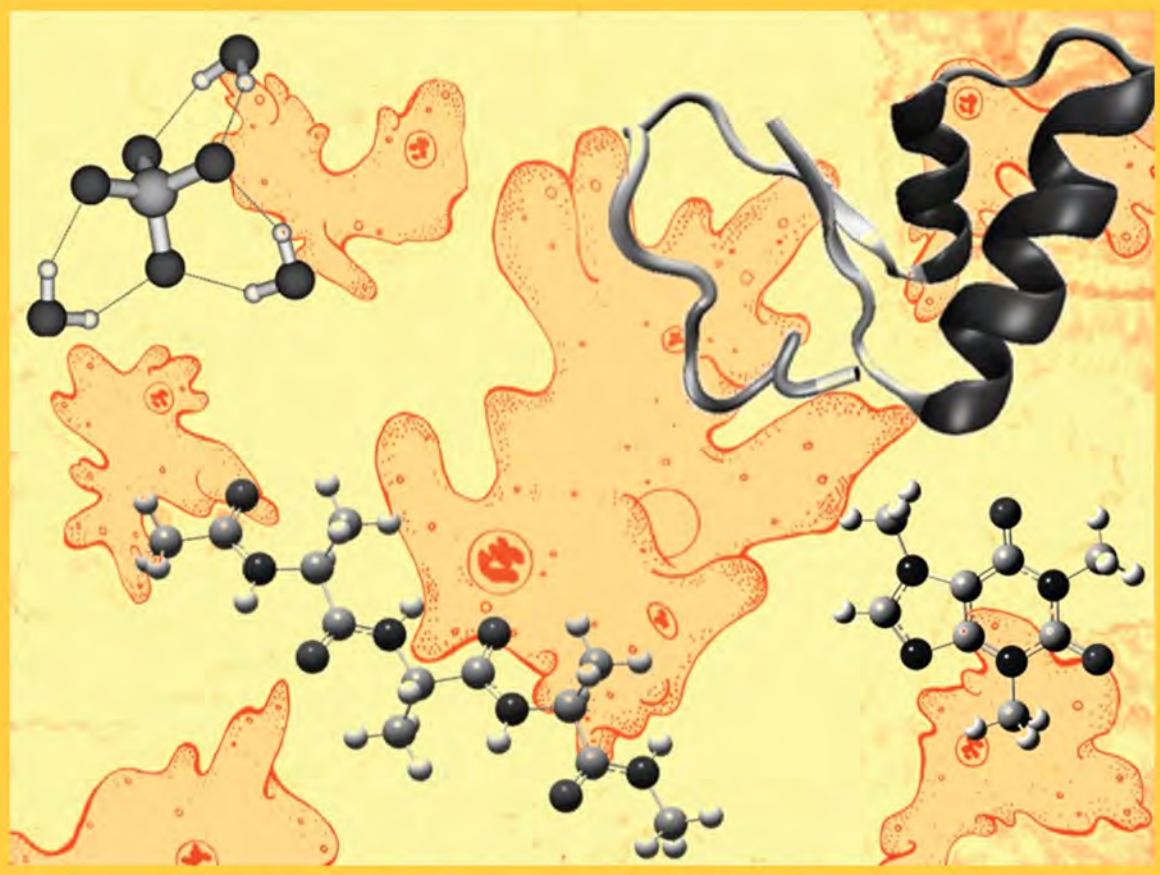
JPCBFK

MARCH 4, 2010
VOLUME 114
NUMBER 8

pubs.acs.org/JPCB

THE JOURNAL OF PHYSICAL CHEMISTRY B



Reviewing the Status of the AMOEBA Force Field (see page 9A)

SOFT CONDENSED MATTER AND BIOPHYSICAL CHEMISTRY

ACS Publications
High quality. High impact.

www.acs.org

Current Status of the AMOEBA Polarizable Force Field

Jay W. Ponder and Chuanjie Wu

Department of Biochemistry and Molecular Biophysics, Washington University, St. Louis, Missouri 63110

Pengyu Ren

Department of Biomedical Engineering, University of Texas, Austin, Texas 78712-1062

Vijay S. Pande,[†,‡] John D. Chodera,[†] Michael J. Schnieders,[†] and Imran Haque[‡]

Departments of Chemistry, Computer Science, Stanford University, Stanford, California 94305

David L. Mobley

Department of Chemistry, University of New Orleans, New Orleans, Louisiana 70148

Daniel S. Lambrecht, Robert A. DiStasio, Jr., and Martin Head-Gordon

Department of Chemistry, University of California, Berkeley, California 94720

Gary N. I. Clark, Margaret E. Johnson, and Teresa Head-Gordon*

Department of Bioengineering, University of California, Berkeley, California 94720

Received: November 9, 2009; Revised Manuscript Received: December 17, 2009

Molecular force fields have been approaching a generational transition over the past several years, moving away from well-established and well-tuned, but intrinsically limited, fixed point charge models toward more intricate and expensive polarizable models that should allow more accurate description of molecular properties. The recently introduced AMOEBA force field is a leading publicly available example of this next generation of theoretical model, but to date, it has only received relatively limited validation, which we address here. We show that the AMOEBA force field is in fact a significant improvement over fixed charge models for small molecule structural and thermodynamic observables in particular, although further fine-tuning is necessary to describe solvation free energies of drug-like small molecules, dynamical properties away from ambient conditions, and possible improvements in aromatic interactions. State of the art electronic structure calculations reveal generally very good agreement with AMOEBA for demanding problems such as relative conformational energies of the alanine tetrapeptide and isomers of water sulfate complexes. AMOEBA is shown to be especially successful on protein−ligand binding and computational X-ray crystallography where polarization and accurate electrostatics are critical.

Introduction

Molecular simulation is now an accepted and integral part of contemporary chemistry, biology, and material science. The allure of molecular simulation is that most if not all relevant structural, kinetic, and thermodynamic observables of a chemical system can be calculated at one time, in the context of a molecular model that can provide insight and new hypotheses. The predictive quality of these observables depends on the accuracy of the potential energy surface and the ability to characterize it through effective sampling of configurations or phase space. Over the last two decades, the field of molecular simulation has been dominated by research problems such as protein folding where dynamical time scales or configurational sampling are the biggest bottlenecks to reaching testable hypotheses or comparisons to experimental results. Given the demands of sampling over so many degrees of freedom to convergence, potential energy surfaces for molecular simulations rely on approximations and empirical input in order to formulate tractable descriptions of the (bio)material in a realistic chemical environment.

Nonpolarizable (fixed charge) models provide an inexpensive description or "effective" potential with approximations that cannot fully capture many-body effects such as electronic polarization. Fixed charge protein and water models went through an extensive period of validation for several decades after they were introduced.[1] The consensus of a number of validation studies is that, while fixed charge models offer tractable descriptions and are robust for equilibrium properties for homogeneous systems, evident discrepancies were identified between simulations and experiments away from ambient conditions, for dynamical properties, and for heterogeneous chemical systems in general.[1,2] Polarizable empirical force fields, which offer a clear and systematic improvement in functional form by including many-body effects, have been introduced into the chemical and biochemical simulation community over the past two decades, and only recently for biomolecular simulation.[1,3] The question in molecular computation currently is whether new

[†] Department of Chemistry, Stanford University.
[‡] Department of Computer Science, Stanford University.

10.1021/jp910674d
Published on Web 02/05/2010

Jay Ponder (Ph.D. 1984, Harvard University, 1985−1990, Postdoctoral Fellow, Yale University) is on the faculty at Washington University, where his group has a longstanding interest in development of software tools for molecular modeling, with particular emphasis on accurate conformational analysis and calculation of intermolecular interactions.

Chuanjie Wu (Ph.D. 2006, Tianjin University) did his doctorate in force field development and application and now is a postdoctoral researcher in AMOEBA force field parametrization and related methodology development at Washington University, St. Louis.

Pengyu Ren (Ph.D. 1999, University of Cincinnati; 2000−2005 Postdoctoral Researcher, Washington University) is an Assistant Professor of Biomedical Engineering at University of Texas at Austin, developing a range of computational tools to study proteins and nucleic acids, with emphasis on protein−ligand binding and computational drug discovery.

Vijay Pande (Ph.D. 1995, MIT; 1996−1999 Miller Fellow, UC Berkeley) is a Professor of Chemistry, Structural Biology, and Computer Science at Stanford University, working on theoretical and computational methods for biomolecular simulation and applications to the biophysics of protein folding and misfolding diseases.

John D. Chodera (Ph.D. 2006, UC San Francisco; 2006−2008 Postdoctoral Researcher, Stanford University) is a Distinguished Postdoctoral Fellow with the California Institute of Quantitative Biosciences (QB3) at the University of California, Berkeley, working on the statistical mechanics of biomolecular function, with emphasis on conformational dynamics, single-molecule experiments, and drug discovery.

Michael J. Schnieders (Ph.D. 2007, Washington University, St. Louis) is a Postdoctoral Fellow at Stanford University working on computational X-ray crystallography.

Imran Haque (B.S. 2006, UC Berkeley) is pursuing a Ph.D. at Stanford University working on computer-aided drug design and high-performance methods for molecular simulation.

David L. Mobley (Ph.D. 2004, UC Davis; 2005−2008 Postdoctoral Researcher, UC San Francisco; 2008 Chief Science Officer, Simprota Corporation) is an Assistant Professor at University of New Orleans applying computational and theoretical methods to understand and quantitatively predict protein−ligand binding, solvation, and solubility.

Daniel Sebastian Lambrecht (Dipl.-Chem. 2003, U. Düsseldorf, Dr. rer. nat. 2007, U. Tübingen) completed his doctorate on the development of linear-scaling approaches in quantum chemistry. He is now a postdoctoral scholar at UC Berkeley.

Robert Distasio, Jr. (Ph.D. 2009, UC Berkeley) completed his doctorate on development and application of fast electronic structure methods that include electron correlation. He is now a Postdoctoral Research Associate at Princeton University.

Martin Head-Gordon (Ph.D. 1989, Carnegie Mellon, 1989−1992, Postdoctoral Member of Technical Staff, AT&T Bell Laboratories) leads a group at UC Berkeley that develops and applies electronic structure theory.

Gary N. I. Clark (B.Sc. 2002, U. Loughborough; Ph.D. 2008, Imperial College London) is a postdoctoral researcher at UC Berkeley working on water structure, clustering, and dynamics at hydrophobic/hydrophilic interfaces.

Margaret E. Johnson (B.S. 2004 Columbia U. Ph.D. 2009, UC Berkeley) completed her doctorate at UC Berkeley on studies of bulk and hydration water structure and dynamics and is now an NIH postdoctoral researcher.

Teresa Head-Gordon (Ph.D. 1989, Carnegie Mellon, 1990−1992, Postdoctoral Member of Technical Staff, AT&T Bell Laboratories) leads a group at UC Berkeley that develops theoretical/experimental methods to study biomaterials assembly and bulk and hydration water properties.

polarizable force field parametrizations have successfully reached a new level of predictive power over their nonpolarizable predecessors.

Demonstrable testing of empirical biomolecular and water force fields is something that the simulation community requires, since so many academic and industry researchers use molecular mechanics and molecular dynamics methodology to tackle biological, chemical, and material science problems of interest. For example, the TINKER software program for molecular mechanics and dynamics simulation[4] has been downloaded by close to 60 000 separate external users, including essentially every major research university and many biotech and pharmaceutical companies, and even greater numbers are expected for other simulation software packages such as Amber,[5] CHARMM,[6] GROMACS,[7] and NAMD.[8] Force field validation and subsequent improvement has been the admirable history of the large community effort on fixed charge force fields led by developers of the Amber,[9] CHARMM,[10] GROMOS,[11] and OPLS[12] potential energy models over many decades. In this Feature Article, we hope to continue that tradition by summarizing some important early validation tests by a consortium of research groups at Washington University St. Louis, University of Texas at Austin, UC Berkeley, and Stanford University conducted on the general purpose polarizable force field AMOEBA (atomic multipole optimized energetics for biomolecular applications) developed by Ponder and co-workers.[13]

The first level of comprehensive testing of any force field will include predictions made by that potential against the best experiments and theoretical calculations available on a wide array of small molecule data in both gas phase and condensed phase environments. In fact, AMOEBA belongs to the class of molecular mechanics force fields that aims for high fidelity to *ab initio* calculations but at a computational cost that makes it suited for both small molecule and biomolecule condensed phase studies where statistical mechanical sampling is necessary. In practical terms, AMOEBA is intermediate in computational cost between other transferable polarizable force fields such as SIBFA (sum of interactions between fragments ab initio),[14] NEMO (non-empirical molecular orbital),[15] and QM/MM approaches such as DRF[16] and inexpensive polarizable biomolecular force fields from the Amber,[3i] CHARMM,[3e,f,h] and OPLS/PFF consortiums.[3d,g] In this paper, we review the AMOEBA model and its performance in several areas including gas phase properties against state-of-the-art quantum mechanical calculations, aqueous peptide solvation, structure and dynamics, solvation free energies of small molecule protein analogues and drug-like molecules with high precision, early structural stability studies of aqueous solvated proteins, computational X-ray crystallography, and protein−ligand binding.

The AMOEBA Force Field

The AMOEBA force field has the following general functional form for the interactions among atoms

$$U = U_{\text{bond}} + U_{\text{angle}} + U_{b\theta} + U_{\text{oop}} + U_{\text{torsion}} + U_{\text{vdW}} + U_{\text{ele}}^{\text{perm}} + U_{\text{ele}}^{\text{ind}} \quad (1)$$

where the first five terms describe the short-range valence interactions (bond stretching, angle bending, bond-angle cross term, out-of-plane bending, and torsional rotation) and the last three terms are the nonbonded vdW and electrostatic contributions. AMOEBA contains a number of differences from "traditional" biomolecular potentials such as the current Amber ff99SB,[9c] CHARMM27,[10] OPLS-AA,[12b,c] and GROMOS 53A6[11b] in the use of bond-angle cross terms, a formal Wilson−Decius−Cross decomposition of angle bending into in-plane and out-of-plane components, and a "softer" buffered 14−7 vdW form. However, the major difference is replacement

of the fixed partial charge model with polarizable atomic multipoles through the quadrupole moments. One advantage of the AMOEBA model is its emphasis on replication of molecular polarizabilities and electrostatic potentials, instead of just interaction energies. The use of permanent dipoles and quadrupoles allows accurate reproduction of molecular electrostatic potentials, and fine-tuning of subtle directional effects in hydrogen bonding and other interactions. The inclusion of explicit dipole polarization allows the AMOEBA model to respond to changing or heterogeneous molecular environments, and allows direct parametrization against gas phase experimental data and high-level quantum mechanical results. The AMOEBA model also presents a consistent treatment of intra- and intermolecular polarization that is achieved through a physically motivated damping scheme for local polarization effects.[17] A further attractive aspect of AMOEBA is its use of multipole moments derived directly from *ab initio* quantum mechanical electron densities for small molecules and molecular fragments. The design goal for AMOEBA has been to achieve a "chemical accuracy" of 0.5 kcal/mol or better for small molecule and protein−ligand interactions. We describe the functional form of the AMOEBA force field below and provide the current standard parameter set for small molecules and proteins in the Supporting Information, while further details of its parametrization are given in ref 13.

Short-Ranged Valence Interactions. The AMOEBA model includes full intramolecular flexibility. For atoms directly bonded (1−2) and separated by two bonds (1−3), the covalent energy is represented by empirical functions of bond lengths and angles. The functional forms for bond stretching (eq 2), angle bending (eq 3), and the coupling between the stretching and bending (eq 4) are those of the MM3 force field,[18] and include an accounting of anharmonicity through the use of higher-order deviations from ideal bond lengths (b_0) and angles (θ_0):

$$U_{\text{bond}} = K_b(b - b_0)^2[1 - 2.55(b - b_0) + (7/12)2.55(b - b_0)^2] \quad (2)$$

$$U_{\text{angle}} = K_\theta(\theta - \theta_0)^2[1 - 0.014(\theta - \theta_0) + 5.6 \times 10^{-5}(\theta - \theta_0)^2 - 7.0 \times 10^{-7}(\theta - \theta_0)^3 + 2.2 \times 10^{-8}(\theta - \theta_0)^4] \quad (3)$$

$$U_{b\theta} = K_{b\theta}[(b - b_0) + (b' - b_0')](\theta - \theta_0) \quad (4)$$

$$U_{\text{oop}} = K_\chi \chi^2 \quad (5)$$

where the bond length, b or b', and bond angle, θ, and energies are in units of Å, degrees, and kcal/mol, with the force constants, K, given in corresponding units. In addition, a Wilson−Decius−Cross function is used at sp^2-hybridized trigonal centers to restrain the out-of-plane bending (eq 5),[19] where for sequentially bonded centers i, j, k, and l, χ refers to the angle between the jl vector and the ijk plane.

A traditional Fourier expansion (a 1-fold through 6-fold trigonometric form) torsional functional

$$U_{\text{torsion}} = \sum_n K_{n\phi}[1 + \cos(n\phi \pm \delta)] \quad (6)$$

is used to aid in merging the short-range "valence" terms with the long-range "nonbonded" interactions. For dihedral angles involving two joined trigonal centers, such as the amide bond of the protein backbone, a Bell torsion[20] functional is applied in addition to the regular torsional terms, where ϕ used in eq 6 is the dihedral angle computed from the p-orbital directions at the two trigonal centers, rather than from the usual bond vectors. The rotational barrier around the amide bond is much higher than for a single covalent bond, and the bigger barrier is largely due to the double bond nature originating in the overlap of the adjacent p-orbitals. Use of the Bell torsion allows appropriately increased flexibility of atoms bonded to trigonal centers (e.g., aromatic hydrogen atoms).[21] The torsional parameters are refined *after* the nonbonded parameters are determined with the hope that the improved AMOEBA intramolecular electrostatic model will lead to a more "physical" balance between the local (vdW + electrostatic + torsional) and long-range (vdW + electrostatic) interactions in the conformational energy.

van der Waals Interactions. The pairwise additive van der Waals (vdW) interaction in AMOEBA adopts the buffered 14−7 functional form[22]

$$U_{\text{vdw}}(ij) = \varepsilon_{ij}\left(\frac{1.07}{\rho_{ij} + 0.07}\right)^7\left(\frac{1.12}{\rho_{ij}^7 + 0.12} - 2\right) \quad (7)$$

where ε_{ij} in kcal/mol is the potential well depth and $\rho_{ij} = R_{ij}/R_{ij}^0$, where R_{ij} in angstroms is the actual separation between i and j and R_{ij}^0 is the minimum energy distance. For heterogeneous atom pairs, the combination rules are given by

$$\varepsilon_{ij} = \frac{4\varepsilon_{ii}\varepsilon_{jj}}{(\varepsilon_{ii}^{1/2} + \varepsilon_{jj}^{1/2})^2} \quad \text{and} \quad R_{ij}^0 = \frac{(R_{ii}^0)^3 + (R_{jj}^0)^3}{(R_{ii}^0)^2 + (R_{jj}^0)^2} \quad (8)$$

The buffered 14−7 function yields a slightly "softer" repulsive region than the Lennard-Jones 6−12 function but achieves a steeper repulsion at very short range than typical Buckingham exp-6 formulations. The buffered 14−7 form was considered superior, as it provides a better fit to gas phase *ab initio* results and liquid properties of noble gases.[22] The AMOEBA van der Waals parameters are derived by fitting to both gas phase and bulk phase experimental properties.

Each atom in AMOEBA possesses a vdW site. For non-hydrogen atoms, the site is located at the position of the atomic nucleus. For a hydrogen atom connected to an atom X, it is placed along the H−X bond such that the distance between the atom X and the vdW site of H is a percentage of the full bond length, namely, the "reduction factor". Application of reduction factors to shift hydrogen sites off of the nuclear centers dates from early work by Stewart et al.,[23] and X-ray structural analyses of glycylglycine and sulfamic acid also support this view.[24] A similar approach is used in MM3 and other force fields from the Allinger group.[18] The use of a reduction factor was found to simultaneously improve the fit to accurate QM water dimer structures and energies for several configurations.

Permanent Electrostatic Interactions. The electrostatic energy in AMOEBA includes contributions from both permanent and induced multipoles. The permanent atomic multipoles (PAM) at each atomic center include the monopole (charge), dipole, and quadrupole moments

$$M_i = [q_i, \mu_{ix}, \mu_{iy}, \mu_{iz}, Q_{ixx}, Q_{ixy}, Q_{ixz}, \ldots, Q_{izz}]^t \quad (9)$$

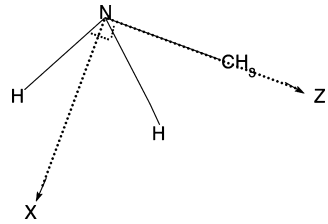


Figure 1. An example of the "*z*-then-bisector" local frame definition. Shown for methylamine that ensures that the atomic multipoles remain constant with time within this local reference frame.

where q_i is the point charge located at the center of atom i, μ is the dipole, and Q is the quadrupole, all in Cartesian representation, and t is the transpose. In the Cartesian polytensor formalism,[25] the interaction energy between atoms i and j separated by r_{ji} is represented as $U_{\text{elec}}^{\text{perm}}(r_{ij}) = M_i^T T_{ij} M_j$ where

$$T_{ij} = \begin{bmatrix} 1 & \frac{\partial}{\partial x_j} & \frac{\partial}{\partial y_j} & \frac{\partial}{\partial z_j} & \text{L} \\ \frac{\partial}{\partial x_i} & \frac{\partial^2}{\partial x_i \partial x_j} & \frac{\partial^2}{\partial x_i \partial y_j} & \frac{\partial^2}{\partial x_i \partial z_j} & \text{L} \\ \frac{\partial}{\partial y_i} & \frac{\partial^2}{\partial y_i \partial x_j} & \frac{\partial^2}{\partial y_i \partial y_j} & \frac{\partial^2}{\partial y_i \partial z_j} & \text{L} \\ \frac{\partial}{\partial z_i} & \frac{\partial^2}{\partial z_i \partial x_j} & \frac{\partial^2}{\partial z_i \partial y_j} & \frac{\partial^2}{\partial z_i \partial z_j} & \text{L} \\ \text{M} & \text{M} & \text{M} & \text{M} & \text{O} \end{bmatrix} \left(\frac{1}{r_{ji}}\right) \quad (10)$$

There are typically five independent quadrupole components due to symmetry ($Q_{\alpha\beta} = Q_{\beta\alpha}$) and the use of traceless moments ($\Sigma Q_{\alpha\alpha} = 0$). Furthermore, the μ_y, Q_{xy}, and Q_{yz} components are zero except for chiral atoms such as the backbone C_α in amino acids. Therefore, most nonchiral atoms will carry six unique, permanent electrostatic multipole parameters.

As previously described for the AMOEBA water model, the dipole and quadrupole are defined with respect to a local reference frame defined by neighboring atoms.[13b] A new "*z*-then-bisector" local frame definition has been developed for atoms with single lone pairs such as the sp^3 nitrogen. An example of this new frame is given for the N atom in methylamine in Figure 1. In principle, the choice of frame should respect local symmetry such that axes are placed along major chemical determinants. As the molecules vibrate, rotate, and diffuse over the course of a dynamic simulation, the atomic multipoles remain constant with respect to the local frame definition.

Atomic multipole moments in this study are derived from *ab initio* calculations of the small molecules using Stone's distributed multipole analysis (DMA).[26] Convergence to reasonable chemical accuracy goals of 0.5 kcal/mol requires inclusion of terms through quadrupole moments. Alternative approaches, such as electrostatic potential fitting and electron density partitioning, have also been explored.[13a] For molecules such as alanine dipeptide that possess conformational degrees of freedom, an extra step is necessary to obtain the conformation-independent permanent atomic multipoles (PAM), as will be discussed in the intramolecular polarization section below. The original DMA-derived multipoles[26a] are converted to the final electrostatic parameters in the corresponding local frame for each atom type:

$$M_{\text{local}} = \mathscr{R}^{-1} M \quad (11)$$

where $\mathscr{R}$ is the rotation matrix transforming the local into the global reference frame.[27]

Electronic Polarization. Electronic polarization refers to the distortion of electron density under the influence of an external field. It represents a major contribution to the overall many-body energetic description of molecular clusters and condensed phases, even though there are situations where other contributions related to dispersion and repulsion are not negligible.[28] In AMOEBA, a classical point dipole moment is induced at each polarizable atomic site according to the electric field felt by that site. Molecular polarization is achieved via an interactive induction model with distributed atomic polarizabilities based on Thole's damped interaction method.[17a] This interactive or mutual induction scheme requires that an induced dipole produced at any site i will further polarize all other sites, and such mutual induction will continue until the induced dipoles at each site reach convergence. One key aspect of Thole's approach is damping of the polarization interaction at very short range to avoid the so-called polarization catastrophe, a well-known artifact of point polarizability models. The damping is effectively achieved by smearing one of the atomic multipole moments in each pair of interaction sites (the result is independent of which one is smeared).[29] The smearing function for charges adopted by AMOEBA has the functional form

$$\rho = \frac{3a}{4\pi} \exp(-au^3) \quad (12)$$

where $u = r_{ij}/(\alpha_i \alpha_j)^{1/6}$ is the effective distance as a function of linear separation r_{ij} and atomic polarizabilities of sites i (α_i) and j (α_j). The factor "a" is a dimensionless width parameter of the smeared charge distribution, and effectively controls the damping strength. Corresponding damping functions for charge, dipole, and quadrupole interactions were derived through their chain rule relationships.[13b]

The Thole model has the advantages of simplicity and transferability, as evidenced by the fact that it reasonably reproduces the molecular polarizability tensor of numerous small molecules using just one isotropic atomic polarizability for each element, plus a universal damping factor.[17a] However, there has been controversy as to whether polarizability decreases, and if so to what extent, when a molecule moves from gas to condensed phase. Morita recently estimated that water polarizability decreases by 7−9%,[30] reduced from 13−18% reported in an earlier publication.[31] Mennuci et al. showed that the effect of Pauli exclusion is to reduce the dipole polarizability of a solute by 2%. In contrast, Gubskaya and Kusilik suggested an increase of the polarizability of water in condensed phases.[32] In light of uncertainty in theoretical estimates of liquid polarizability, we have chosen to use the same atomic polarizability values for both gas and condensed phase. The resulting average dipole moment of AMOEBA liquid water, using a constant gas phase polarizability value, is 2.8 D, only slightly lower than recent quantum mechanical estimates of 2.95 D.[33] Furthermore, in the AMOEBA polarization model, the damping factor provides another control over the ability of an atom to polarize; the universal damping factor adopted by AMOEBA is $a = 0.39$, which effectively leads to a stronger damping and less short-range polarization than the original value of 0.572 suggested by Thole. We have kept the same atomic polarizabilities (Å^3) given by Thole, i.e., 1.334 for carbon, 0.496 for hydrogen, 1.073 for nitrogen, and 0.837 for oxygen. The only exception is for carbon and hydrogen in aromatic rings, where we found the use of somewhat larger values greatly improves the molecular polarizability tensor of benzene and polycyclic aromatics.

Validation Studies against Electronic Structure Calculations

A necessary if not sufficient condition for robust performance of a force field is its ability to reproduce or predict relative conformational energies of model systems complex enough to contain realistic features but simple enough to be treated by electronic structure methods[34] that can yield reliable benchmark results. First principles electronic structure calculations can provide benchmarks of uncompromising accuracy for relative energies of molecules in different conformations, since quantum mechanics provides an essentially exact description of the behavior of electrons in small molecules.[35] However, in practice, approximate electronic structure calculations for larger molecules suffer from errors associated with the imperfect treatment of electron correlations and the use of incomplete atomic orbital basis sets, which can render the results too inaccurate to be useful, or even worse, potentially misleading. Incomplete treatments of electron correlation such as commonly used density functional theory (DFT) methods omit dispersion interactions that are very important in biological macromolecules, while incomplete basis sets give rise to, for example, intramolecular basis set superposition error (BSSE), which favors compact relative to extended conformations.

In the work summarized here, we cannot claim to have completely eliminated either of these problems, but we have certainly reduced some limitations of earlier calculations, using new algorithms and faster computers. With respect to electron correlation, we have used the second order Møller−Plesset (MP2) method, which includes long-range electron correlation effects in a reasonably accurate manner, and then tested for remaining errors by using local coupled cluster theory with a smaller basis set. With respect to basis set errors, we have performed calculations with the Dunning augmented correlation consistent basis sets up to the aug-cc-pVQZ level, which we used with aug-cc-pVTZ results to perform an extrapolation to the complete basis set limit (TQ extrapolation). Comparisons against smaller basis sets show that this level of theory is required to obtain reasonable convergence of MP2 relative energies. On the basis of these quantum mechanical benchmarks, we evaluate the AMOEBA performance on nanosolvation and conformational energetics of alanine tetrapeptide.

Conformational Searching for Global and Low-Lying Energy Minima for Water Cluster Systems. Nanodroplets and nanosolvation are interesting and demanding test cases for modeling water via polarizable force fields because they contain water molecules in different extremes of environment, ranging from surface molecules exposed to vacuum to buried molecules that experience bulk-like environment. Beyond direct use as a "stress test" for polarizable force fields, the science of nanodroplets is interesting in itself, since these species are intermediate between small clusters (20−30 molecules and below) that are currently intensively studied by high-accuracy electronic structure theory as well as beam experiments, and solvation in the bulk liquid solvent. They will have some of the features of water in confined regions, and may well exhibit interesting structural motifs that lie in between small cluster building blocks and the hydrogen-bonding patterns of the bulk. The behavior of a solute in these clusters in terms of whether it appears on the surface or in the interior may have similarities to the partitioning of solutes at interfaces.

Studies of the properties of the nanodroplets require an effective sampling technique due to the exponentially fast rise in the number of minima with cluster size. In hybrid energy approaches, a cheap energy function is used to provide configurations for the sampling of an expensive *ab initio* energy function. The primary problem in using hybrid energy schemes is that we have no knowledge or guarantee that the distribution of configurations generated with the lower quality energy function overlaps sufficiently with the higher quality energy function. However, we have found AMOEBA to be a reliable generator of viable minima with sound energy ordering when benchmarked against a reliable *ab initio* theoretical model. In a recent study on $n = 3$, 4, and 5 water−sulfate anion clusters $(H_2O)_nSO_4^{2-}$, we used replica exchange simulations over the temperature range from 140 to 500 K using the AMOEBA model, and all samples collected every 0.5 ps at every temperature were energy minimized using the BFGS local optimization algorithm. Sampling for all cluster sizes considered appeared to be exhaustive, since all of the 10 000 structures collected for each cluster size reduced to a smaller set of up to 200 local minima.

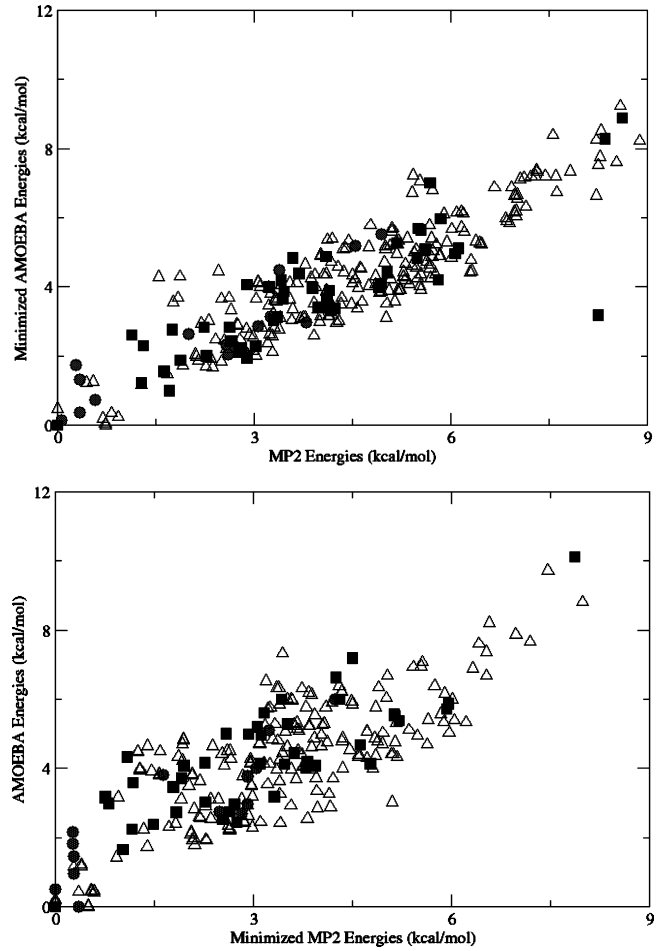


Figure 2. Energy correlations between AMOEBA and MP2 energies for (a) AMOEBA minimized water−sulfate anion clusters and (b) MP2 minimized water−sulfate anion clusters. Shown for $(H_2O)_nSO_4^{2-}$, where $n = 3$ (⊗), 4 (■), and 5 (△). Correlation coefficients are 0.88 ($n = 3$), 0.77 ($n = 4$), and 0.79 ($n = 5$) for AMOEBA geometries and correlation coefficients are 0.92 ($n = 3$), 0.92 ($n = 4$), and 0.90 ($n = 5$) for MP2 geometries.

Figure 2a shows that the quantitative correlation between AMOEBA and the *ab initio* theory is very good (correlation coefficient, $r^2 \sim 0.9$) while the qualitative comparison is excellent given the agreement on the global minimum structure for $n = 3$ and 4 that will likely dominate the nanosolvation properties of this system size, and very competitive low lying minima for $n = 5$. The lowest minimum energy structures determined from the empirical polarizable model were in turn energy minimized

TABLE 1: Comparison of Relative Energies (kcal/mol) for Sulfate−Water Clusters $(H_2O)_3SO_4^{2-}$ [a]

isomer	benchmark	RIMP2/ aug-cc-pVQZ	RIMP2/ aug-cc-pVDZ	AMOEBA
1	0.00	0.00	0.00	0.00
2	0.29	0.29	0.25	0.72
3	0.57	0.60	0.33	0.37
4	0.65	0.54	0.32	1.31
5	0.71	0.59	0.29	1.74
6	2.38	2.68	2.08	2.63
7	2.66	2.80	3.04	2.04
8	3.62	3.54	3.27	2.27

[a] The geometries of each cluster isomer were optimized at the RIMP2/aug-cc-pVTZ level, and single point quantum mechanical energies were calculated at a benchmark level (RIMP2/aug-cc-pVQZ+ΔCCSD(T)/6-31+G*), as well as RIMP2 using aug-cc-pVQZ and aug-cc-pVDZ basis sets. AMOEBA results are reported for the AMOEBA minimized structures.

by the RI-MP2 level of theory using an augmented cc-pVDZ basis set. Figure 2b shows that for minimized MP2 structures the quantitative correlation between AMOEBA single point energies and the *ab initio* theory is still very good ($r^2 \sim 0.8$), showing that AMOEBA geometries are in very good agreement with the benchmark calculation.

For the smaller $n = 3$ clusters, we can benchmark against high-level QM results. Table 1 shows the relative energies for the eight lowest-lying configurations. The RIMP2+ΔCCSD(T) reference energies were obtained at the RI-MP2/aug-cc-pVQZ level of theory, which were corrected at the CCSD(T)/6-31+G* level for higher-order correlation effects. Comparing the RIMP2/aug-cc-pVDZ and RIMP2/aug-cc-pVQZ results, we find basis set effects of up to 0.4 kcal/mol. Higher-order correlation effects are on the order of up to 0.3 kcal/mol, as seen by comparing RIMP2 and RIMP2+ΔCCSD(T) results. This emphasizes the importance of both effects, given that the energy differences between most low-lying isomers are on the same order of magnitude. We see that the AMOEBA force field has quite good energy ordering of the isomers relative to the highest level of theory, showing its validity outside the quantum chemistry levels of theory used in the parametrization scheme reported in ref 24.

Electronic Structure Calculations of Conformational Energies of the Alanine Tetrapeptide. Alanine tetrapeptide is a system which has at least several dozen low-lying conformational minima ranging from globular to extended, and includes hydrogen bonding and packing interactions that make it quite a rich biochemical system even in the gas phase. For this reason, benchmark calculations on alanine tetrapeptide first appeared roughly a dozen years ago.[36] In this section, we summarize, and in some instances extend, recent calculations[37] that significantly improve the accuracy and reliability of the earlier benchmarks. Table 2 contains relative energies calculated with different popular electronic structure methods all using geometries optimized at the same Hartree−Fock level of theory with the 6-31G** basis set, for 27 conformations of alanine tetrapeptide using the labels reported in ref 36. Our best (benchmark) level of theory (MP2 with TQ extrapolation) as well as the MP2 theory with a less complete basis set (DT extrapolation) are beyond originally published benchmarks at the lower double-ζ level of quality.[36] The comparison of the first two columns of Table 2 indicates how troublesome obtaining fully converged results is using an electron correlation method such as MP2. The DT level of theory is already beyond most literature calculations yet in some cases is not converged to within 1 kcal/mol of the larger TQ results. One must also assume that there would be a further shift on the order of perhaps up to 0.1 kcal/mol upon further improvement of the basis set beyond the TQ extrapolation.

The second comparison of importance in Table 2 is with standard DFT and the widely used B3LYP functional, using a very large cc-pVQZ basis set. While DFT calculations cannot be soundly extrapolated to the complete basis set limit, they also converge more rapidly with basis set size than MP2 theory, so we can consider these results to be quite well-converged. However, B3LYP is known to perform fairly poorly for intermolecular interactions (and hence conformational energies), as a result of limitations in its exchange and correlation functionals (for instance, it neglects dispersion interactions). Indeed, this causes serious discrepancies relative to the best MP2 results. The overall energy ranking of conformers is quite poor using this conventional DFT method, emphasizing its lack of suitability as sources of benchmark conformational energies.

The development of new functionals that improve exchange functionals and include empirical van der Waals corrections is likely to yield significantly improved performance. We assess the role of improved exchange with the range-separated ωB97 and ωB97X functionals,[38] and the additional effect of dispersion with the recently proposed ωB97X-D functional.[39] Table 2 shows conformer energies for the ωB97, ωB97X, and ωB97X-D long-range corrected functionals. All of them yield a significant improvement over B3LYP and exhibit an excellent agreement with the benchmark energies (correlation coefficients of 0.910, 0.932, and 0.908, respectively). Interestingly, however, the dispersion correction in ωB97X-D does not improve the performance of the functional in the present test case, suggesting that intramolecular dispersion effects may be adequately captured by other parts of the functional.

Finally, we report the original LMP2 results but using more tightly converged geometries than reported originally,[36] as well as the AMOEBA results which used LMP2 conformational energies of the alanine dipeptide as part of the parametrization of the AMOEBA protein model. It is interesting to see that AMOEBA (using AMOEBA relaxed geometries) gives a competitive energy ranking over all the conformations compared to the RI-MP2 benchmark ($r^2 \sim 0.88$), comparable to that exhibited by the LMP2 level of theory ($r^2 \sim 0.95$), and far better than conventional DFT ($r^2 \sim 0.46$). AMEOBA is essentially competitive with the new generation density functionals, ωB97, ωB97X, and ωB97X-D, which illustrates that it is very well balanced for polypeptide conformation energies.

From a biophysical viewpoint, one of the most important comparisons is between the extended conformation (conformer 1) and a compact globular conformation with a tight hairpin turn (conformer 3). This type of energy difference is particularly sensitive to basis set convergence problems because limitations of the basis set will favor the globular conformation, where atoms in nonbonded contact can artificially lower their energy by making fractional use of the functions on their nonbonded neighbors. This intramolecular basis set superposition error is essentially absent in the extended conformation. As a result, the benchmark extended-globular energy gap, $E_{gap} = 3.56$ kcal/mol, is overestimated by ~1.3 kcal/mol at the DT extrapolated level. This emphasizes the importance of carrying out the calculations to the largest feasible basis set size. Errors associated with neglect of dispersion interactions, which are relatively nonspecific, can sometimes approximately cancel out for conformations of approximately similar compactness. However, the energy difference between extended and globular

TABLE 2: Comparison of Benchmark RI-MP2 Calculations Approaching the Basis Set Limit against Other Electronic Structure Methods and AMOEBA for 27 Alanine Tetrapeptide Conformations[a]

conf.	MP2/TQ	MP2/DT	ωB97/LP	ωB97X/LP	ωB97X-D/LP	B3LYP/Q	LMP2/ cc-pVTZ (-f)	AMOEBA
11	0.000	0.000	0.000	0.000	0.000	2.184	0.000	0.090
12	0.290	0.346	1.099	1.187	0.902	3.266	0.699	0.372
3	0.571	0.693	0.723	0.425	0.523	1.251	0.195	0.000
26	0.674	1.223	2.367	2.187	2.398	1.806	0.373	1.509
20	1.755	2.335	3.032	2.666	3.190	1.270	1.061	2.432
18	1.913	2.468	1.944	1.577	2.668	0.000	0.718	1.938
15	2.194	1.707	2.136	2.347	2.591	5.291	2.261	1.164
25	2.495	3.118	3.575	3.389	4.030	2.442	1.784	2.935
6	2.895	3.148	2.601	2.433	3.196	3.228	2.383	2.422
21	2.918	3.009	2.474	2.547	2.749	3.336	2.300	2.828
17	3.418	3.414	2.474	2.547	2.749	4.638	2.980	2.520
16	3.549	3.784	4.713	4.292	4.438	3.925	3.021	2.575
13	3.655	4.538	4.034	3.474	4.509	1.225	1.965	3.519
19	3.816	4.319	3.950	3.682	4.648	2.033	3.029	3.610
24	3.976	4.115	4.280	4.241	5.131	4.521	3.171	3.424
27	4.020	4.513	5.197	4.989	5.423	4.155	3.378	4.355
1	4.130	5.553	4.745	4.088	5.576	0.742	2.690	4.162
2	4.190	5.390	4.892	4.358	5.650	1.056	2.780	4.001
8	4.640	4.477	5.024	5.050	5.390	6.333	4.364	4.258
14	4.679	5.395	5.811	5.336	5.758	3.862	3.877	4.029
5	5.261	6.353	6.431	5.835	5.758	3.263	4.074	4.152
4	5.730	6.884	6.907	6.219	7.317	3.350	4.062	4.831
23	5.815	5.979	5.944	5.809	6.383	6.335	5.018	5.618
22	5.824	5.899	5.667	5.520	6.126	6.318	5.019	4.295
7	6.665	6.931	6.648	6.614	6.126	7.730	5.927	4.385
10	7.791	7.766	7.637	7.707	8.286	8.367	7.189	5.613
9	7.923	8.197	7.932	7.631	8.033	6.264	7.129	8.066

[a] All relative energies are in kcal/mol and geometries optimized at the HF/6-31G** level. AMOEBA results used minimized structures based on the AMOEBA force field for each conformation.

TABLE 3: Effect of the Level of Theory Used for Geometry Optimization on the Energy Difference (in kcal/mol) between the Extended and Globular Conformations of Alanine Tetrapeptide[a]

energy evaluation	level of theory for geometry optimization			
	RI-MP2/T	B3LYP/T	HF/T	HF/6-31G**
RI-MP2/TQ	**4.994**	4.414	2.884	3.559
RI-MP2/DT	6.720	5.582	3.942	4.860
B3LYP/Q	−1.320	−0.093	−0.560	−0.590

[a] The benchmark value of 4.994 kcal/mol for the energy gap is highlighted in bold.

conformations is quite sensitive to the neglect of dispersion, resulting in a calculated B3LYP $E_{gap} = -0.51$, a large error that underestimates the benchmark calculation by roughly 4 kcal/mol. The corresponding gap measured by LMP2 is underestimated by ~1.1 kcal/mol, likely due to basis set size limitations and the local approximation of the model. AMOEBA performs the best on this benchmark, overshooting the RI-MP2/TQ result by only ~0.6 kcal/mol (close to the AMOEBA chemical accuracy goal of 0.5 kcal/mol), although again it is based on a comparison using AMOEBA relaxed geometries and not the HF/6-31G** geometries.

The effect of geometry optimization on the extended-globular gap is probed further with the calculations shown in Table 3, where large basis set geometry optimizations at the MP2, DFT, and HF levels are compared via single point energy calculations using the three different sets of structures. While it has been shown that the MP2 geometries are superior to HF geometries, it is commonly assumed (generally for good reason) that DFT or HF structures are adequate, because errors in electron correlation treatment cancel for small displacements of the geometry. However, there are significant shifts in relative conformational energy at the highest level of theory (RI-MP2/TQ) depending upon the geometry that is used, with a new benchmark value of $E_{gap} = 4.994$ kcal/mol. There is a shift of over 2 kcal/mol between HF and MP2 geometries, with the DFT geometry in much closer agreement (0.65 kcal/mol) with RI-MP2. Even between small basis HF (Table 1) versus the larger basis results shown in Table 2, there is a shift of roughly 0.7 kcal/mol. Against the new MP2 geometry benchmark, E_{gap} is overestimated by ~1.7 kcal/mol at the RI-MP2/DT extrapolated level, while DFT underestimates the gap by now roughly 6 kcal/mol. By contrast, AMOEBA now undershoots the benchmark result by ~0.9 kcal/mol, showing that AMOEBA geometries are the most robust when compared to the RI-MP2 geometries and energies.

Validation against Hydration Free Energies

The evaluation of hydration free energies is a natural test of any force field, since it incorporates many challenging aspects of a heterogeneous chemical environment that are not involved in the parametrization of the protein fragments or water force fields by themselves.[40] The AMOEBA solvation free energies were computed using a free energy perturbation procedure based on three thermocycle steps and processed with the Bennett acceptance ratio (BAR) method.[41] For each small molecule, the thermodynamic cycle corresponded to first solute discharging in a vacuum over 7 windows, followed by a soft core modification of eq 7 to introduce the solute−solvent van der Waals coupling over 16 windows, and finally solute recharging in water over 7 windows. The statistical samples of the first thermocycle step in a vacuum were collected every 0.5 ps from a 10 ns stochastic dynamics simulation with an integration time step of 0.1 fs, while the thermocycle steps in the condensed phase were run for 1 ns in the NVT ensemble with the density fixed at 1.000 g cm^{-3}. Induced dipoles were converged to 10^{-5}

TABLE 4: Accuracy of AMOEBA Solvation Free Energies for Small Molecules[a,b]

compound	AMOEBA	experiment	compound	AMOEBA	experiment
isopropanol	−4.21 ± 0.34	−4.74	propane	1.69 ± 0.17	1.96
methylether	−2.22 ± 0.38	−1.92	methane	1.73 ± 0.13	1.98
H_2S	−0.41 ± 0.17	−0.44	methanol	−4.79 ± 0.23	−5.10
p-cresol	−5.60 ± 0.23	−6.61	*n*-propanol	−4.85 ± 0.27	−4.85
ethylsulfide	−1.74 ± 0.24	−1.14	toluene	−1.53 ± 0.25	−0.89
dimethylsulfide	−1.85 ± 0.21	−1.83	ethylbenzene	−0.80 ± 0.28	−0.79
phenol	−5.05 ± 0.28	−6.62	*N*-methylacetamide	−8.66 ± 0.30	−10.00
benzene	−1.23 ± 0.23	−0.90	water	−5.86 ± 0.19	−6.32
ethanol	−4.69 ± 0.25	−4.96	acetic acid	−5.63 ± 0.20	−6.69
ethane	1.73 ± 0.15	1.81	methylsulfide	−1.44 ± 0.27	−1.24
n-butane	1.11 ± 0.21	2.07	methylethylsulfide	−1.98 ± 0.32	−1.50
dinitrogen	2.26 ± 0.12	2.49	imidazole	−10.25 ± 0.30	−9.63
methylamine	−5.46 ± 0.25	−4.55	acetamide	−9.30 ± 0.27	−9.71
dimethylamine	−3.04 ± 0.26	−4.29	ethylamine	−4.33 ± 0.24	−4.50
trimethylamine	−2.09 ± 0.24	−3.20	pyrrolidine	−4.88 ± 0.29	−5.48

[a] The uncertainty is the statistical uncertainty in the BAR free energy calculation. All units are kcal/mol. When compared to the experimental results, the RMS error for the 30 AMOEBA solvation free energies is 0.68 kcal/mol and the mean signed error is +0.14 kcal/mol. [b] Experimental values are reported in ref 42a, except for imidazol taken from ref 42b, *N*-methylacetamide taken from ref 42c, and methylsulfide and water taken from ref 42d.

D per step per atom for simulations in a vacuum and 10^{-2} D in the liquid during the trajectory, and the energies of the condensed phase snapshots (saved every 0.5 ps) were reevaluated with the induced dipole converged to 10^{-5} D. BAR was then used to estimate the free energy between the neighboring steps, and the final free energy was taken as the sum over all windows.

Table 4 reports the AMOEBA solvation free energies of common small molecules found in biochemistry, including common amino acid side chain analogues, with corresponding statistical uncertainties obtained via a block averaging applied to each simulation step, and the final statistical error bar is a sum of the uncertainties over all steps. When compared to the experimental results,[42] the rms error for AMOEBA solvation free energies is 0.68 kcal/mol, with a mean signed error of +0.14 kcal/mol. Calculated solvation free energies using traditional fixed charge force fields typically have an average rms error of 1.0−1.25 kcal/mol compared to available experiments for similar sets of molecules and a general shift in solvation free energy with a mean error of approximately 1 kcal/mol,[40] demonstrating that for chemical spaces similar to proteins AMOEBA offers significant improvement over corresponding fixed charge force fields.

Prediction of Solvation Free Energies for 2009 OpenEye SAMPL Competition. The Statistical Assessment of the Modeling of Proteins and Ligands (SAMPL) blind challenge is an assessment of force fields and sampling methods for protein and ligand modeling. One prediction aspect highlighted in the first SAMPL contest in 2008 consisted of predicting 63 vacuum−water transfer energies. A number of research groups using fixed charge force field models with water represented explicitly calculated solvation free energies using standard free energy perturbation MD calculations, and ultimately their blind prediction results were compared to available experimental literature numbers. The overall performance of these approaches gave an rms error of over 3 kcal/mol compared to the SAMPL reported experimental data.[43] The goal of these blind assessment approaches is to not criticize the underperformance of fixed charge force fields but to better understand when they do well, and when additional physics of the computational model is needed for predicting more challenging classes of compounds.

The AMOEBA force field was used to predict vacuum-to-water solvation free energies of 43 drug-like and other organic molecules for the 2009 SAMPL exercise (http://sampl.eyesopen.com/). Alchemical hydration free energy calculations to compute the transfer free energy from 1 M gas phase to 1 M aqueous solution were carried out in a manner similar to that described in ref 44 using a preview release of Tinker 5 modified to add a numerically computed analytical long-range dispersion correction,[45] soft-core forms of the Halgren potential, and the ability to periodically evaluate potential energies at all alchemical intermediates. In a vacuum and solvent, seven discharging intermediate states were used to scale charges, multipoles, and polarizabilities by factors lambda, crudely optimized to reflect the quadratic dependence of charging self-energies, while torsional barriers were correspondingly scaled by linear factors. In solvent, a decoupling parameter λ_h was used to modify the Halgren potential shift constants and well depth to mimic a soft-core potential at intermediate values of λ_h. Vacuum simulations (discharging only) at each alchemical intermediate were run for 5 ns using Langevin dynamics with a collision rate of 5/ps, with energies at all alchemical states written every 10 ps. Solvated systems consisted of the solute molecule immersed in an isotropic box containing 850 AMOEBA water. Solvated simulations (discharging and decoupling) for each alchemical intermediate were run for 300−600 ps using the Berendsen weak-coupling algorithm[46] for both thermal (coupling time 0.1 ps) and volume (coupling time 2 ps) control which are available in Tinker, though the distribution generated by Berendsen should approach the correct NPT ensemble in the thermodynamic limit. Potential energies from solvated simulations were computed at all alchemical intermediates and stored every 0.5 ps. Particle mesh Ewald was employed with a real-space cutoff of 7 Å, interpolation order of 5, and a grid of 42 × 42 × 42 points. Dynamics were integrated using the "better Beeman" algorithm with a time step of 1 fs.

Correlation times were computed for the potential energy history and the trajectories subsampled to produce a set of uncorrelated samples. All recorded samples were processed with the multistate Bennett acceptance ratio (MBAR)[47] to estimate free energies and uncertainties for each leg of the thermodynamic cycle corresponding to transfer from 1 M gas to 1 M aqueous solution: discharging in vacuum, decoupling in water, and discharging in water. The first 1 ns of vacuum simulations and 50 ps of solvated simulations were discarded to equilibration, and the remainder (up to 4 ns for vacuum simulations)

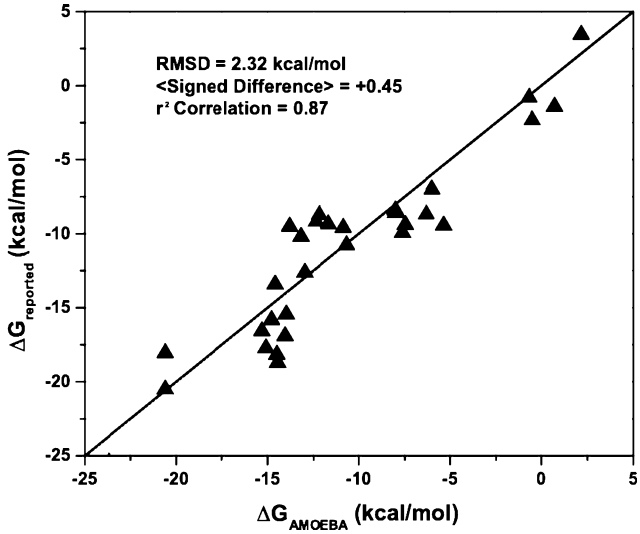


Figure 3. Comparison of the AMOEBA solvation free energies vs reported values from SAMPL2009. See Table 5 for details. All units are kcal/mol.

analyzed with MBAR; each leg of the thermodynamic cycle was processed individually, but all simulations within the leg were used together to obtain the most accurate estimates of free energies and their uncertainties.

Figure 3 shows the AMOEBA prediction against the OpenEye reported experimental literature values for a few classes of compounds, while Table 5 reports all of the results submitted to SAMPL2009. It is evident from Table 5 that AMOEBA did especially well in areas where traditional force fields failed, especially for very soluble molecules such as D-xylose and D-glucose. In addition, the 2009 SAMPL experimental values for glycerol and cyanuric acid were later revised (after submission of this manuscript), bringing AMOEBA into far better agreement with experiment. Other compounds such as the uracils, parabens, and NSAIDs have a range of reported experimental values. For example, the AMOEBA predictions for the uracils are between the SAMPL experimental values (taken from Cabani et al.[48]) and other more recently reported experimental values, suggesting that experimental uncertainty is much greater than the SAMPL error bars that are typically reported to be below 1 kcal/mol.

It is noteworthy that AMOEBA tended to do poorly on the polyhalogenated compounds, which typically have large atomic polarizabilities on the halogen atoms, values that were not derived in the original work by Thole. The AMOEBA force field derived the atomic polarizabilities for the halogens by fitting to just a couple of monohalogenated organic liquids. The results suggest that the reason AMOEBA underestimates the solvation free energy is that the atomic polarizabilities need to increase. The nitro compounds were also a challenge, with some evidence of large bond length changes between gas phase and liquid (as there are for amides!), as well as more complicated "push−pull" polarization that is not fully captured by the current "simple" polarization model.

Condensed Phase Structure and Dynamics

We have completed molecular dynamics simulations using nonpolarizable and polarizable protein force fields to contrast the water dynamics near hydrophilic, *N*-acetyl-glycine-methylamide (NAGMA), and amphiphilic, *N*-acetyl-leucine-methylamide (NALMA), peptides as a function of temperature, as models for understanding temperature dependent hydration dynamics near chemically heterogeneous protein surfaces.[49] These simulations are tightly coupled to X-ray diffraction and quasi-elastic neutron scattering (QENS) perfomed on these same systems at the same concentrations. Unlike a majority of macromolecular simulations that model a single solvated protein, these studies included ∼30−50 individual peptides that can interact with one another as well as the water molecules. The ability to accurately model the interactions of individual peptide fragments in a crowded solution is important for eventual studies of protein−ligand binding and protein−protein interactions, wherein the proteins can form temporary and reversible complexes. Hence, these peptide simulation studies represent an important biological environment with which to test any force field.

For the fixed charge case, we used the AMBER ff03[9b] all-atom protein force field and potential parameters to model the NALMA and NAGMA solutes, and the rigid, nonpolarizable TIP4P-Ew model[50] for the water. We have chosen a nonstandard protein−water model combination because we know that transport properties of TIP4P-Ew are excellent over a large temperature range, unlike the default TIP3P model typically used with biomolecular solutes. Unfortunately, we found that the simulated solution structure with nonpolarizable force fields predicts too much aggregation of both the hydrophobic and hydrophilic peptide solutes (Figure 4), in disagreement with our liquid diffraction experiments. This in turn frees up too much bulk-like water, so as to yield water diffusion constants that are faster and with an Arrhenius temperature dependence, contradicting our quasi-elastic neutron scattering experiments. However, when we fix the solutes to remain solvent-separated as that determined from the structural experiments, we find that the simulated hydration dynamics with the nonpolarizable force fields are close to quantitative with respect to the experimental dynamical trends with temperature for NAGMA (Figure 5a) and NALMA. It is clear that reparameterization of a biomolecular force field such as Amber ff03 (or other fixed charge force fields) to improve solvation properties using TIP4P-Ew is an important direction for future nonpolarizable force field efforts.

Due to the unphysical perturbation introduced by fixing the solutes, we also performed the same simulations with the AMOEBA polarizable force field.[13b] In contrast to the fixed-charge simulations, the polarizable force field nicely reproduces a nonaggregated, uniform distribution of solutes throughout the volume (Figure 4). It appears from these results that the ability of the peptides to respond dynamically to their electrostatic environment via polarization is important for reproducing a correct uniform mixture of peptides in water. Given the qualitative improvement in solution structure using the AMOEBA model, we also compared the changes in water dynamics as a function of temperature against our experimental data. On the basis of quasi-elastic neutron scattering (QENS) experiments, the amphiphilic NALMA peptide solution exhibits two translational relaxations at low temperatures, while the hydrophilic peptide shows only a single translational process, with transport properties of water near both peptide chemistries being very suppressed with respect to bulk dynamics.[49d,e] This is a real stress test for any force field given the range of dynamical trends that depend on amino acid chemistry and temperature. We note that we converged the induced dipoles very tightly in order to ensure energy conservation in the NVE ensemble under which we collected time correlation functions for calculating the diffusion coefficients.

AMOEBA provides reasonable agreement with the experimental temperature trends in regards to translational diffusion

TABLE 5: Comparison of AMOEBA Solvation Free Energies vs Reported Values from SAMPL2009[a]

molecule	AMOEBA	SAMPL2009	molecule	AMOEBA	SAMPL2009
cyanuric acid	−20.59 ± 0.30	−18.06	ibuprofen	−6.00 ± 0.39	−7.00
glycerol	−14.59 ± 0.72	−13.43	6-chlorouracil	−14.78 ± 0.37	−15.83
methyl paraben	−13.80 ± 0.44	−9.51	uracil	−15.30 ± 0.35	−16.59
butyl paraben	−12.16 ± 0.55	−8.72	5-trifluoromethyluracil	−13.97 ± 0.28	−15.46
ethyl paraben	−12.33 ± 0.56	−9.20	*d*-glucose	−23.69 ± 0.40	−25.47
naproxen	−13.17 ± 0.43	−10.21	hexachlorobenzene	−0.51 ± 0.34	−2.33
propyl paraben	−11.66 ± 0.57	−9.37	diflunisal	−7.47 ± 0.42	−9.40
octafluorocyclobutane	2.17 ± 0.22	3.43	hexachloroethane	0.72 ± 0.20	−1.41
phthalimide	−10.84 ± 0.38	−9.61	acetylsalicylic acid	−7.62 ± 0.39	−9.94
caffeine	−12.96 ± 0.47	−12.64	trimethylphosphate	−6.30 ± 0.28	−8.70
d-xylose	−20.60 ± 0.38	−20.52	5-chlorouracil	−15.08 ± 0.26	−17.74
ketoprofen	−10.67 ± 0.44	−10.78	5-fluorouracil	−14.05 ± 0.27	−16.92
trimethylorthotrifluoroacetate	−0.68 ± 0.28	−0.80	5-bromouracil	−14.49 ± 0.29	−18.17
flurbiprofen	−8.00 ± 0.45	−8.42	4-nitroaniline	−5.34 ± 0.34	−9.45
sulfolane	−7.99 ± 0.28	−8.61	5-iodoracil	−14.44 ± 0.27	−18.72

[a] The statistical uncertainty in the BAR free energy calculation is reported. All units are kcal/mol.

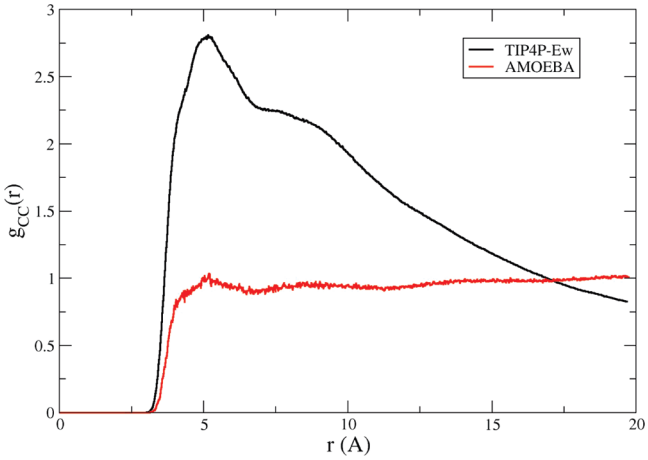


Figure 4. Solute carbon−carbon radial distribution functions for the 1 M NALMA solution at 298 K in the fixed charge (black) vs AMOEBA (red) force fields. Figure reproduced with permission from ref 49a. Copyright 2009 American Chemical Society.

for the glycine peptide (Figure 5b), although the dynamics are far too slow at the lowest temperatures for the amphiphilic NALMA peptide. Even so, calculations of the intermediate scattering function (ISF)

$$F_T^H(Q,t) = \langle \exp\{i\mathbf{Q}\cdot[\mathbf{r}_H(t) - \mathbf{r}_H(0)]\}\rangle \qquad (13)$$

using the AMOEBA model showed that the fits to its decay at low temperatures required two relaxation time scales for NALMA, while the same quantity calculated for NAGMA decayed with a single relaxation process. This reproduced the experimental trends observed in the QENS data with respect to peptide chemistry. What the AMOEBA simulations revealed is that the inner hydration layer nearest the amphiphilic solute relaxed on a much slower time scale than the outer hydration layers, while the hydrophilic peptide showed no differences in relaxation times in the two regions. Given that water dynamics for the amphiphilic peptide system reproduces all known rotational and translational hydration dynamical anomalies exhibited by hydration water near protein surfaces, our analysis using the AMOEBA model provided the critical evidence that hydration dynamics near biological interfaces is induced by chemical heterogeneity, as opposed to just topological roughness, of the protein surface.[49a]

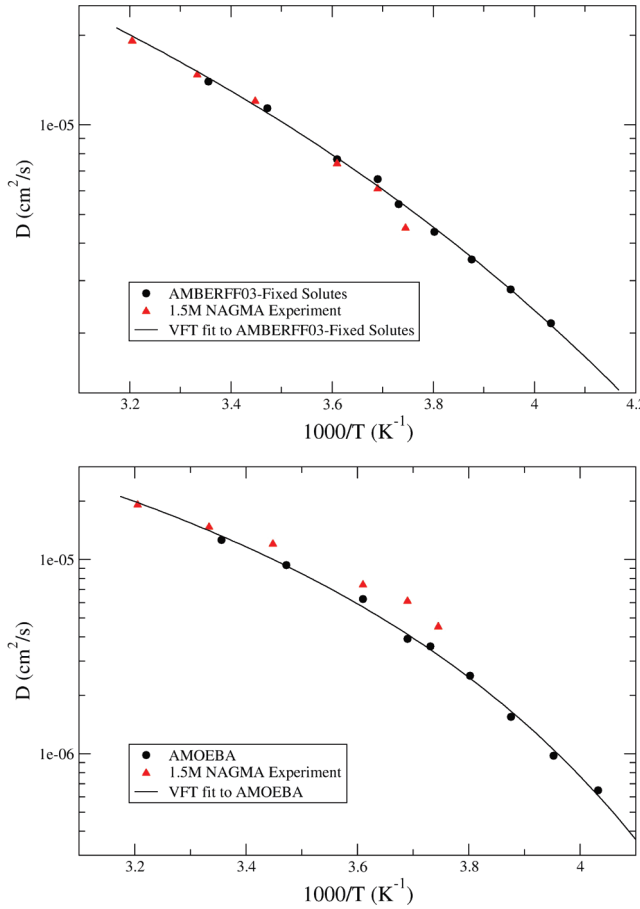


Figure 5. Arrhenius representation of the (a) fixed charge force field and (b) AMOEBA force field compared to the experimentally determined D_t for the 1.5 M NAGMA solution. The VFT fit (solid line) is to the simulation data (black circles). Figure reproduced with permission from ref 49a. Copyright American Chemical Society.

We have also used the AMOEBA polarizable model to investigate changes in solution structure and hydration dynamics of the 1 M NALMA peptide solution upon the addition of two small molecule cosolvents, the protein stabilizer glycerol and the protein denaturant dimethyl sulfoxide (DMSO).[49b] There continues to be debate in regards to the mechanism of protein stabilization or destabilization by cosolvents[51] (although that debate is often focused more on ionic additives). An indirect mechanism proposes that chaotropes disrupt water structure so as to enhance solubilization of hydrophobic groups, thus shifting

the equilibrium to the unfolded state, whereas kosmotropes increase water structure so as to diminish the solubilization of hydrophobic groups, thus stabilizing the folded state. A more direct mechanism proposes that chaotropes or denaturants preferentially bind to the protein, thereby dehydrating the protein surface to promote the unfolded state, while stabilizing kosmotropic agents do not interact with the biological macromolecule, leading to a preferential hydration of the protein surface that favors the folded state.

In our simulations, we found that, with the addition of DMSO, water was preferentially excluded from the hydrophobic leucine surface, while the opposite occurred with the addition of glycerol, consistent with experimental expectations.[52] While the AMOEBA simulated hydrogen bonds formed between water molecules and the peptide backbone agreed well with our neutron diffraction data for the glycerol solution,[49c] the simulated DMSO solution maintained peptide backbone−water hydrogen bonds, contradicting our experimental results, indicating a need to reparameterize the DMSO molecule to better reproduce solution properties. This was done in 2009 for the SAMPL competition, and it is clear that the new modified Lennard-Jones parameters show excellent agreement with solvation free energy data (Table 4), and we would expect that corresponding solution structure would improve as a result. Nonetheless, using the older parameter set, DMSO does displace water near the hydrophobic side chain, consistent with a preferential exclusion mechanism we found from our experiments.

For both cosolvent solutions, the quantitative values of the translational diffusion constants from the AMOEBA simulations were too slow compared to our QENS experiments[49c] for all temperatures studied. Clearly, there is strong directionality and longer hydrogen-bonding lifetimes between AMOEBA water and all solutes and cosolvents that explain why the diffusion constants of these solutions are an order of magnitude slower than the experiments. However, the observed dynamical trends were consistent with the experiment: mechanistically, we showed that the glycerol cosolvent preserves the hydration structure near the peptide, which in turn preserves the dynamical temperature trends of two water relaxation processes observed in the cosolvent free solution. By contrast, the DMSO solution disrupts the water structure near the peptide surface and destroys the inner hydration layer relaxation process, to show a single time scale for translational water dynamics that is consistent with experiment. Together, the AMOEBA theoretical model and the corresponding experiments showed that the direct mechanism was the most fully encompassing predictor of cosolvent behavior.

Protein Stability

As an initial evaluation of the AMOEBA force field for use in general protein simulation, the stability of some small globular proteins has been tested via a series of short molecular dynamics trajectories in aqueous solution. The proteins studied include crambin, villin headpiece, BPTI, Trp cage, GB3, and a SUMO-2 domain. All systems contained a single polypeptide without counterions in a periodic cubic box of AMOEBA water, ranging in size from 49 to 62 Å on a side, and chosen to provide a minimum of 10 Å of water between protein atoms and the closest box edge. Simulations were started from partially minimized systems, slowly heated in stages over 300−500 ps, and finally equilibrated at 298 K and 1 atm. Production simulations were then collected for 2−20 ns using 1.0 fs time steps under a modified Beeman integrator. van der Waals interactions were smoothly reduced to zero over a window from 10.8 to 12.0 Å. Multipole electrostatics and polarization were treated via particle-mech Ewald summation with a "tinfoil" boundary. Average production period rmsd values from the original PDB structure over backbone α-carbon atoms are reported in Table 6. While the rmsd from a reported crystal or NMR structure is a very imperfect measure of the overall quality and fidelity of a force field, these preliminary results show the promise of AMOEBA for modeling of larger biological structures. Some cursory comments are provided below, and more detailed analysis will be the subject of future work.

TABLE 6: Results of AMOEBA Protein Simulations Showing the Average α-Carbon RMSD between PDB Structure and MD Snapshots

protein	PDB code	no. of residues	simulation time (ns)	⟨rmsd⟩
Crambin	1EJG	46	19.6	0.73
Villin	1VII	36	18.1	1.80
BPTI	1BPI	58	2.0	1.20 (0.85)[a]
Trp Cage	1L2Y	20	5.0	1.40 (1.00)[b]
GB3	2OED	56	3.0	1.56
SUMO-2	1WM3	72	3.0	1.23

[a] rmsd computed over residues 1−56, omitting 57 and 58. [b] rmsd computed over residues 2−19, omitting 1 and 20.

The longest MD simulations, approaching 20 ns, were performed for the disulfide-containing crambin, and the three-helical villin headpiece. Crambin has an extremely hydrophobic sequence, and remains remarkably close to its high-resolution X-ray crystal structure throughout the AMOEBA simulation. The individual helices of villin generally remain intact across the simulation, but relative motions of the helices via their connecting hinge regions lead to a larger overall rmsd from the NMR-derived PDB structure. For both BPTI and Trp cage, a significant portion of the deviation from the PDB structure during the simulation is accounted for by fraying of the terminal residues. As indicated in Table 6, the α-carbon rmsd for each protein is reduced nearly one-third by omitting only two residues. The reported average rmsd for the relatively short simulation of GB3 is not converged, and this protein exhibits partial unfolding of an aromatic hydrophobic core at one end of the single domain, with some water infiltrating to solvate surface area occluded in the PDB structure. Another group (David Case, personal communication) has also noted a relatively high rmsd vs the NMR structure for GB3 in a short AMOEBA simulation performed with the Amber software package. Whether this is simply a random fluctuation in a short simulation, or a reproducible characteristic of GB3 modeled with AMOEBA is currently under investigation.

Protein−Ligand Binding

AMOEBA has been utilized in calculating the binding free energy between trypsin and a series of six benzamidine-like ligands.[53] The positively charged benzamidine and its derivatives form a salt bridge with the negatively charged D189 aspartic acid in the S1 site of trypsin.[54] The ability to capture the specific recognition between proteins and ligands requires an accurate description of atomic interactions between ligand−water and ligand−protein. The trypsin−benzamidine system has been selected for the study due to the availability of experimental data, the subtle chemical changes in the ligand series, the charged nature, and small size of the ligands. To calculate the absolute binding free energy of benzamidine to trypsin, free energy perturbation calculations have been performed using the AMOEBA potential for the protein, water, and ligand molecules.

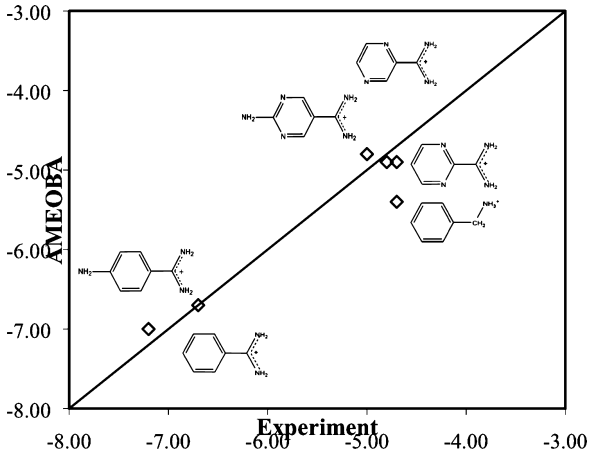


Figure 6. Comparison of experimental and calculated ligand binding free energy using AMOEBA potential. The ligand chemical structures are shown from left to right roughly according to their experimental binding free energy.

The interaction between the benzamidine and the environment (neat water or trypsin-in-water) was gradually decoupled via the scaling of the ligand electrostatic parameters (permanent multipole and polarizability) and the vdW interactions using a soft-core treatment following the double decoupling procedure.[55] Up to 3 ns MD simulations were performed at each of the 20 uniform decoupling steps. A rather large hydration free energy, −45.8 kcal/mol, was obtained for benzamidine. The total binding free energy was calculated to be 6.7 kcal/mol,[53b] in good agreement with the experimental value that ranges between −6.3 and −7.3 kcal/mol.[56]

To achieve quantitative understanding of the polarization effect as the benzamidine moves from water into the trypsin binding site, the dipole induction between benzamidine and water or trypsin-in-water was "turned off" to evaluate the polarization free energy. In this experiment, the "permanent" atomic multipoles in trypsin−water or benzamidine no longer polarized each other; however, the induction within water or trypsin−water remained, as it was an integral part of the potential. The calculations showed that the polarization between water and benzamidine was responsible for −4.5 kcal/mol out of the total −45.8 kcal/mol hydration free energy. In contrast, the polarization between trypsin-in-water and benzamidine weakened the attraction between benzamidine and trypsin by 22.4 kcal/mol. It may seem counterintuitive that turning on polarization would increase the system energy, as at any given state the polarization effect always lowers the system energy. However, in the "on" state, the trypsin-in-water sees both aspartic acid and benzamidine together as a dipole moment, whereas in the "off" state the system only sees a negatively charged aspartic acid, which gives rise to much more significant polarization. Thus, our observation indicates that, when the medium (trypsin-in-water in this case) is capable of responding via electronic polarization, it will screen the "permanent" electrostatic interaction.

In addition, the binding free energy of five ligands relative to benzamidine has been evaluated using AMOEBA via free energy perturbation[53b] (Figure 6). The RMSE of the computed binding free energy is 0.4 kcal/mol, and the largest error is 0.7 kcal/mol. When the amidine group in benzamidine is replaced by an amine or the phenyl ring is substituted by a diazine or an aniline, the free energy changes in both trypsin and water are on the order of several tens of kcal/mol and are mostly due to the electrostatic interaction.[53c] The two changes mostly cancel so that the relative binding free energy changes are on the order of 0.0−3.0 kcal/mol. These substitutions in the benzamidine also result in notable change in the molecular dipole moment of the ligand. There seems to be a correlation between the ligand molecular polarizability, instead of the molecular dipole moment, and its binding free energy.[53c] Note that the accuracy of the computed binding energy (RMSE = 0.4 kcal/mol) is slightly better than that of the hydration free energy of the 30 small molecules reported in Table 4 (RMSE = 0.68 kcal/mol). On the other hand, the HFE of the drug-like compounds in SAMPL (Table 5) show much greater error due to the various factors discussed in the previous section, such as uncertainty in the experimental data, problems with halogenated molecules and nitro compounds. The distributed multipole analysis used to derive atomic multipole moments has not been extensively tested on such molecular systems. Nonetheless, it is likely that better accuracy can be achieved in the binding free energy than in the individual solvation free energy in water or protein, as the systematic error may cancel between the two. Study of a broader range of protein−ligand complexes will be necessary to bring further insight.

The initial application of AMOEBA to protein−ligand binding suggests that the polarization effect plays an important role in the specific recognition, and the polarizable atomic multipole is able to capture the chemical details of the substituted benzamidine ligands. The fact that the finite binding free energy arises from a small difference between some large free energy of solvation stresses the need for a highly accurate interaction potential in order to achieve robust chemical accuracy in the binding free energy prediction. We are in the process of extending the AMOEBA polarizable model to a broader range of protein−ligand systems.

X-Ray Crystallography Refinement

X-ray crystallography is one of only a few experimental methods capable of yielding atomic resolution structural information. During refinement of a model against diffraction data, a force field offers a rich source of prior chemical knowledge. However, widely used crystallography programs such as CNS[57] and PHENIX[58] are not yet coupled to polarizable force fields. Furthermore, particle mesh Ewald (PME) summation is almost exclusively limited to P1 symmetry within biomolecular simulation codes.[59] Given these limitations, chemical features that are not seen clearly in the electron density are typically left out of the model. Our major goal in applying AMOEBA to X-ray crystallography is to consistently explain important structural features that are ambiguous based only on the experimental data.

The first work by Schnieders et al. in applying AMOEBA to crystallography focused on the development of a scattering model based on Cartesian Gaussian multipoles that for the first time allowed structure factors to be computed from an aspherical (i.e., multipolar) and anisotropic description of molecular electron density via FFT.[60] After beginning with peptide crystals, we scaled up to high resolution lysozyme, trypsin, and nucleic acid data sets to demonstrate that our AMOEBA based refinement method precisely orients water within hydrogen bonding networks while reducing R and R_{free} relative to deposited values by 5−6%.[61]

A limitation of this work has been the requirement to expand to P1 in order to use the TINKER energy and gradient routines. This motivates our current efforts to develop an AMOEBA code capable of taking advantage of space group symmetry to reduce memory requirements and accelerate the calculation of energies and gradients for any system size, unit cell dimensions, or space

TABLE 7: Timings for the Space Group PME Implementation of AMOEBA on a 3 Ghz 8 Core MacPro Workstation[a]

		number of atoms		time (s)	
PDB ID	space group	unit cell	asymmetric unit	expand to P1	native
1WQY	$P3_221$	5370	895	0.3	0.1
3DAI	$P6_522$	31752	2646	2.2	0.5
2J00	$P2_12_12_1$	1 954 100	488 525		72.4

[a] The self-consistent field was converged to 0.01 RMS Debye. Expansion to P1 for the ribosome system (2J00) was not attempted due to limited RAM.

group encountered in biomolecular crystallography. We have recently completed such an engine, named "Force Field X", in pure Java code. To the best of our knowledge, this represents the first formulation of AMOEBA or PME that includes support for all 230 space groups. Although the details of our space group PME version of AMOEBA are beyond the scope of the present article, we present timings in Table 7 to demonstrate that we have opened the door to routine use of AMOEBA within the X-ray crystallography community even for large, challenging data sets such as ribosome crystals.

Software Infrastructure for AMOEBA

The Ponder lab introduced AMOEBA as one of the first new polarizable protein and water force fields released in the public domain, and it is available in the TINKER package via the Web site http://dasher.wustl.edu. All of the force field parameters are made freely available to both academic and commercial parties. Currently, AMOEBA serves as the force field engine for a number of independent efforts, including the Folding@Home distributed computing project,[62] the TINKERATE kinetic rate calculation software,[63] and the GAMESS QM/MM implementation.[64] Versions are available for essentially all modern computer systems, and prebuilt executables are supplied for Linux, Windows, and Apple OSX. A User's Guide as well as many examples and test cases are available online. A variety of potential functions and parameter sets are available to the user, including MM2, MM3, AMBER, CHARMM, OPLS, OPLS-AA, and our own AMOEBA force field parameters. TINKER is one of only a few molecular mechanics packages to implement each of the major protein force fields within a single unified body of code. The AMOEBA force field also supports the study of nucleic acids and small molecules, and therefore, the TINKER package supports a more broad chemistry computational infrastructure of molecular models.

In 2009, we released the TINKER 5 modeling software and AMOEBA force field which contains a number of software improvements such as increased efficiency of particle mesh Ewald (PME) and neighbor list calculations, and shared-memory parallelization of the TINKER modeling package for the AMOEBA force field under the OpenMP protocol. Recent performance advances allow a speedup of 5.5X out of a possible 8 on commodity dual quad-core machines for AMOEBA-based MD simulation of medium-sized proteins in explicit water (~25 000 atoms). Simple application of Amdahl's law would indicate that nearly 95% of the computational cost of the TINKER simulation is now parallelized. Much of the speedup has been in the AMOEBA implementation of PME summation, but challenges also remain in this area. Future work will optimize the initial placement of AMOEBA multipoles onto the PME "charge" grid (the current parallelization bottleneck) as well as working on a shared memory spatial decomposition algorithm to achieve further speed gains. The Pande group is also working on accelerating AMOEBA and other molecular mechanics force fields on graphical processor units (GPUs), and is collaborating with the Simbios National Center for Biomedical Computing to disseminate the software. This collaboration has also led to novel methods for programming molecular dynamics on GPUs on implicit solvent and explicit solvent on fixed charge force fields.[65] This has been accomplished within the OpenMM software package, now available at http://simtk.org/home/openmm, and we hope to release the same for AMOEBA in 2010.

Two major new software programs have been added to the TINKER package with the goal of substantially automating development and refinement of AMOEBA force field parameters for arbitrary organic molecules. The first program, POTENTIAL, is a facility for comparing and fitting parameters to the electrostatic potential surrounding a molecule. It reads the potential from *ab initio* results or computes the potential from a force field model on a user-controlled radial grid. Force field parameters (partial charges, atomic multipoles) can be fit to potentials with a great deal of flexibility regarding terms and regions to optimize and parameter restraints. The program supports multiple molecular conformations and cluster configurations, which we find to be critical in obtaining robust parametrizations. The second new program, VALENCE, aids in determination of local valence force field parameters for bond stretching, angle bending, stretch−bend coupling, out-of-plane bending, and torsional amplitude. It also uses *ab initio* quantum results and operates in two modes: force fitting and structure fitting. Force fitting refines parameters for a static structure against the *ab initio* forces, and Hessian matrix. Structure fitting finds the force field parameters via repeated structural optimization and comparison to *ab initio* bond and angle values and vibrational frequencies. In addition, another program called TORSFIT is under development for the fitting of force field torsional parameters to energy benchmark results for rotation about specific bonds. Taken together, these programs represent a major advance in AMOEBA parameter development both in terms of accuracy and consistency of the resulting parameters. For example, using these tools, we were able to complete the AMOEBA parametrization of the 43 drug-like organic molecules in the 2009 SAMPL solvation free energy test set reported above in about 1 week.

Conclusions

Biomolecular simulations lie at the heart of physically driven atomistic approaches to computational biology. Empirical force fields are the core of all biomolecular simulations, with the computer programs that implement them, and together, they define the central community intellectual property and infrastructure in this field. While sustained advances in computing hardware have helped the broad adoption of simulation as an equal to theory and experiment, an equally important advance is the development of theoretical models that have proven predictive power. We have shown that the AMOEBA force field offers a significant improvement over nonpolarizable models for more accurate structural and thermodynamic of small protein-like fragments, and good transport properties such as diffusion constants near ambient temperatures. Given its parametrization strategy involving careful decomposition, AMOEBA shows excellent agreement with benchmark electronic structure data, and should be advocated as an excellent molecular mechanics choice for QM/MM schemes. Further fine-tuning is

necessary to describe solvation free energies of drug-like small molecules, dynamical properties away from ambient conditions, with possible further improvements on aromatic group interactions that may impact structural stability of proteins like GB3. However, AMOEBA has demonstrated that polarizability is a necessary intermolecular interaction for prediction of protein−ligand binding, and its improved treatment of electrostatics is likely to open up a new level of protein structural refinement in X-ray crystallography.

Acknowledgment. The work reported here is supported by a joint NSF Cyber-Infrastructure Award 0344670 to the Berkeley, Stanford, and Washington, St. Louis consortium. Development of the TINKER software program that implements the AMOEBA force field was supported by NIH grant R01 GM58712 to J.W.P. The protein-ligand binding study is supported by NIH R01GM079686 to P.R.

Supporting Information Available: The current standard parameter set for small molecules and proteins. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Ponder, J. W.; Case, D. A. Force fields for protein simulations. *Protein Simul.* **2003**, *66*, 27+.

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

(3) (a) Halgren, T. A.; Damm, W. Polarizable Force Fields. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236–242. (b) Rick, S. W.; Stuart, S. J. Potentials and Algorithms for Incorporating Polarizability in Computer Simulations. *Rev. Comput. Chem.* **2002**, *18*, 89–146. (c) 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. (d) 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* (4), 621–627. (e) Patel, S.; Brooks, C. L., 3rd. CHARMM fluctuating charge force field for proteins: I parameterization and application to bulk organic liquid simulations. *J. Comput. Chem.* **2004**, *25* (1), 1–15. (f) 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* (12), 1504–1514. (g) Harder, E.; Kim, B. C.; Friesner, R. A.; Berne, B. J. Efficient simulation method for polarizable protein force fields: application to the simulation of BPTI in liquid. *J. Chem. Theory Comput.* **2005**, *1*, 169–180. (h) 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. (i) Wang, Z. X.; Zhang, W.; Wu, C.; Lei, H.; Cieplak, P.; Duan, Y. Strike a balance: optimization of backbone torsion parameters of AMBER polarizable force field for simulations of proteins and peptides. *J. Comput. Chem.* **2006**, *27* (6), 781–790. (j) Cieplak, P.; Dupradeau, F.-Y.; Duan, Y.; Wang, J. Polarization Effects in Molecular Mechanical Force Fields. *J. Phys.: Condens. Matter* **2009**, *21*, 333102.

(4) Ponder, J. W. *TINKER: Software Tools for Molecular Design, 5.0*; Washington University School of Medicine: Saint Louis, MO, 2009.

(5) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber BioMolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

(6) Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **2009**, *30* (10), 1545–1614.

(7) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(8) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802.

(9) (a) Corongiu, G. Molecular Dynamics Simulation for Liquid Water Using a Polarizable and Flexible Potential. *Int. J. Quantum Chem.* **1992**, *42*, 1209–1235. (b) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24* (16), 1999–2012. (c) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65* (3), 712–725.

(10) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102* (18), 3586–3616.

(11) (a) van Gunsteren, W. F.; Daura, X.; Mark, A. E. GROMOS Force Field. *Encyclopedia of Computational Chemistry* **1998**, *2*, 1211–1216. (b) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25* (13), 1656–1676.

(12) (a) Jorgensen, W. L.; Tirado-Rives, J. The OPLS Potential Functions for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666. (b) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118* (45), 11225–11236. (c) Kaminski, G.; 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*, 6474–6487.

(13) (a) Ren, P. Y.; Ponder, J. W. Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. *J. Comput. Chem.* **2002**, *23* (16), 1497–1506. (b) Ren, P. Y.; Ponder, J. W. Polarizable atomic multipole water model for molecular mechanics simulation. *J. Phys. Chem. B* **2003**, *107* (24), 5933–5947. (c) Grossfield, A.; Ren, P. Y.; Ponder, J. W. Ion solvation thermodynamics from simulation with a polarizable force field. *J. Am. Chem. Soc.* **2003**, *125* (50), 15671–15682. (d) Ren, P. Y.; Ponder, J. W. Temperature and pressure dependence of the AMOEBA water model. *J. Phys. Chem. B* **2004**, *108* (35), 13427–13437. (e) Ren, P.; Wu, C.; Ponder, J. W. Polarizable atomic multipole-based intermolecular potentials for organic molecules. *J. Chem. Theory Comput.* **2010**, *submitted*.

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

(15) Astrand, P.-O.; Linse, P.; Karlstrom, G. Molecular dynamics study of water adopting a potential function with explicit atomic dipole moments and anisotropic polarizabilities. *Chem. Phys.* **1995**, *191*, 195–202.

(16) Swart, M.; van Duijnen, P. T. DRF90: A Polarizable Force Field. *Mol. Simul.* **2006**, *32*, 471–484.

(17) (a) Thole, B. T. Molecular Polarizabilities Calculated with a Modified Dipole Interaction. *Chem. Phys.* **1981**, *59*, 341–350. (b) van Duijnen, P. T.; Swart, M. Molecular and Atomic Polarizabilities: Thole's Model Revisited. *J. Phys. Chem. A* **1998**, *102* (14), 2399–2407.

(18) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. Molecular Mechanics. The MM3 Force-Field for Hydrocarbons 0.1. *J. Am. Chem. Soc.* **1989**, *111* (23), 8551–8566.

(19) Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations*; McGraw-Hill: New York, 1955.

(20) Bell, R. P. Bond Torsion in the Vibrations of the Benzene Molecule. *Trans. Faraday Soc.* **1945**, *41*, 293b–295.

(21) Mannfors, B.; Sundius, T.; Palmo, K.; Pietila, L. O.; Krimm, S. Spectroscopically determined force fields for macromolecules. Part 3. Alkene chains. *J. Mol. Struct.* **2000**, *521*, 49–75.

(22) Halgren, T. A. Representation of van der Waals (vdW) Interactions in Molecular Mechanics Force Fields: Potential Form, Combination Rules, and vdW Parameters. *J. Am. Chem. Soc.* **1992**, *114*, 7827–7843.

(23) Stewart, R. F.; Davidson, E. R.; Simpson, W. T. Coherent X-Ray Scattering of the Hydrogen Atom in the Hydrogen Molecule. *J. Chem. Phys.* **1965**, *42* (9), 3175–3187.

(24) Coppens, P.; Guru Row, T. N.; Leung, P.; Stevens, E. D.; Becker, P. J.; Yang, Y. W. Net Atomic Charges and Molecular Dipole Moments from Spherical-Atom X-Ray Refinements, and the Relation between Atomic Charge and Shape. *Acta Crystallogr.* **1979**, *A35* (1), 63–72.

(25) (a) Dykstra, C. E. Efficient Calculation of Electrically Based Intermolecular Potentials of Weakly Bonded Clusters. *J. Comput. Chem.* **1988**, *9*, 476–487. (b) Applequist, J. Traceless Cartesian Tensor Forms for

Spherical Harmonic-Functions - New Theorems and Applications to Electrostatics of Dielectric Media. *J. Phys. A: Math. Gen.* **1989**, *22* (20), 4303–4330.

(26) (a) Stone, A. J. Distributed Multipole Analysis, or How to Describe a Molecular Charge Distribution. *Chem. Phys. Lett.* **1981**, *83* (2), 233–239. (b) Stone, A. J. Distributed Multipole Analysis: Stability for Large Basis Sets. *J. Chem. Theory Comput.* **2005**, *1*, 1128–1132.

(27) Kong, Y. Multipole Electrostatic Methods for Protein Modeling with Reaction Field Treatment. Ph.D. thesis, Washington University Medical School, St. Louis, MO, 1997.

(28) Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: Oxford, U.K., 1996.

(29) Burnham, C. J.; Li, J. C.; Xantheas, S. S.; Leslie, M. The Parametrization of a Thole-type All-Atom Polarizable Water Model from First Principles and Its Application to the Study of Water Clusters (n=2−21) and the Phonon Spectrum of Ice Ih. *J. Chem. Phys.* **1999**, *110* (9), 4566–4581.

(30) Morita, A. Water polarizability in condensed phase: Ab initio evaluation by cluster approach. *J. Comput. Chem.* **2002**, *23* (15), 1466–1471.

(31) Morita, A.; Kato, S. An ab initio analysis of medium perturbation on molecular polarizabilities. *J. Chem. Phys.* **1999**, *110* (24), 11987–11998.

(32) Gubskaya, A. V.; Kusalik, P. G. The multipole polarizabilities and hyperpolarizabilities of the water molecule in liquid state: an ab initio study. *Mol. Phys.* **2001**, *99* (13), 1107–1120.

(33) (a) Gubskaya, A. V.; Kusalik, P. G. The total molecular dipole moment for liquid water. *J. Chem. Phys.* **2002**, *117* (11), 5290–5302. (b) Silvestrelli, P. L.; Parrinello, M. Water Molecule Dipole in the Gas and in the Liquid Phase. *Phys. Rev. Lett.* **1999**, *82* (16), 3308–3311.

(34) Head-Gordon, M.; Artacho, E. Chemistry on the Computer. *Phys. Today* **2008**, *61*, 58–63.

(35) Helgaker, T.; Klopper, W.; Tew, D. P. Quantitative Quantum Chemistry. *Mol. Phys.* **2008**, *106*, 2107–2143.

(36) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. Accurate ab Initio Quantum Chemical Determination of the Relative Energetics of Peptide Conformations and Assessment of Empirical Force Fields. *J. Am. Chem. Soc.* **1997**, *119* (25), 5908–5920.

(37) (a) DiStasio, R. A.; Jung, Y.; Head-Gordon, M. A Resolution-Of-The-Identity Implementation of the Local Triatomics-In-Molecules Model for Second-Order Møller-Plesset Perturbation Theory with Application to Alanine Tetrapeptide Conformational Energies. *J. Chem. Theory Comput.* **2005**, *1*, 862−876. (b) DiStasio, R. A., Jr.; Steele, R. P.; Rhee, Y. M.; Shao, Y.; Head-Gordon, M. An improved algorithm for analytical gradient evaluation in resolution-of-the-identity second-order Moller-Plesset perturbation theory: application to alanine tetrapeptide conformational analysis. *J. Comput. Chem.* **2007**, *28* (5), 839–856.

(38) Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 084106.

(39) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom−atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

(40) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. Extremely Precise Free Energy Calculations of Amino Acid Side Chain Analogs: Comparison of Common Molecular Mechanics Force Fields for Proteins. *J. Chem. Phys.* **2003**, *119*, 5740–5761.

(41) (a) Bennett, C. H. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comput. Phys.* **1976**, *22*, 245–268. (b) Jiao, D.; King, C.; Grossfield, A.; Darden, T. A.; Ren, P. Y. Simulation of Ca2+ and Mg2+ solvation using polarizable atomic multipole potential. *J. Phys. Chem. B* **2006**, *110* (37), 18553–18559. (c) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. Equilibrium free energies from nonequilibrium measurements using maximum likelihood methods. *Phys. Rev. Lett.* **2003**, *91* (14), 140601–140604.

(42) (a) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. Group contribution to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution. *J. Solution Chem.* **1981**, *10*, 563–595. (b) Wolfenden, R.; Liang, Y.-L.; Matthews, M.; Williams, R. Cooperativity and anti-cooperativity in solvation by water: imidazoles, quinones, nitrophenols, nitrophenolate, and nitrothiophenolate ions. *J. Am. Chem. Soc.* **1987**, *109* (2), 463–466. (c) Wolfenden, R. Interaction of the peptide bond with solvent water: A vapor phase analysis. *Biochemistry* **1978**, *17*, 201–204. (d) Abraham, M. H.; Whiting, G. S. Thermodynamics of solute transfer from water to hexadecane. *J. Chem. Soc. Perkin Trans* **1990**, *2*, 291–300.

(43) Nicholls, A.; Mobley, D. L.; Guthrie, J. P.; Chodera, J. D.; Bayly, C. I.; Cooper, M. D.; Pande, V. S. Predicting small-molecule solvation free energies: An informal blind test for computational chemistry. *J. Med. Chem.* **2008**, *51* (4), 769–779.

(44) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. Comparison of charge models for fixed-charge forcefields: small molecule hydration free energies in explicit solvent. *J. Phys. Chem. B* **2007**, *111*, 2242–2254.

(45) Shirts, M. R.; Mobley, D. L.; Chodera, J. D.; Pande, V. S. Accurate and efficient corrections for missing dispersion interactions in Molecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 13052−13063.

(46) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(47) Minh, D. D. L.; Chodera, J. D. Optimal estimators and asymptotic variances for nonequilibrium path-ensemble averages. *J. Chem. Phys.* **2009**, *131*, 134110.

(48) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. Group Contributions to the Thermodynamic Properties of Non-Ionic Organic Solutes in Dilute Aqueous Solution. *J. Solution Chem.* **1981**, *10*, 563–595.

(49) (a) Johnson, M. E.; Malardier-Jugroot, C.; Murarka, R. K.; Head-Gordon, T. Hydration water dynamics near biological interfaces. *J. Phys. Chem. B* **2009**, *113*, 4082–4092. (b) Johnson, M. E.; Malardier-Jugroot, C.; Head-Gordon, T. Effects of co-solvents on peptide hydration water structure and dynamics. *Phys. Chem. Chem. Phys.* **2010**, *12* (2), 393−405. (c) Malardier-Jugroot, C.; Bowron, D. T.; Soper, A. K.; Johnson, M. E.; Head-Gordon, T. Structure and dynamics of aqueous peptide solutions in the presence of co-solvents. *Phys. Chem. Chem. Phys.* **2010**, *12* (2), 382−392. (d) Malardier-Jugroot, C.; Head-Gordon, T. Separable cooperative and localized translational motions of water confined by a chemically heterogeneous environment. *Phys. Chem. Chem. Phys.* **2007**, *9* (16), 1962–1971. (e) Malardier-Jugroot, C.; Johnson, M. E.; Murarka, R. K.; Head-Gordon, T. Aqueous peptides as experimental models for hydration water dynamics near protein surfaces. *Phys. Chem. Chem. Phys.* **2008**, *10* (32), 4903–4908.

(50) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for bio-molecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120* (20), 9665–9678.

(51) (a) Marcus, Y. Effect of ions on the structure of water: Structure making and breaking. *Chem. Rev.* **2009**, *109*, 1346–1370. (b) Auton, M.; Bolen, D. W.; Rösgen, J. Structural thermodynamics of protein preferential solvation: osmolyte solvation of Proteins, aminoacids, and peptides. *Proteins* **2008**, *73* (4), 802–813.

(52) Timasheff, S. N. The Control of Protein Stability and Association by Weak-Interactions with Water - How Do Solvents Affect These Processes. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 67–97.

(53) (a) Jiao, D.; Golubkov, P. A.; Darden, T. A.; Ren, P. Calculation of protein-ligand binding free energy by using a polarizable potential. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6290–6295. (b) Jiao, D.; Zhang, J.; Duke, R. E.; Li, G.; Schnieders, M. J.; Ren, P. Trypsin-Ligand Binding Free Energies from Explicit and Implicit Solvent Simulations with Polarizable Potential. *J. Comput. Chem.* **2009**, *30*, 1701–1711. (c) Shi, Y.; Jiao, D.; Schnieders, M. J.; Ren, P. Trypsin-Ligand Binding Free Energy Calculation with AMOEBA. *IEEE Engineering in Medicine and Biology Society* 2009, EMBC proceedings, 2328−2331.

(54) Katz, B. A.; Finer-Moore, J.; Mortezaei, R.; Rich, D. H.; Stroud, R. M. Episelection - Novel K-I-Similar-to Nanomolar Inhibitors of Serine Proteases Selected by Binding or Chemistry on an Enzyme Surface. *Biochemistry* **1995**, *34*, 8264–8280.

(55) (a) Boresch, S.; Karplus, M. Absolute binding free energies: A quantitative approach for their calculation. *J. Phys. Chem. B* **2003**, *107*, 9535–9551. (b) Hamelberg, D.; McCammon, J. A. Standard free energy of releasing a localized water molecule from the binding pockets of proteins: Double-decoupling method. *J. Am. Chem. Soc.* **2004**, *126*, 7683–7689.

(56) (a) Katz, B. A.; Elroda, K.; Luonga, C.; Ricea, M. J.; Mackmana, R. L.; Sprengelera, P. A.; Spencera, J.; Hatayea, J.; Janca, J.; Linka, J.; Litvaka, J.; Raia, R.; Ricea, K.; Siderisa, S.; Vernera, E.; Young, W. A novel serine protease inhibition motif involving a multi-centered short hydrogen bonding network at the active site. *J. Mol. Biol.* **2001**, *307*, 1451–1486. (b) Schwarzl, S. M.; Tschopp, T. B.; Smith, J. C.; Fischer, S. Can the calculation of ligand binding free energies be improved with continuum solvent electrostatics and an ideal-gas entropy correction? *J. Comput. Chem.* **2002**, *23*, 1143–1149.

(57) Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* **2007**, *2*, 2728.

(58) Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Loerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr., Sect. D* **2002**, *58*, 1948.

(59) (a) Sagui, C.; Pedersen, L. G.; Darden, T. A. Towards an accurate representation of electrostatics in classical force fields: efficient implementation of multipolar interactions in biomolecular simulations. *J. Chem. Phys.* **2004**, *120*, 73. (b) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577. (c) Darden, T.; York, D.; Pedersen, L. Particle mesh ewald: An n log(n) method for ewald sums. *J. Chem. Phys.* **1993**, *98*, 10089.

(60) Schnieders, M. J.; Fenn, T. D.; Pande, V. S.; Brunger, A. T. Polarizable atomic multipole X-ray refinement: application to peptide crystals. *Acta Crystallogr., Sect. D* **2009**, *65*, 952.

(61) Fenn, T. D.; Schnieders, M. J.; Brunger, A. T.; Pande, V. S. *J. Mol. Biol.*, submitted for publication, **2009**.

(62) Larson, S. M.; Snow, C.; Pande, V. S. *Folding@Home and Genome@Home: Using distributed computing to tackle previously intractable problems in computational biology*; Horizon Press: New York, 2003.

(63) Tishchenko, O.; Higashi, M.; Albu, T. V.; Corchado, J. C.; Kim, Y.; Vill, J.; Xing, J.; Lin, H.; Truhlar, D. G. *MC-Tinker*; University of Minnesota: Minneapolis, MN, 2009.

(64) Gordon; M. S. Schmidt, M. W. *Advances in electronic structure theory: GAMESS a decade later*; Elsevier: Amsterdam, The Netherlands, 2005.

(65) Friedrichs, M.; Eastman, P.; Vaidyanathan, V.; Houston, M.; LeGrand, S.; Beberg, A.; Ensign, D.; Bruns, C.; Pande, V. S. Accelerating Molecular Dynamic Simulation on Graphics Processing Units. *J. Comput. Chem.* **2009**, *30*, 864–872.

JP910674D