Title
Advanced Potential Energy Surfaces for Condensed Phase Simulation

Journal
Annual Review of Physical Chemistry, 65(1)

ISSN
0066-426X

Authors
Demerdash, Omar
Yap, Eng-Hui
Head-Gordon, Teresa

Publication Date
2014-04-01

DOI
10.1146/annurev-physchem-040412-110040

Peer reviewed

Advanced Potential Energy Surfaces for Condensed Phase Simulation

Omar Demerdash[1], Eng-Hui Yap[2], and Teresa Head-Gordon[1,2,3,4*]

[1]Department of Chemistry, [2]Department of Bioengineering, [3]Department of Chemical and Biomolecular Engineering, [4]Chemical Sciences Division, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720

ABSTRACT

Computational modeling at the atomistic and mesoscopic levels has undergone dramatic development in the last 10 years in order to meet the challenge of adequately accounting for the many-body nature of intermolecular interactions. At the heart of this challenge is the ability to identify the strengths and specific limitations of pairwise-additive interactions, and improving classical models to explicitly account for many-body effects, and consequently enhance their ability to describe a wider range of reference data and to build confidence in their predictive capacity. However the corresponding computational cost of these advanced classical models increases significantly enough so that statistical convergence of condensed phase observables becomes more difficult to achieve. Here we review a hierarchy of potential energy surface models used in molecular simulations for systems with many degrees of freedom that best meet the trade off between accuracy and computational speed in order to define a “sweet spot” for a given scientific problem of interest.

***Corresponding author: thg@berkeley.edu**

KEYWORDS

Many-body interactions, polarization, Poisson-Boltzmann, electrostatics, empirical force field

1. INTRODUCTION

Molecular simulation has realized broad adoption by academic researchers and industry scientists due in part to the tractability of the classical model assumption of pairwise additivity of molecular interactions. Molecular models, or force fields, are widely available and often highly successful on a variety of problems in chemistry, biochemistry and materials science. And yet we know pairwise additive treatments do break down, for example for heterogeneous environments or in areas of the phase diagram outside of which they were parameterized, although we can't always anticipate why or for which chemical systems. In principle, mutually polarizable models offer a significant improvement in the physics of multi-body electrostatic interactions. However the corresponding cost of an advanced polarizable force field increases significantly enough so that statistical convergence of condensed phase observables using molecular dynamics (MD), the so-called sampling problem, becomes more difficult to achieve.

While MD simulations routinely and beneficially address questions centered around detailed nanoscale chemistry, there is another set of problems on the supramolecular or mesoscale where the limitations of size and timescales in MD are reached. Coarse graining of the participating nanoscale structures and their environment, combined with simulations using stochastic dynamics, can be just as insightful as the all-atom deterministic dynamics when this regime is reached. Solutions to the Poisson Boltzmann equation (PBE) provides an example of a very different electrostatic model to that of an atomistic polarizable model, reducing molecular features of complex macromolecules in salty aqueous solvent to embedded charge distributions polarized by a dielectric continuum with the ability to describe ionic strength of the solution. However the accuracy of the PBE solution degrades rapidly as the coarse-grained system geometries become too complex and/or the lengthscales become too large, limiting their application to the mesoscale regimes for which they are most needed.

These two extremes in molecular modeling would seem to have little in common for a unified review, but here we define a theme around advances made in the treatment of electrostatics and polarization at the two scales of interest. These advanced potential energy surfaces that now include mutual polarization have encountered obstacles that inhibit their application to challenging chemical problems including accuracy and computational cost of the theoretical models and lack of innovation in better algorithms or controlled approximations that can mitigate the cost. However there have been important strides made in these areas that we

anticipate signals a new era in more widespread development and use of advanced potential energy surfaces in condensed phase simulations. This warrants an appraisal of the current state of the art in the area of atomistic force fields and PBE solvers, both of which can be formulated to better describe anisotropic interactions and many-body polarization.

2. RECENT ADVANCES IN ATOMISTIC FORCE FIELDS

A critical aspect of accurate modeling of physical properties of condensed systems is the development of models that reflect the complex potential energy surface of these systems. Naturally, these systems are described by the solution of the many-electron Schrödinger equation, but this is computationally intractable for many degrees of freedom simulated over long timescales, and we must resort to the development of empirical force fields for use with classical Newtonian mechanics. The typical components of a classical force field include

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

where the first five terms describe covalent interactions and the last three terms are the non-covalent contributions. U_{bond} and U_{angle} are commonly modeled with harmonic functions, or carry higher order terms to model anharmonicity as in force fields such as AMOEBA (1) and MM3 (2), and $U_{torsion}$ is typically modeled with a sinusoidal form. The second two covalent terms are less commonly included. $U_{b\theta}$, describes coupling between bond stretching and angle bending, which is important for reproducing vibrational frequencies. U_{oop} is a term to restrict out of plane bending to enforce planar geometry at sp2 hybridized centers, and is modeled with a harmonic Wilson-Decius-Cross term (3) in the AMOEBA force field (1).

While the description and parameterization of short-range valence terms is relatively straightforward, modeling of short- and long-range non-covalent interactions has remained more elusive. This is due in large part to the many-body, non-additive nature of these interactions. When describing complex intermolecular interactions, it is convenient and physically justified to distinguish between short-range interactions that decay exponentially as a function of distance R (i.e., as $exp(-\alpha R)$) and those that decay over a longer range as inverse powers of R, (i.e., as $1/R^n$) (4). Short-range many-body effects include exchange-repulsion and charge transfer and are best understood using quantum mechanics (QM), as the distance range over which these effects are strong is short enough to allow for overlap of a molecule's wavefunction with that of a nearby molecule. In most current force fields, only exchange repulsion is routinely represented

explicitly, typically as the repulsive $1/R^{12}$ component of a Lennard-Jones potential, and infrequently in an exponential form, such as Buckingham potential, or as Halgren's buffered 14-7 potential (5), as in the AMOEBA force field (1, 6).

Longer ranged effects include dispersion, electrostatics, and induction or polarization. The electrostatic energy arises from the fixed charge distributions of the molecules and this interaction decays very slowly (as $1/R$) for monopole-monopole interactions; interactions between higher order multipoles decay more rapidly but are still long-range. In most present day fixed charge force fields, charge distributions are approximated as atom-centered partial charges obtained by a fit to a QM-derived potential, and the long-range decay is accounted for through the Ewald sum. Within this pairwise-additive approximation we show in Section 2.1 that immediate improvements in predicted properties can be made in biomolecular simulation through direct reparameterization of existing protein force fields with more quantitative water models developed under the Ewald approximation.

While one may naively conclude that the additive nature of the electrostatic interaction renders it straightforward to model in classical force fields, a challenge arises in how one describes the charge distribution. Charge distributions approximated as atom-centered partial charges are a severe simplification of the complex orbital picture of real molecules which may have lone pair electrons that are not atom-centered or diffuse, delocalized electron clouds, as in the case of π orbitals in aromatic systems. Modeling efforts in the development of charge distributions that extend beyond the standard atom-centered charges, i.e. the introduction of the distributed multipole analysis (DMA) of Stone(4, 7), will be reviewed in Section 2.2.

Dispersion and polarization are both many-body, non-additive effects, although dispersion is typically approximated using a pairwise additive term, decaying as $-1/R^6$ or $-1/R^7$ in the Halgren potential. A notable exception to the approximation of pairwise additivity in dispersion is the well-known three-body Axilrod-Teller potential and more recently in other formulations of three-body potentials (8, 9). Polarization arises from the distortion of a molecule's charge distribution in the presence of an external field and is also non-additive in nature. It is longer in range than dispersion, as induced dipole-induced dipole interactions decay as $-1/R^4$. Current efforts in force field enhancement with respect to treatment of intermolecular interactions have focused on polarization since it is the slowest decaying of the aforementioned

many-body effects and therefore justifiably receives special focus in force field development as reviewed in Section 2.3.

2.1. New Compatible Protein and Water Force Fields

Empirical, fixed point-charge force fields for biomolecular simulation have undergone significant development since the pioneering work of Lifson and coworkers in the development of their Consistent Force Field (10-12). This has consisted of a large-scale, community-wide effort to refine force field parameters, extend the parameter sets to encompass biologically important small molecules, carbohydrates, and lipids, and appraise force field quality by benchmarking against experimental data. In particular, fixed charge force fields will certainly continue to be a frontline chemical model when the demands of sampling is paramount, such as in the simulation of protein folding or characterizing sluggish dynamics at cold temperatures.

Enhancements in computing power have also enabled pairwise-additive force fields to be assessed in terms of their ability to adequately describe the behavior of biological molecules on longer time scales. Such important “stress tests” include structural stability tests, free energy calculations, and protein folding simulations. Solvation free energy studies of protein side chains (13) and small organic molecules (14) both showed that solvation free energies were systematically too positive on average. In a study of structure and hydration dynamics in concentrated peptide solutions, Johnson et al. showed that fixed-charge force fields predicted too much aggregation of the peptides(15). The excessive aggregation is an indicator that significant improvements of fixed-charge force fields are required to correctly model extended peptide conformations, a necessity for protein folding simulations and protein-protein interactions. Interestingly, the same simulations performed with the AMOEBA polarizable force field did not yield unphysical aggregation (15).

In addition to performance in modeling solvation, fixed-charge force fields have also been assessed regarding their ability to reproduce key structural features of proteins, namely, experimentally determined dihedral angle propensities. Compared with the other bonded degrees of freedom in typical force fields, dihedral angles exhibit the greatest deviations from their equilibrium values in protein motions and thus are of utmost importance in elucidating large-scale conformational changes and folding/unfolding transitions. A number of deficiencies have been reported. Jiang et al. showed that simulated peptides exhibited deviations in their χ-angle

propensities from those of a library of χ angles culled from the Protein Data Bank (PDB) (16); this discrepancy extends to backbone dihedral angles. Indeed, Hummer and coworkers found that the AMBER protein has a bias towards predicting α–helices, and there is negligible dependence of helix formation on temperature, that is contrary to NMR data(17, 18).

More under recognized is that these deficiencies may stem in part from the model of water that is used, the primary molecular solvent for biological macromolecules. The early and successful water models such as the TIP models by Jorgensen and co-workers (19, 20) and the SPC variants by Berendsen and co-workers (21, 22) were designed to work under non-bonded interaction schemes involving simple spatial truncation of electrostatics and van der Waals interactions. Given that one of their primary design purposes was the ability to simulate systems with a large number of degrees of freedom, these early water models sought to describe bulk and solvation properties with a computationally tractable r-space cutoff of 9-12Å. Even though monopole-monopole electrostatic interactions are still significant at this spatial distance, model compensations could be made by adjusting the parameters to reproduce primarily *ambient* bulk water properties in the context of the spatial truncations. While the TIP and SPC models were a significant step forward in the description of bulk water at the time, their simulated discrepancies with experimental reference data over a range of temperature and pressure were harbingers for their limitations as an aqueous solvent model in biomolecular simulation.

The original TIP and SPC bulk water models have been superseded by more sophisticated parameterization schemes that account for the missing van der Waals interactions ignored beyond the cutoff, are optimized over a range of temperatures and/or pressures, and in some cases describe long-ranged electrostatics through the Ewald approximation. Examples of these newer generation non-polarizable water models include TIP4P-Ew (23) and TIP4P/2005 (24) and TIP5P (25) (which, unlike the former two models, continues to use truncation schemes). Even so, a vast majority of biomolecular simulations continue to use the TIP3P model as the default aqueous solvent model, and furthermore abuse its original parameterization conditions by combining it with particle mesh Ewald and long-range van der Waals corrections- simulation conditions that have been shown to degrade the accuracy of the TIP3P solvent model (26)!

Currently there is a move toward adopting water models such as TIP4P-Ew and TIP4P/2005 in biomolecular simulations since early evidence has shown that they are providing quantitative improvements in performance. For example, the simulation of dynamical T1 and T2

spin relaxations and ROESY intensities for the Aβ21-30 peptide by Fawzi and co-workers(27) found that the combination of AMBER ff99SB and the TIP4P-Ew model yielded far better predictions than the default TIP3P model, and similar conclusions were reached that a range of NMR data was better predicted for Aβ40 and Aβ42 using TIP4P-Ew compared to TIP3P (28-30). Using the same ff99SB protein force field, Wickstrom et al. showed that TIP4P-Ew predicted NMR J-couplings for $(Ala)_3$ and $(Ala)_5$ with better accuracy than that generated with TIP3P (31). Hydrophobic solvation was also found to be improved under the TIP4P-Ew model (32).

Given these positive initial results that a better water model could yield improvements in simulated properties without a single parameter change has in turn led to more recent work which seeks to make these water models more fully compatible over a range of structural, thermodynamic, and dynamical data. Best and Mittal developed a correction to the ψ backbone dihedral angle potential for AMBER ff03 when simulated with the TIP4P/2005 water model to give a more cooperative helix-to-coil transition and a more realistic collapse of the unfoldeded state with increasing temperature (33). Nerenberg and co-workers developed a perturbation to the ϕ' backbone dihedral potential of the ff99SB protein force field, so that when combined with the TIP4P-Ew model, yields agreement for a diverse set of J-coupling data over a temperature range of 275-350K for $(Gly)_3$, $(Ala)_3$, and $(Val)_3$ (34).

Nerenberg et al. have continued to advance a larger goal of creating a next generation fixed-charge protein force field combined with the TIP4P-Ew water model by next optimizing solute-water van der Waals interactions to reproduce experimental solvation free energy data (35). It is well known that current fixed charge force fields yield aqueous solvation free energies that are *systematically* unfavorable by ~1.0-2.5 kcal/mole (Figure 1). To better interface the protein and water force fields that were developed in independent communities, the parameterization strategy adopted in (35) was to jettison the use of simple mixing rules for the Lennard-Jones parameters in favor of optimizing the solute-solvent well depths and van der Waals radii directly against experimental solvation free energy data. Although there are now additional parameters in the new protein-water force field, the potential problem of overfitting was reduced in two ways; first by a novel optimization approach and secondly because the relative abundance of solution phase data allowed for the formulations of both a training set and independent testing set of the solvation free energies of model side chain chemistries(35). Such an "alternative" philosophy of force field parameterization- i.e. parameterization using complex

condensed phase experimental data as opposed to relying solely on neat organic liquid or gas phase *ab initio* data- appears to be gaining traction within the simulation community (18, 33-36). Figure 1 shows that Nerenberg and co-workers(35) were able to reproduce the solvation free energies of the *test* set of side chain analogue molecules to within 1.0 k_BT and with random error, while simultaneously reducing the aggregation propensities of dipeptide-water solutions, as well as improving the conformational preferences of short disordered peptides.

Not surprisingly, the ubiquitin protein unfolded under the new more favorable solute-water interactions(35)- which is to be expected as the protein-protein interactions are to be the final stage of parameterization under the ff99SB-TIP4P-Ew force field. Encouragingly, the hydrophobic core of ubiquitin remained intact under the new solute-solvent model, suggesting that the protein-protein van der Waals interactions are still relatively well-balanced with the protein-water van der Waals interactions. Instead it was found that protein-protein hydrogen bonds unraveled in parts of ubiquitin. By resurrecting the 10-12 potential for protein-protein backbone hydrogen bonds, ubiquitin remained stable over a 100 ns trajectory and maintained near quantitative agreement with experimentally measured Lipari-Szabo order parameters(35). Therefore, it is worthwhile considering for the future the best functional form for capturing short-ranged directionality and cooperativity, the essence of hydrogen-bonding, as discussed in the following section.

2.2. Next Generation Fixed Charge Models

Due to their long-range character and consequent strong influence on molecular association, solvation, and phase behavior, proper treatment of permanent electrostatic effects is of paramount importance. Naturally, a key component of this is the determination of fixed charges associated with each atom that are accurate within the framework of Coulombic electrostatics. This poses a strong challenge, as the complex orbital description of the electron density must be reduced to a set of fixed atom-centered point charges that are determined from a restrained fitting of a QM determined electrostatic potential (37, 38). While this approach is feasible for the determination of point charges, it is difficult to apply it to the case of multipole models, as the increased number of parameters associated with multipoles poses a difficult fitting challenge.

Rather than fitting to an electrostatic potential, Stone (7) developed an approach known as distributed multipole analysis (DMA) where multipoles are derived from the QM determined electron density according to the following equation:

$$Q_{LM}(s) = \int R_{LM}(r-s)\,\rho(r)\,dv \tag{2}$$

where R_{LM} is a spherical harmonic operator, r is the spatial coordinate, s is the origin, ρ is the charge density (in turn determined by the many-electron wavefunction), and dv is the differential volume element. This is simply an alternative expression of the multipole expansion which has been discussed in other excellent reviews (39, 40). A key feature of DMA is the use of multiple sites S for the multipole expansion, typically at atom centers, allowing for the multipole expansion to be accurate at short intermolecular distances, as opposed to a regime in which there is only one site S, in which case the multipole expansion will be convergent far outside the distance range of interest at which intermolecular association occurs.

Early studies of point multipole models showed that use of higher-order multipoles yielded better agreement with the quantum mechanical reference potential than point charges (41). Dykstra demonstrated the importance of higher order multipoles in more accurately reproducing the electrostatic potential just outside the van der Waals surface in particular, which is the distance range over which key determinants of specific intermolecular association such as hydrogen bonding occur(42). Fowler and Buckingham showed that multipoles are essential to obtain correct geometries of small molecule complexes that are dominated by hydrogen bonds or non-polar interactions (43, 44). Beyond offering an ostensibly more realistic approximation of a molecule's true electron distribution, use of off-center atomic charges, such as derived by DMA, have been shown to provide more accurate modeling of physical properties(45, 46) such as improved crystallographic refinement and agreement with electron densities.

The multipole model has been adopted by a number of force fields, including AMOEBA(1, 6, 47), SIBFA(48), and NEMO(49). Other electrostatic models have been adopted, including augmenting atom-centered charges with lone-pair sites (50), use of exponential charge distributions in QMPFF (51), and Gaussian charge distributions for the monopole(48). While different approaches to representing charge distributions have their strengths and weaknesses, the multipole method is attractive due to its straightforward connection with classical multipole expansions, which is nothing more than a Taylor expansion of the potential about a point(s) within the molecule's charge distribution.

Hydrogen bonding is a non-covalent interaction that consists of both electrostatics, induction, and charge-transfer components (although the relative contributions of each is a

current topic of debate!), whose description may benefit greatly by approaches that go beyond atom-centered charges. Known deficits in prediction of protein conformational preferences by current fixed-charge force fields, which typically lack explicit hydrogen bonding, have been attributed to failure to account for directional electrostatic interactions, explicitly or implicitly (18). The geometry dependence of hydrogen bonding has typically been modeled with potentials dependent on the donor-hydrogen-acceptor angle in force fields that include an explicit term for hydrogen bonding. Force fields that include an explicit hydrogen bonding term have demonstrated improved NMR protein structure prediction (52) and protein-protein complex prediction (53). Therefore it is reasonable to hypothesize that hydrogen bonding may benefit from an electrostatic model that affords directional dependency, such as the multipole model. This elegant approach would need to be weighed against simpler geometric definitions of explicit hydrogen bonding, which would be cheaper to evaluate but may be inherently limited in accuracy.

2.3. Mutual Polarization Models

The importance of many-body effects in intermolecular energies has been appreciated by the molecular modeling field for quite some time, beginning with the seminal work of Axilrod, Teller, and Muto, where they developed a three-body potential for the dispersion interaction (54, 55). Therefore, taking many-body interactions into account has become a vibrant area of research, both within the community that develops fragment-based QM methods (56-58) and the empirical force field community. The components of the energy that are recognized to be many-body in character include polarization (often referred to as induction), exchange-repulsion, dispersion, and charge transfer. Typical empirical force fields account for exchange-repulsion and dispersion, however under the assumption of pairwise additivity, as in the Lennard-Jones and Halgren potentials.

When enhancements to force fields that account for many-body effects are considered, polarization usually receives special attention, as it decays more slowly than dispersion, exchange-repulsion, or charge transfer with a $1/R^4$ dependence. The multipole description of the charge distribution fits naturally within the framework of polarizable models according to a Taylor expansion of the electrostatic energy U in the field $\bar{E}$:

$$U = E\left(\frac{\partial U}{\partial E}\right)_{E=0} + \frac{1}{2!}E^2\left(\frac{\partial^2 U}{\partial E^2}\right)_{E=0} + \frac{1}{3!}E^3\left(\frac{\partial^3 U}{\partial E^3}\right)_{E=0} + \ldots$$
$$= \mu E + \frac{1}{2!}\alpha E^2 + \frac{1}{3!}\beta E^3 \quad (3)$$

where μ is the dipole moment, α is the dipole polarizability, and β is the dipole hyperpolarizability (42). Therefore, major efforts have been undertaken throughout the empirical force field community to account for polarization. This began early with the seminal work of Warshel and Levitt in 1976 where they expound their Langevin dipole method (59). Beginning in the late 1980s and early 1990s, forerunners of modern treatments of polarization were introduced (60, 61). Since then, development of polarizable force fields has become an active area of research, in particular because it is known that correct treatment of the most dominant many-body intermolecular interaction, polarization, leads to improved descriptions of numerous phenomena, including phase transitions (62, 63), solvation free energies (64-67), behavior of solvent at hydrophilic and hydrophobic interfaces (15, 68), binding free energies(69-72), and intrinsic conformational preferences of organic compounds and DNA oligomers(73, 74). Polarizable models afford the ability to model physical phenomena across the phase diagram, whereas fixed charge models have difficulty simultaneously reproducing gas and condensed-phase properties (75).

There have emerged three main approaches to calculating polarization in empirical force fields: the fluctuating charge method(61, 76), adopted by CHARMM (77, 78) and OPLS (79, 80); the drude-oscillator method (81, 82), as adopted by CHARMM (83, 84) and GROMOS (85, 86); and the well-studied induced dipole method (87, 88), implemented in TINKER(47), SIBFA (48), NEMO (49), SDFF (89), AMBER (90), CHARMM (91), and OPLS (92). The fluctuating charge and Drude oscillator approaches have been excellently reviewed in detail elsewhere (40, 93), and are unique from the induced dipole model in that the former two approaches are essentially attempts to extend previous fixed, atom-centered charge models to accommodate polarization.

By contrast, the induced dipole model incorporates multipole moments beyond the point charge in a formalism where the natural link between the fixed off-atom charge and the polarizabilities is clear from the fact they are terms of a Taylor expansion of the energy in the

field as is evident from Eq. 2. This model allows for off-atom, through space polarization by means of inducible dipoles determined by the following equation:

$$\mu_{i,\beta} = \alpha_i \left(\sum_{\{j\}} T_{\beta}^{ij} M_j^{perm} + \sum_{\{j'\}} T_{\beta\delta}^{ij'} \mu_{j'\delta} \right) \qquad \beta,\delta = x,y,z \tag{4}$$

where M_j^{perm} is the permanent multipole moments, T is a matrix of derivatives of $1/R$, and α_i is a scalar, isotropic polarizability. In the AMOEBA model multipoles are represented trough quadrupoles and μ_i is solved for by iterating Eq. (4) to self-consistency using the conservative successive over relaxation (SOR) method, and the computational cost of tight convergence of the induced dipoles to 10^{-6} to 10^{-8} D is a factor of 15-30 that of a standard fixed charge model. Recently, the SOR method has been replaced by a more stable version of the pre-conditioned conjugate gradient SCF developed by Wang and Skeel(94) that reduces that factor to ~3.

Table 1 reports the AMOEBA solvation free energies of common small molecules found in biochemistry, including common amino acid side chain analogues, with corresponding statistical uncertainties. When compared to the experimental results, the RMS error for AMOEBA solvation free energies is 0.68 kcal/mol, with a mean signed error of +0.14 kcal/mol(95), which is a qualitative improvement over traditional fixed charge force fields (Figure 1), and without needing to parameterize against solvation free energy data.

3. RECENT ADVANCES IN POISSON BOLTZMANN ELECTROSTATICS

On the other extreme of modeling, questions involving supramolecular assemblies and long timescales are beyond the reach of current all atom MD methods. An excellent illustration is the cellular scale dynamical simulation of *E. coli* cytoplasm with ~ 1000 macromolecules modeled with PB electrostatics and simulated for 20 μs using Brownian dynamics (96). To extend dynamical simulations to such regimes requires judicious reduction of the computational complexity per timestep on several fronts.

First the inherent multiscale nature of macromolecular interactions is generally comprised of an initial long timescale diffusional search dominated by long-range electrostatics, followed by a docking phase characterized by more complicated short-range forces. Furthermore since we are principally interested in the behavior of the macromolecular solutes, explicit solvent

molecules, and ions, can be replaced with an implicit solvation model, since these smaller molecules relax to their equilibrium positions and momenta faster than the macromolecular solutes. This mean-field treatment is comprised of both a non-polar contribution to the solvation energy (97) (98), as well as a polar contribution that is modeled using continuum electrostatics. In this case the solvent is modeled as a uniform dielectric constant ε_{out} which typically differs substantially from the low dielectric envelope of the macromolecules (ε_{in}) within which are embedded complex charge distributions. The contributions of salt or ions are modeled as a continuous charge density profile outside the macromolecular cavities, which can be determined by solving the Poisson-Boltzmann equation.

The classical DLVO theory (99, 100) was originally developed to describe the interaction between charge-stabilized colloidal particles via a pair potential that includes repulsive screened-Coulomb interactions characterized by the Debye length λ_D. Despite its success at long-range, the DLVO approximation breaks down when λ_D is much greater than the average distance between particles or low salt concentrations. This corresponds to the scenario where the ion clouds of particles overlap with each other and many-body effects come into play. The Poisson-Boltzmann (PB) treatment for electrostatics provides a powerful coarse-grained model to account for this missing mutual polarization effect. In this review we primarily focus on the linearized form of the PB equation which gives the potential Φ at any point $\mathbf{r}$ in space $\Re^3$ as

$$-\nabla \times \left(\varepsilon(r) \nabla \Phi(r) \right) + \kappa^2 \Phi(r) = 4\pi \rho_{fixed}(r) \tag{5}$$

where ε is the relative dielectric function, ρ_{fixed} is the charge density due to the fixed macromolecular partial charges, and $\kappa = \sqrt{8\pi \bar{n} e^2 / \varepsilon k_B T}$ is the inverse of the Debye length λ_D, e is the fundamental electronic charge, k_B the Boltzmann constant, and T the absolute temperature.

Traditionally, PB electrostatics have been limited to rigid-body solute descriptions(101-103), because introducing internal degrees of freedom to the solute necessitates a smaller timestep and calls to PB solvers are computationally expensive. In most cases, MD using PB rigid bodies are based on finite-difference or boundary element solvers, and various numerical limitations to high quality description of the complex polarization fields have curtailed the dynamical simulations to one macromolecule (104-108) or just a few (109). Consequently MD using implicit solvent has typically relied on the faster (but less accurate) generalized Born

model. However, new PB and LPB equation solvers are extending the ability of these methods to describe mutual polarization accurately for larger and more complex systems.

3.1. Treatment of Mutual Polarization for Poisson-Boltzmann Solvers

The primary bottleneck for rapid evaluation of the PB equation is the need to fully converge the mutual polarization. Solutes immersed in a high-dielectric, salty continuum experience two sources of polarization. Self-polarization refers to a solute's combined response to the dielectric discontinuity across its surface and the salt exclusion from its interior. Mutual polarization refers to the response of a solute to the presence of other charges from other solutes in the system. Accounting for mutual polarization is a many-body problem that must be iteratively solved until self-consistency is achieved or via direct matrix calculation. While mutual polarization is negligible for well separated cavities, it becomes dominant at smaller separation distances(110). Lotan and Head-Gordon showed using a system of barstar protein molecules that forces and torques computed with and without mutual polarization essentially agree when molecules are separated beyond 40Å, but the exponential rise in the importance of mutual polarization below this separation results in differences that rise to more than 80% at 2 Å separation (111). However, most Brownian dynamics studies driven by PB electrostatics use an 'effective-charge' approximation, in which self-polarization charges are added to the macromolecule at the start of the simulation, but no mutual polarization is computed during the simulation (101). The cellular scale *E. coli* simulation uses such an approximation (96).

3.2 Recent Advances in Poisson-Boltzmann Solvers

Solution of the PB equation can be broadly categorized into analytical and numerical approaches. Analytical methods typically allow rapid solution of the PB theory under specialized geometries such as spheres or cylinders. The complete LPB solution for one spherical macromolecule was developed by Kirkwood (112) more than 80 years ago. Generalization of this solution to two or more spherical macromolecules proved to be more difficult, and many different partial and approximate solutions were proposed (113-116). Lotan and Head-Gordon derived a complete analytical LPB solution for computing the screened electrostatic interaction between arbitrary numbers of spheres of arbitrarily complex charge distributions, separated by arbitrary distance (111). This was accomplished by exploiting multipole expansion theory for the screened Coulomb potential by Greengard and co-workers (117), in order to describe screened charge-

charge interactions and all significant higher-order cavity polarization effects between low dielectric spherical cavities containing their charges, while treating mutual polarization correctly at all separation distances.

The full derivation is too complex to describe here, and the interested reviewer is referred to (111). Here we give the final result of the system of linear equations that describes the fully analytical LPB solution for simple spherical geometries:

$$H = \Gamma \times (C \times T \times H + F) \tag{6}$$

where the vectors H represent the *effective* multipole expansion of the charge distributions of each molecule, for a given configuration of N macromolecules (the matrices T). Thus Eq. (6) can be understood intuitively: the external potential field induced by a molecule is the sum of the contribution of its free charges, *F*, and the contribution of polarization charges induced by other molecules in a salty environment, $C \times T \times H$ (where C is the cavity polarization operator and T is the operator that converts the multipole expansion to a local Taylor expansion), transformed by the effect of its dielectric boundary, Γ. This solution allows for the study of full mutual polarization and provides the first complete benchmark for numerical solutions of the LPBE.

Numerical methods to solve for more realistic dielectric boundaries can be based on finite difference (FD) (118-120), finite-element (FE)(118, 121), or boundary element (BE) (122-126) approaches. FD and FE methods require that the solution be solved on a grid, limiting their practical application to spatial domains of either two to three typical macromolecules at reasonably high resolution (~0.2Å), or to larger numbers of macromolecules with greatly diminished resolution and thus solution accuracy (127). In contrast, BE methods by construction focus on solutions only on the macromolecular surfaces, thus removing spatial resolution limitations imposed by the 3D grid of the FD or FE solutions, making them more applicable to large-scale, multi-molecular dynamic simulations. Other methods such as stochastic ‘walk on sphere’ algorithms (128) are capable of computing the electrostatic energies, but as yet formulation for forces, needed for dynamics, is not available. Interface modules supporting PB electrostatics to run APBS (129) have been implemented in several widely used MD packages such as AMBER (107, 108), CHARMM(108, 130), and NAMD (131).

For one-time static computations, PB-solvers can afford to fully account for mutual polarization, although error control differs between the numerical methods. FD and FE methods by construction compute the solution globally at each iteration step; although focusing

techniques (132, 133) can be used to refine the solution near points of interest (e.g. macromolecular surfaces), the mutual polarization responses must still propagate back and forth through space from one molecule to another. BE methods, in contrast, solve the surface polarization charge on each macromolecule directly and allows a more physically intuitive control of mutual polarization. For instance, Bordner and Huber (124) approximate the mutual polarization effects of molecule 2 on molecule 1 via a one-step computation that solves the BE-based PB equation of the dielectric cavity of molecule 1 (i.e. without partial charges) in the presence of potential due to molecule 2.

Recently, Yap and Head-Gordon developed a semi-analytical PB method (PB-SAM) using BE that accounts for complete mutual polarization(127, 134). This new numerical approach represents the macromolecular surface as a collection of spheres in which the surface charges can then be iteratively solved by the analytical multipole methods previously introduced by Lotan & Head-Gordon(111). The PB-SAM solution has been recently extended to calculate forces and torques to simulate Brownian dynamics(135). The strength of the PB-SAM method was illustrated by the study of the association kinetics for a system of 125 barnase / barstar molecules- a classic example of an electrostatically steered diffusion-limited association – but now under conditions of high concentration relevant to so-called “crowding” conditions. In particular this work showed that it is possible to perform a LPB dynamics simulation under periodic boundary conditions by calculating the mean first passage time for a barnase-barstar docking event(135). The two order of magnitude increase in the number of macromolecules that can be treated with full mutual polarization in a MD scheme should allow for systematic study of larger mesoscale systems where electrostatics dominate(135). In summary, with current advancements in PB solvers, PB-based MD with far greater complexity may be within reach (111, 127, 136, 137).

4. New Results and Future Directions

Mutually polarizable models offer a significant improvement in the physics of classical atomistic and coarse-grained force fields as described in Sections 2 and 3. The sidebar provides recent developments of mutually polarizable QM/MM schemes for applications that can now model chemical reactivity, excited states, and non-adiabatic interactions. Current challenges in treatments of polarizability include enhancing computational efficiency and improving the ability

of these models to reproduce physical observables, in order to deliver on a better model for a new range of systems accessible to molecular and coarse-grained simulations. In this final section we consider systematic approximations to a full mutually polarizable model, atomistic or coarse-grained, that may define a sweet spot for accuracy and tractability for condensed phase simulations in the future.

The many-body expansion (MBE) to the total potential energy of an N-body system

$$\Delta U = \Delta U_1 + \Delta U_2 + \Delta U_3 ... \tag{7a}$$

where

$$\Delta U_1 = \sum_{i=1}^{N} U(i) \qquad \Delta U_2 = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} U(i,j) - U(i) - U(j)$$

$$\Delta U_3 = \sum_{i=1}^{N-2} \sum_{j=i+1}^{N-1} \sum_{k=j+1}^{N} U(i,j,k) - U(i,j) - U(i,k) - U(j,k) + U(i) + U(j) + U(k) \tag{7b}$$

states that the potential energy can be evaluated as the set of atomic or molecular interactions for increasing cluster size starting from monomers, progressing to dimers, trimers, etc. The MBE has been known since the 1970's (138, 139), and it remains a current topic of interest for those developing QM-based fragment approaches and embedding schemes (56-58, 140). Here the question is whether truncation of an N-body mutual polarization model such as AMOEBA or PB-SAM to lower order is sufficient for both accuracy and cost savings. As discussed above, truncation of Eq. (7) at the level of two-body interactions is exact for a fixed charge model potential since the point charge or multipole electrostatics are by definition pairwise additive.

It is at the next level of approximation that the analysis becomes more interesting for classical electrostatic models. One such approximation is a direct polarization scheme whereby the inducible dipoles respond only to the field due to permanent multipoles, eliminating mutual polarization. This entails omitting the second term on the right-hand side of Eq. 4. Inspection of the expression for the energy due to interacting induced dipoles in this regime shows that this is effectively a three-body potential (42). Written in tensor notation, the expression for the energy is

$$U = \tfrac{1}{2} \underline{\mu}^T \underline{\underline{T}} \underline{\mu} \tag{8}$$

where single and double underscores denote rank one and rank two tensors, respectively. When the first term of Eq. 4 is substituted in Eq. 9, it is clear that energy and force terms that arise depend on 2- and 3-body direct polarization, with no contributions from higher order terms of

Eq. (7). The 2-body direct polarization energy is the interaction of a polarizable site with the electric field of the permanent moments of another site. The 3-body direct polarization energy arises from the interaction of a polarizable site with the *interaction* electric field of the permanent moments of two sites. Consequently, it is easy to show that the direct polarization energy of an N water molecule system is equivalent to the direct polarization of the many-bodied expansion truncated at the level of triplets in Eq. (7).

However, ~20% of the polarization energy is lost when neglecting the mutual polarization that must be recovered through optimization of the model parameters. Fulfillment of accuracy for this direct polarization scheme has been recently realized through a new classical water model that is a direct polarization version of the fully polarizable AMOEBA water model. Wang and co-workers used an optimization method called ForceBalance (141, 142) for parameterizing the iAMOEBA model using a combination of experimental data and high-level QM calculations. The demonstrated accuracy of the iAMOEBA model was validated against a published, comprehensive benchmark of water properties developed by Vega and coworkers (143) that covers a wide range of phases and thermodynamic conditions going far beyond the parameterization data set. Table 2 provides a range of properties of the iAMOEBA model in comparison to other fully mutual polarizable models and experiment, which shows that the iAMOEBA model performs exceptionally well.

It should be recognized that although iAMOEBA performs better on average than the mutually polarizable models, it is not because mutual polarization can always be neglected. Rather it is that the more automated ForceBalance scheme(142) is a better parameterization strategy than hand-tuning which is largely how most other polarizable (or even fixed charge) force fields have been derived. The limitation of the direct polarization functional form likely arises in the reproduction of *ab initio* benchmark energies of large water clusters, in which iAMOEBA systematically underestimates the binding energies(144). Furthermore, while iAMOEBA correctly predicts a higher dielectric constant in ice Ih compared to the liquid, the dielectric constant ratio of ice to the liquid is lower than the experimental value, another indicator of the importance of full polarization(144). The incomplete balance of description of properties across gas phase water cluster data, the bulk liquid, and ice, is evidence that the direct polarization model can’t fully capture the significant variation in electric fields. Thus

development of mutual polarization models using automated parameter approaches such ForceBalance(142) is an attractive future direction.

The excellent ability of iAMOEBA in reproducing a range of bulk water data validates our notion that we can approximate the mutual induction while achieving good agreement with condensed-phase observables. Therefore we would also like to attempt a regime where we preserve some mutual polarization while potentially saving computational cost. A model that recovers successfully higher levels of mutual polarization would involve truncation of Eq. (7) for full mutually polarizable dimers, trimers, or perhaps tetramers (145, 146). Here we show that truncation of Eq. (7) at trimers can be a very good approximation to the N-body polarizable potential given the rapid spatial decay of mutual polarization.

Table 3 compares the potential energy calculated with AMOEBA (our reference state for N-body polarization) and as a function of truncations at 1-, 2- and 3-body mutual polarization for a variety of water clusters ranging in size from 17 to 1000 water molecules and evaluated under the Ewald sum. We also consider a small water cluster with an embedded sulfate anion, as well as solutions of sulfate with no net charge and a sulfate with no permanent electrostatics to mimic polar and hydrophobic species, respectively. Several important observations from Table 3 are noteworthy. The most important of these is that the truncation error of Eq. (7) at the level of 3-body terms can be a small fraction of the total energy with errors below 1% in all cases except for the net charged sulfate water cluster. This opens up the possibility of a cheaper and reasonably accurate polarization model that refactors the N-body algorithm for calculating mutual polarization to one that is highly parallelizable over independent trimer calculations. Secondly, in accord with established observations of the magnitude of many-body effects, the dominant intermolecular term is the 2-body term, which provides the justification for subsuming higher order many-body in popular pairwise additive force fields. Lastly, while the 3-body term is negative in most cases, it is positive in the net charged system, which is interesting in that one observes that polarization can give rise to cooperativity as well as *anti-cooperativity.*

A fortuitous effect of our reduction of computational cost is that we may consider enhancing the 3-body polarization model in a number of ways. For example, the 3-body potential may also include additional many-body effects in the calculation, such as charge transfer, changes to the Thole damping model to allow for some account of charge penetration, coupling of polarization and exchange-repulsion, anisotropic polarizabilities, coupling of polarization with

conformational change, and other enhancements that heretofore were computationally infeasible. The 3-AMOEBA model will necessitate reparameterization, and the ForceBalance suite of programs used in the iAMOEBA reparameterization will also benefit 3-AMOEBA polarization scheme.

The direct polarization model or truncation of the MBE to approximate N-body polarization at lower order are completely generalizable to the LPB and PB-SAM models described in Section 3. To show the validity of these approximate polarizable coarse-grained models, we have calculated the total interaction energy of four spheres of radius 21.8 Å with a net charge of -5e placed at the sphere centers, that are separated by a distance *a*. Since the many-body interaction depends on the interplay between *a* and λ_D, we also tested three monovalent salt concentrations of 0.001 M, 0.05 M and 0.1 M, which correspond to λ_D = 95.997, 13.576 and 9.600 Å, respectively. The direct polarization model means that the surface charge density of a given molecule is not polarized by another molecule; the 2-body and 3-body approximation is based on the sum over the mutual polarization interactions for N(N-1)/2! dimers and N(N-1)(N-2)/3! trimers, respectively.

In Figure 2 the interaction energies scaled by k_bT are compared between the direct polarization, 2-body, 3-body approximations and the full mutual polarization PB-SAM model as a function of *a* and λ_D. In all cases the energy obtained from the 3-body approximation is in excellent agreement with that from full mutual polarization, and the 2-body approximation is also good for the smaller Debye lengths. This is a natural result since more dilute salt solutions have larger Debye lengths, and at large separations the ion cloud around each molecule can still overlap, thus making 3-body terms of the MBE still relevant. In summary, the PB solvers outlined in Section 3 can also benefit from the trivial parallelization afforded by the MBE analysis presented here, thereby opening up an increased accuracy and size range of coarse-grained calculations.

5. CONCLUSIONS

Next generation computational chemistry will be strongly enabled by the deployment of new state of the art theoretical models that include improved pairwise additive potentials, many-body force fields and new algorithms that increase their efficiency and practicality to perform sampling via MD. The increasing complexity of the parameter search problem in development of

new force-fields is also exposing the limitations of hand-tuning parameters which is often a manual and laborious task. The benefits of a more automated process such as ForceBalance may aid in the development of a hierarchy of atomistic theoretical models and PBE solvers that can work seamlessly together.

We have presented a hierarchy of classical force fields that alter the trade off between accuracy and computational speed to define a “sweet spot” for a given scientific application. We believe that fixed charge force fields will benefit from greater integration with more recently developed water models (the most ubiquitous of solvent environment), and more complex functional forms such as multipoles that describe short-ranged anisotropic interactions such as hydrogen-bonding. Direct polarization models and truncation of the N-body interactions to lower 3-body interactions offer well-defined approximations to mutual polarization that are valid for both atomistic and coarse-grained electrostatic models. Finally full mutual polarization interaction will be needed when trying to cover a wide range of environments within one unified model. Here we have highlighted the new accuracy possible with polarizable force fields that is especially needed in computational research that makes connection with actual experimental observables. In summary, a better understanding of which statistical properties of condensed phase chemical systems actually require an advancement over a classical pairwise-additive force field helps give formal understanding about the importance of more complex mutual polarization interactions and a practical guide to the proper application of theoretical models.

ACKNOWLEDGEMENTS. We thank the National Science Foundation grant CHE-1265731 for support of this work. We also acknowledge computational resources obtained under NSF grant CHE-1048789. We thank Dr. Shule Liu for providing the data embodied in Figure 2. We would like to thank Paul Nerenberg, Jay Ponder, Pengyu Ren, Dave Case, Martin Head-Gordon, Julia Rice, Bill Swope, Lee-Ping Wang, and Vijay Pande for helpful discussions and enjoyable collaborations.

SIDEBAR HIGHLIGHT

Since the pioneering work of Warshel and Levitt (59), hybrid QM/MM studies have grown in popularity and sophistication. In addition to adequate treatment of the QM region with the appropriate level of theory that is inherent in all *ab initio* calculations, QM/MM calculations also require methods for modeling the interaction of the MM region with the QM region. In most

modern electrostatically embedded approaches, the electrostatic interaction of the MM region with the QM region is taken into account by explicitly including the fixed partial charges of the MM region in the QM Hamiltonian, effectively permitting polarization of the charge density in the QM region by the MM region. While this represents an advancement over mechanical embedding methods where the QM and MM regions are completely independent, there remains the need to account for polarization of the MM region by the QM region. To address this, a number of groups have incorporated classical polarization methods in the MM region that respond to the field generated by the MM permanent charges and the QM charge density. These include the inducible dipole model, as recently implemented in the polarizable embedding method (147), the fluctuating charge model (148), and the Drude oscillator model (149). Due to the computational cost of iteration to self-consistency, these approaches have been limited to small systems, with the exception of a QM/MM calculation on chlorophyll performed with the relatively inexpensive semiempirical INDO method (150). Low order approximations of the many-body polarization afford an attractive solution to reduce the computational cost.

6. REFERENCES

1. Ren PY, Ponder JW. 2003. Polarizable atomic multipole water model for molecular mechanics simulation. *J. Phys. Chem. B* 107: 5933-47
2. Allinger NL, Yuh YH, Lii JH. 1989. Molecular Mechanics - the Mm3 Force-Field for Hydrocarbons .1. *J. Amer. Chem. Soc.* 111: 8551-66
3. Wilson EB, Decius JC, Cross PC. 1955. *Molecular Vibrations*. New York, New York: McGraw Hill
4. Stone AJ. 1996. *The Theory of Intermolecular Forces*. Oxford,UK: Clarendon Press
5. Halgren TA. 1992. Representation of van der Waals (vdW) interactions in molecular mechanics force fields: potential form, combination rules, and vdW parameters. *J. Amer. Chem. Soc.* 114: 7827-43
6. Ren P, Ponder J. 2002. Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. *J. Comp. Chem.* 23: 1497-506
7. Stone AJ. 1981. Distributed multipole analysis, or how to describe a molecular charge distribution. *Chem. Phys. Lett.* 83: 233-9
8. Wen S, Beran G. 2011. Accurate molecular crystal lattice energies from a fragment QM/MM approach with on-the-fly ab initio force-field parameterization. *J. Chem. Theory Comput.* 7: 3733-42
9. Cieplak P, Lybrand T, Kollman P. 1987. Calculatio of free energy changes in ion-water clusters using non-additive potentials and Monte Carlo methods. *J. Chem. Phys.* 86: 6393-404
10. Lifson S, Warshel A. 1969. Consistent force field for calculations of conformations, vibrational spectra, and enthalpies of cycloalkane and n-alkane molecules. *J. Chem. Phys.* 49: 5116

The Consistent Force Field was the first empirical force field . Most later force fields adopt the basic functional forms introduced here.

11. Hagler A, Huler E, Lifson S. 1974. Energy functions for peptides and proteins. I. Derivation of a consistent force field including the hydrogen bond from amide crystals. *J. Amer. Chem. Soc.* 96: 5319-27
12. Hagler A, Lifson S. 1974. Energy functions for peptides and proteins. II. Amid hydrogen bond and calculation of amide crystal properties. *J. Amer. Chem. Soc.* 96: 5327-35
13. Shirts M, Pande V. 2005. Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. *J. Chem. Phys.* 122: 134508
14. Mobley D, Bayly C, Cooper M, Shirts M, Dill K. 2009. Small molecule hydration free energies in explicit solvent: An extensive test of fixed-charge atomistic simulations. *J. Chem. Theo. Comp.* 5: 350-8
15. Johnson M, Malardier-Jugroot C, Murarka R, Head-Gordon T. 2009. Hydration water dynamics near biological interfaces. *J. Phys. Chem. B* 113: 4082-92
16. Jiang F, Han W, Wu Y. 2010. Influence of side chain conformations on local conformational features of amino acids and implication for force field development. *J. Phys. Chem. B* 114: 5840-50
17. Best R, Buchete N, Hummer G. 2008. Are current molecular dynamics force fields too helical? *Biophys. J.* 95: L7-L9
18. Best R, Hummer G. 2009. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J. Phys. Chem. B* 113: 9004-15
19. Jorgensen WL. 1981. Quantum and Statistical Mechanical Studies of Liquids .10. Transferable Intermolecular Potential Functions for Water, Alcohols, and Ethers - Application to Liquid Water. *J. Amer. Chem. Soc.* 103: 335-40
20. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 79: 926-35
21. Berendsen HJC, Grigera JR, Straatsma TP. 1987. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* 91: 6269-71
22. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. 1981. In *Intermolecular Forces*, ed. B Pullmann, pp. 331-42. Dordrecht,Netherlands: D. Reidel Publishing Company
23. Horn H, Swope W, Pitera J, Madura J, Dick T, et al. 2004. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* 120: 9665
24. Abascal JLF, Vega C. 2005. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* 123: 234505.
25. Mahoney MW, Jorgensen WL. 2000. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* 112: 8910-22
26. Price DJ, Brooks CL. 2004. A modified TIP3P water potential for simulation with Ewald summation. *J. Chem. Phys.* 121: 10096-103
27. Fawzi NL, Phillips AH, Ruscio JZ, Doucleff M, Wemmer DE, Head-Gordon T. 2008. Structure and dynamics of the Abeta(21-30) peptide from the interplay of NMR experiments and molecular simulations. *J. Am. Chem. Soc.* 130: 6145-58
28. Sgourakis NG, Merced-Serrano M, Boutsidis C, Drineas P, Du ZM, et al. 2011. Atomic-Level Characterization of the Ensemble of the A beta(1-42) Monomer in Water Using

Unbiased Molecular Dynamics Simulations and Spectral Algorithms. *J. Mol. Bio.* 405: 570-83
29. Ball KA, Phillips AH, Nerenberg PS, Fawzi NL, Wemmer DE, Head-Gordon T. 2011. Homogeneous and Heterogeneous Tertiary Structure Ensembles of Amyloid-beta Peptides. *Biochemistry* 50: 7612-28
30. Ball KA, Phillips AH, Wemmer DE, Head-Gordon T. 2013. Differences in β-strand populations of monomeric amyloid-β 40 and amyloid-β 42. *Biophys. J.*104: 2714-2724.
31. Wickstrom L, Okur A, Simmerling C. 2009. Evaluating the Performance of the ff99SB Force Field Based on NMR Scalar Coupling Data. *Biophys. J.* 97: 853-6
32. Krouskop PE, Madura JD, Paschek D, Krukau A. 2006. Solubility of simple, nonpolar compounds in TIP4P-Ew. *J. Chem. Phys.* 124:016102.
33. Best RB, Mittal J. 2010. Protein Simulations with an Optimized Water Model: Cooperative Helix Formation and Temperature-Induced Unfolded State Collapse. *J. Phys. Chem. B* 114: 14916-23
34. Nerenberg PS, Head-Gordon T. 2011. Optimizing Protein-Solvent Force Fields to Reproduce Intrinsic Conformational Preferences of Model Peptides. *J. Chem. Theo. Comp.* 7: 1220-30
35. Nerenberg P, Jo B, So C, Tripathy A, Head-Gordon T. 2012. Optimizing solute-water van der Waals interactions to reproduce solvation free energies. *J. Phys. Chem. B* 116: 4524-34
36. Horta BAC, Fuchs PFJ, van Gunsteren WF, Huenenberger PH. 2011. New interaction parameters for oxygen compounds in the GROMOS force field: Improved pure-liquid and solvation properties for alcohols, ethers, aldehydes, ketones, carboxylic acids, and esters. *J. Chem. Theo. Comp.* 7: 1016-31
37. Bayly C, Cieplak P, Cornell W, Kollman P. 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges--The RESP model. *J. Phys. Chem.* 97: 10269-80
38. Breneman C, Wiberg K. 1990. Determining atom-centered monopoles from molecular electrostatic potentials: The need for high sampling density in formamide conformational analysis. *J. Comp. Chem.* 11: 361-73
39. Ren P, Chun J, Thomas DG, Schnieders MJ, Marucho M, et al. 2012. Biomolecular electrostatics and solvation: a computational perspective. *Q Rev Biophys* 45: 427-91
40. Yu H, van Gunsteren W. 2005. Accounting for polarization in molecular simulation. *Comp. Phys. Comm.* 172: 69-85
41. Williams D. 1988. Representation of the molecular electrostatic potential by atomic multipole and bond dipole models. *J. Comp. Chem.* 9: 745-63
42. Dykstra CE. 1993. Electrostatic Interaction Potentials in Molecular-Force Fields. *Chem. Rev.* 93: 2339-53

A scholarly introduction to polarizable force fields.*

43. Fowler P, Buckingham A. 1983. The long range model of intermolecular forces *Mol. Phys.* 50: 1349-61
44. Fowler P, Buckingham A. 1991. Central or distributed multipole moments? Electrostatic models of aromatic dimers. *Chem. Phys. Lett.* 176: 11-8
45. Schnieders MJ, Fenn TD, Pande VS, Brunger AT. 2009. Polarizable atomic multipole X-ray refinement: application to peptide crystals. *Acta Cryst. Sec. D* 65: 952-65

46. Afonine PV, Grosse-Kunstleve RW, Adams PD, Lunin VY, Urzhumtsev A. 2007. On macromolecular refinement at subatomic resolution with interatomic scatterers. *Acta Cryst. Sec. D* 63: 1194-7
47. Ponder JW, Wu CJ, Ren PY, Pande VS, Chodera JD, et al. 2010. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* 114: 2549-64
48. Piquemal JP, Williams-Hubbard B, Fey N, Deeth RJ, Gresh N, Giessner-Prettre C. 2003. Inclusion of the ligand field contribution in a polarizable molecular mechanics: SIBFA-LF. *J. Comp. Chem.* 24: 1963-70
49. Engkvist O, Åstrand P-O, Karlström G. 2000. Accurate intermolecular potentials obtained from molecular wave funcions: Bridging the gap between quantum chemistry and molecular simulations. *Chem. Rev.* 100: 4087-108
50. Cieplak P, Caldwell J, Kollman P. 2001. Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation: Aqueous solution free energies of methanol and N-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform/water partition coefficients of the nucleic acid bases. *J. Comp. Chem.* 22: 1048-57
51. Donchev AG, Ozrin VD, Subbotin MV, Tarasov OV, Tarasov VI. 2005. A quantum mechanical polarizable force field for biomolecular interactions. *Proc. Natl. Acad. Sci. USA* 102: 7829-34
52. Grishaev A, Bax A. 2004. An empirical backbone-backbone hydrogen-bonding potential in proteins and its applications to NMR structure refinement and validation. *J. Amer. Chem. Soc.* 126: 7281-92
53. Kortemme T, Morozov AV, Baker D. 2003. An orientation-dependent hydrogen bonding potential improves prediction of specificity and structure for proteins and protein-protein complexes. *J. Mol. Bio.* 326: 1239-59
54. Axilrod B, Teller E. 1943. Interaction of the van der Waals type between three atoms. *J. Chem. Phys.* 11: 299-300
Early model of 3-body dispersion.
55. Muto Y. 1943. The force between nonpolar molecules. *J. Phys. Math. Soc. Japan* 17: 629-31
56. Gordon M, Fedorov D, Pruitt S, Slipchenko L. 2011. Fragmentation methods: A route to accurate calculations on large systems. *Chem. Rev.* 112:632-672
57. Leverentz H, Maerzke K, Keasler S, Siepmann J, Truhlar D. 2012. Electrostatically embedded many-body method for dipole moments, partial atomic charges, and charge transfer. *Phys. Chem. Chem. Phys.* 14: 7669-78
58. Richard R, Herbert J. 2012. A generalized many-body expansion and a unified view of fragment-based methods in electronic structure theory. *J. Chem. Phys.* 137:064113
59. Warshel A, Levitt M. 1976. Theoretical studies of enzymatic reaction: Dielectric, electrostatic, and steric stabilization of carbonium-ion in reaction of lysozyme. *J. Mol. Bio.* 103: 227-49
The first QM/MM study of an enzyme.
60. Sprik M, Klein M. 1988. A polarizable model for water using distributed charge sites. *J. Chem. Phys.* 89: 7556-60
61. Rick S, Stuart S, Berne B. 1994. Dynamical fluctuating charge force-fields: Application to liquid water. *J. Chem. Phys.* 101: 6141-56
The FLUCU model is introduced.

62. Patel S, Brooks CL. 2006. Revisiting the hexane-water interface via molecular dynamics simulations using nonadditive alkane-water potentials. *J. Chem. Phys.* 124: 204706
63. Bauer BA, Patel S. 2009. Properties of water along the liquid-vapor coexistence curve via molecular dynamics simulations using the polarizable TIP4P-QDP-LJ water model. *J. Chem. Phys.* 131
64. Grossfield A, Ren PY, Ponder JW. 2003. Ion solvation thermodynamics from simulation with a polarizable force field. *J. Amer. Chem. Soc.* 125: 15671-82
65. Jiao D, King C, Grossfield A, Darden TA, Ren P. 2006. Simulation of Ca2+ and Mg2+ solvation using polarizable atomic multipole potential. *J. Phys. Chem. B* 110: 18553-9
66. Yu HB, Whitfield TW, Harder E, Lamoureux G, Vorobyov I, et al. 2010. Simulating Monovalent and Divalent Ions in Aqueous Solution Using a Drude Polarizable Force Field. *J. Chem. Theo. Comp.* 6: 774-86
67. Shi Y, Wu CJ, Ponder JW, Ren PY. 2011. Multipole Electrostatics in Hydration Free Energy Calculations. *J. Comp. Chem.* 32: 967-77
68. Johnson ME, Malardier-Jugroot C, Head-Gordon T. 2010. Effects of co-solvents on peptide hydration water structure and dynamics. *Phys. Chem. Chem. Phys.* 12: 393-405
69. Jiao D, Golubkov PA, Darden TA, Ren P. 2008. Calculation of protein-ligand binding free energy by using a polarizable potential. *Proc. Natl. Acad. Sci. USA* 105: 6290-5
70. Khoruzhii O, Donchev AG, Galkin N, Illarionov A, Olevanov M, et al. 2008. Application of a polarizable force field to calculations of relative protein-ligand binding affinities. *Proc. Natl. Acad. Sci. USA* 105: 10378-83
71. Patel S, Davis JE, Bauer BA. 2009. Exploring Ion Permeation Energetics in Gramicidin A Using Polarizable Charge Equilibration Force Fields. *J. Amer. Chem. Soc.* 131: 13890-+
72. Wang JM, Cieplak P, Li J, Wang J, Cai Q, et al. 2011. Development of Polarizable Models for Molecular Mechanical Calculations II: Induced Dipole Models Significantly Improve Accuracy of Intermolecular Interaction Energies. *J. Phys. Chem. B* 115: 3100-11
73. Babin V, Baucom J, Darden TA, Sagui C. 2006. Molecular dynamics simulations of DNA with polarizable force fields: Convergence of an ideal B-DNA structure to the crystallographic structure. *J. Phys. Chem. B* 110: 11571-81
74. Rasmussen TD, Ren PY, Ponder JW, Jensen F. 2007. Force field modeling of conformational energies: Importance of multipole moments and intramolecular polarization. *Int. J. Quant. Chem.* 107: 1390-5
75. Giese TJ, York DM. 2004. Many-body force field models based solely on pairwise Coulomb screening do not simultaneously reproduce gas-phase and condensed-phase polarizability limits. *J. Chem. Phys.* 120: 9903-6
76. Rick SW, Stuart SJ, Bader JS, Berne BJ. 1995. Fluctuating Charge Force-Fields for Aqueous-Solutions. *J. Mol. Liq.* 65-6: 31-40
77. Patel S, Mackerell AD, Brooks CL. 2004. CHARMM fluctuating charge force field for proteins: II - Protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model. *J. Comp. Chem.* 25: 1504-14
78. Patel S, Brooks CL. 2004. CHARMM fluctuating charge force field for proteins: I parameterization and application to bulk organic liquid simulations. *J. Comp. Chem.* 25: 1-15
79. Kaminski GA, Stern HA, Berne BJ, Friesner RA, Cao YXX, et al. 2002. Development of a polarizable force field for proteins via ab initio quantum chemistry: First generation model and gas phase tests. *J. Comp. Chem.* 23: 1515-31

80. Stern HA, Kaminski GA, Banks JL, Zhou RH, Berne BJ, Friesner RA. 1999. Fluctuating charge, polarizable dipole, and combined models: Parameterization from ab initio quantum chemistry. *J. Phys. Chem. B* 103: 4730-7
81. Jacucci G, Mcdonald IR, Singer K. 1974. Introduction of Shell-Model of Ionic Polarizability into Molecular-Dynamics Calculations. *Phys. Lett. A* A50: 141-3
82. Mitchell PJ, Fincham D. 1993. Shell-Model Simulations by Adiabatic Dynamics. *J. Phys. Cond. Matt.* 5: 1031-8
83. Lamoureux G, MacKerell AD, Roux B. 2003. A simple polarizable model of water based on classical Drude oscillators. *J. Chem. Phys.* 119: 5185-97
84. Lamoureux G, Roux B. 2003. Modeling induced polarization with classical Drude oscillators: Theory and molecular dynamics simulation algorithm. *J. Chem. Phys.* 119: 3025-39
85. Geerke DP, van Gunsteren WF. 2007. On the calculation of atomic forces in classical simulation using the charge-on-spring method to explicitly treat electronic polarization. *J. Chem. Theo. Comp.* 3: 2128-37
86. Yu HB, Hansson T, van Gunsteren WF. 2003. Development of a simple, self-consistent polarizable model for liquid water. *J. Chem. Phys.* 118: 221-34
87. Applequist J, Carl JR, Fung KK. 1972. Atom Dipole Interaction Model for Molecular Polarizability - Application to Polyatomic-Molecules and Determination of Atom Polarizabilities. *J. Amer. Chem. Soc.* 94: 2952-&
88. Thole BT. 1981. Molecular Polarizabilities Calculated with a Modified Dipole Interaction. *Chem. Phys.* 59: 341-50
89. Palmo K, Mannfors B, Mirkin NG, Krimm S. 2006. Inclusion of charge and polarizability fluxes provides needed physical accuracy in molecular mechanics force fields. *Chem. Phys. Lett.* 429: 628-32
90. Wang ZX, Zhang W, Wu C, Lei HX, Cieplak P, Duan Y. 2006. Strike a balance: Optimization of backbone torsion parameters of AMBER polarizable force field for simulations of proteins and peptides. *J. Comp. Chem.* 27: 781-90
91. Xie WS, Pu JZ, MacKerell AD, Gao JL. 2007. Development of a polarizable intermolecular potential function (PIPF) for liquid amides and alkanes. *J. Chem. Theo. Comp.* 3: 1878-89
92. Jorgensen WL, Jensen KP, Alexandrova AN. 2007. Polarization effects for hydrogen-bonded complexes of substituted phenols with water and chloride ion. *J. Chem. Theo. Comp.* 3: 1987-92
93. Cieplak P, Dupradeau FY, Duan Y, Wang JM. 2009. Polarization effects in molecular mechanical force fields. *J. Phys. Cond. Matt.* 21:333102.
94. Wang W, Skeel RD. 2005. Fast evaluation of polarizable forces. *J. Chem. Phys.* 123: 164107
95. Shirts MR, Pitera JW, Swope WC, Pande VS. 2003. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *J. Chem. Phys.* 119: 5740-61
96. McGuffee SR, Elcock AH. 2010. Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLoS Comp. Bio.* 6: e1000694
97. Hermann RB. 1972. Theory of hydrophobic bonding. II. Correlation of hydrocarbon solubility in water with solvent cavity surface area. *J. Phys. Chem.* 76: 2754-9

98. Barone V, Cossi M, Tomasi J. 1997. A new definition of cavities for the computation of solvation free energies by the polarizable continuum model. *J. Chem. Phys.* 107: 3210-21
99. Verwey EJW. 1945. Theory of the stability of lyophobic colloids. *Phillips Res. Rep.* 1: 33-49
100. Derjaguin B, Landau L. 1993. Theory of the Stability of Strongly Charged Lyophobic Sols and of the Adhesion of Strongly Charged-Particles in Solutions of Electrolytes. *Prog. Surf. Sci.* 43: 30-59
101. Gabdoulline RR, Wade RC. 2002. Biomolecular diffusional association. *Curr. Opin. Struct. Bio.* 12: 204-13
102. Northrup S, Allison S, McCammon A. 1984. Brownian dynamics simulation of diffusion-influenced bimolecular reactions. *J. Chem. Phys.* 80: 1517-24
103. Madura JD, Briggs JM, Wade RC, Davis ME, Luty BA, et al. 1995. Electrostatics and Diffusion of Molecules in Solution - Simulations with the University-of-Houston Brownian Dynamics Program. *Comp. Phys. Comm.* 91: 57-95
104. Lu BZ, Wang CX, Chen WZ, Wan SZ, Shi YY. 2000. A stochastic dynamics simulation study associated with hydration force and friction memory effect. *J. Phys. Chem. B* 104: 6877-83
105. Lu BZ, Chen WZ, Wang CX, Xu XJ. 2002. Protein molecular dynamics with electrostatic force entirely determined by a single Poisson-Boltzmann calculation. *Proteins* 48: 497-504
106. Luo R, David L, Gilson MK. 2002. Accelerated Poisson-Boltzmann calculations for static and dynamic systems. *J. Comp. Chem.* 23: 1244-53
107. Lu Q, Luo R. 2003. A Poisson-Boltzmann dynamics method with nonperiodic boundary condition. *J. Chem. Phys.* 119: 11035-47
108. Prabhu NV, Zhu PJ, Sharp KA. 2004. Implementation and testing of a stable fast implicit solvation in molecular dynamics using the smooth-permittivity finite difference Poisson-Boltzmann method. *Biophys. J.* 86: 413A-A
109. Xin WD, Juffer AH. 2007. A boundary element formulation of protein electrostatics with explicit ions. *J. Comp. Phys.* 223: 416-35
110. Lu B, Cheng X, Huang J, McCammon JA. 2006. Order N algorithm for computation of electrostatic interactions in biomolecular systems. *Proc. Natl. Acad. Sci. USA* 103: 19314-9
111. Lotan I, Head-Gordon T. 2006. An Analytical Electrostatic Model for Salt Screened Interactions between Multiple Proteins. *J. Chem. Theo. Comp.* 2: 541-55
The analytical generalization of the Kirkwood PB solution for one molecule to multiple molecules and their mutual polarization.*
112. Kirkwood JG. 1934. Theory of Solutions of Molecules Containing Widely Separated Charges with Special Application to Zwitterions. *J. Chem. Phys.* 2
113. Phillies GD. 1974. Effects of Intermacromolecular Interactions on Diffusion .1. 2-Component Solutions. *J. Chem. Phys.* 60: 976-82
114. McClurg RB, Zukoski CF. 1998. The electrostatic interaction of rigid, globular proteins with arbitrary charge distributions. *J. Coll. Int. Sci.* 208: 529-42.
115. Sader JE, Lenhoff AM. 1998. Electrical double-layer interaction between heterogeneously charged colloidal particles: A superposition formulation. *J. Coll. Int. Sci.* 201: 233-43

116. Fenley AT, Gordon JC, Onufriev A. 2008. An analytical approach to computing biomolecular electrostatic potential. I. Derivation and analysis. *J. Chem. Phys.* 129
117. Greengard LF, Huang JF. 2002. A new version of the fast multipole method for screened Coulomb interactions in three dimensions. *J. Comp. Phys.* 180: 642-58
118. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* 98: 10037-41
119. Nicholls A, Honig B. 1991. A Rapid Finite-Difference Algorithm, Utilizing Successive over-Relaxation to Solve the Poisson-Boltzmann Equation. *J. Comp. Chem.* 12: 435-45
120. Yu S, Zhou Y, Wei G. 2007. Matched interface and boundary (MIB) method for elliptic problems with sharp-edged interfaces. *J. Comp. Phys.* 224: 729-56
121. Chen L, Holst MJ, Xu JC. 2007. The finite element approximation of the nonlinear Poisson-Boltzmann equation. *Siam J. Num. Anal.* 45: 2298-320
122. Zauhar RJ, Morgan RS. 1985. A New Method for Computing the Macromolecular Electric-Potential. *J. Mol. Bio.* 186: 815-20
123. Boschitsch AH, Fenley MO, Zhou HX. 2002. Fast boundary element method for the linear Poisson-Boltzmann equation. *J. Phys. Chem. B* 106: 2741-54
124. Bordner AJ, Huber GA. 2003. Boundary element solution of the linear Poisson-Boltzmann equation and a multipole method for the rapid calculation of forces on macromolecules in solution. *J. Comp. Chem.* 24: 353-67
125. Altman MD, Bardhan JP, White JK, Tidor B. 2009. Accurate Solution of Multi-Region Continuum Biomolecule Electrostatic Problems Using the Linearized Poisson-Boltzmann Equation with Curved Boundary Elements. *J. Comp. Chem.* 30: 132-53
126. Bajaj C, Chen SC, Rand A. 2011. An Efficient Higher-Order Fast Multipole Boundary Element Solution for Poisson-Boltzmann Based Molecular Electrostatics. *SIAM J. Sci. Comp* 33: 826-48
127. Yap E-H, Head-Gordon T. 2010. New and Efficient Poisson-Boltzmann Solver for Interaction of Multiple Proteins. *J. Chem. Theo. Comp.* 6: 2214-24
128. Mackoy T, Harris RC, Johnson J, Mascagni M, Fenley MO. 2013. Numerical Optimization of a Walk-on-Spheres Solver for the Linear Poisson-Boltzmann Equation. *Comm. Comp. Phys.* 13: 195-206
129. Holst M, Baker N, Wang F. 2001. Adaptive multilevel finite element solution of the Poisson-Boltzmann equation I. Algorithms and examples. *J. Comp. Chem.* 22: 475-
130. Im W, Beglov D, Roux B. 1998. Continuum solvation model: computation of electrostatic forces from numerical solutions to the Poisson-Boltzmann equation. *Comput. Phys. Comm.* 111: 59-75
131. Konecny R, Baker NA, McCammon JA. 2012. iAPBS: a programming interface to Adaptive Poisson-Boltzmann Solver (APBS). *Comp. Sci. & Dis.* 5
132. Holst M. 2001. Adaptive numerical treatment of elliptic systems on manifolds. *Adv. Comp. Math.* 15: 139-91
133. Boschitsch AH, Fenley MO. 2011. A Fast and Robust Poisson-Boltzmann Solver Based on Adaptive Cartesian Grids. *J. Chem. Theo. Comp.* 7: 1524-40
134. Yap E-H, Head-Gordon T. 2010. New and Efficient Poisson Boltzmann Solver for Interaction of Multiple Proteins. *J. Chem. Theo. Comp.* 6: 2214-24
135. Yap E-H, Head-Gordon TL. 2012. Calculating the Bimolecular Rate of Protein-Protein Association with Interacting Crowders. *J. Chem. Theo. Comp.* 9: 2481-9

136. Yap E-H, Head-Gordon TL. Calculating the Bimolecular Rate of Protein-Protein Association with Interacting Crowders. *J. Chem. Theo. Comp.*
137. Lu BZ, Zhou YC, Holst MJ, McCammon JA. 2008. Recent progress in numerical methods for the Poisson-Boltzmann equation in biophysical applications. *Comm. Comp. Phys.* 3: 973-1009
138. Hankins D, Moskowitz J. 1970. Water molecule interactions. *J. Chem. Phys.* 53: 4544-54
Classic study showing the convergence of the many-body expansion at low order.
139. Stoll H, Preuss H. 1977. Direct Calculation of Localized Hf Orbitals in Molecule Clusters, Layers and Solids. *Theo. Chim. Acta* 46: 11-21
140. Fedorov DG, Kitaura K. 2009. *Theoretical background of the fragment molecular orbital (FMO) method and its implementation in GAMESS*. In *The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems*, ed. DG Fedorov, K Kitaura, pp. 5-36. Boca Raton, FL: CRC
141. Wang LP, Chen JH, Van Voorhis T. 2013. Systematic Parametrization of Polarizable Force Fields from Quantum Chemistry Data. *J. Chem. Theo. Comp.* 9: 452-60
142. Wang LP. Simtk.org: ForceBalance: Systematic Force Field Optimization.
A nice example of new automated methods for force field development.
143. Vega C, Abascal JLF. 2011. Simulating water with rigid non-polarizable models: a general perspective. *Phys. Chem. Chem. Phys.* 13: 19663-88
144. Wang L-P, Head-Gordon T, Ponder JW, Ren P, Chodera JD, et al. 2013. Systematic Improvement of a Classical Molecular Model of Water. *J. Phys. Chem. B* just accepted manuscript
145. Cui J, Liu HB, Jordan KD. 2006. Theoretical characterization of the (H2O)(21) cluster: Application of an n-body decomposition procedure. *J. Phys. Chem. B* 110: 18872-8
146. Cobar EA, Horn PR, Bergman RG, Head-Gordon M. 2012. Examination of the hydrogen-bonding networks in small water clusters (n= 2–5, 13, 17) using absolutely localized molecular orbital energy decomposition analysis. *Phys. Chem. Chem. Phys.* 14: 15328-39
147. Sneskov K, Schwabe T, Kongsted J, Christiansen O. 2011. The polarizable embedding coupled cluster method. *J. Chem. Phys.* 134:104108.
148. Zhang Y, Lin H, Truhlar DG. 2007. Self-consistent polarization of the boundary in the redistributed charge and dipole scheme for combined quantum-mechanical and molecular-mechanical calculations. *J. Chem. Theo. Comp.* 3: 1378-98
149. Boulanger E, Thiel W. 2012. Solvent Boundary Potentials for Hybrid QM/MM Computations Using Classical Drude Oscillators: A Fully Polarizable Model. *J. Chem. Theo. Comp.* 8: 4527-38
150. Thompson MA, Schenter GK. 1995. Excited-States of the Bacteriochlorophyll-B Dimer of Rhodopseudomonas-Viridis - a Qm/Mm Study of the Photosynthetic Reaction-Center That Includes Mm Polarization. *J. Phys. Chem.* 99: 6374-86

TABLES

Table 1. *Accuracy of AMOEBA solvation free energies for small molecules compared to experiment*. All units are kcal/mol. Reprinted with permission from ((47)). Copyright 2010 ACS.

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	-10.0
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

Table 2. *Properties of water calculated using several polarizable models and compared to experimental measurements.* Liquid bulk properties are measured at 298 K, 1 bar; TMD and T_m are measured at 1 bar, and T_c is determined for the critical pressure of the model. Reprinted with permission from ((144)). Copyright 2013 ACS.

Property	Experiment	AMOEBA	SWM4-NDP	TTM3-F	GCPM	SWM6	BK3	iAMOEBA
ρ / g cm^{-3}	0.997	1.000	0.994 (2)	0.994	1.007	0.996 (2)	0.9974 (2)	0.997
ΔH_{vap} / kcal mol^{-1}	10.52	10.48	10.44	11.4	11.30	10.52	10.94	10.94
α / 10^{-4} K^{-1}	2.56	1.9 (6)			4.2		3.01 (8)	2.5 (1)
κ_T / 10^{-6} bar^{-1}	45.3	66 (1)					44.4 (7)	41.1 (4)
C_p / cal mol^{-1} K^{-1}	18.0	21.3 (5)			22.5		22.0 (2)	18.0 (2)

ε(0)	78.5	81.4 (14)	78.0 (14)	67.7	84	78.1 (28)	79 (3)	80.7 (11)
D_0 / 10^{-5} cm^2 s^{-1}	-10.68	2.0	2.85 (28)	2.37	2.26	2.14 (19)	2.28 (4)	2.54 (2)
η / mPa s	0.896	1.08 (5)	0.66 (9)			0.87 (12)	0.95 (1)	0.85 (2)
TMD / K	277	292 (2)	< 220		255	235	275 (3)	277 (1)
T_m / K	273.15		< 120	248 (2)			250 (3)	261 (2)
T_c / K	647.1	581 (2)	576		642		629 (5)	622

Table 3. Various test systems of the many-body expansion (Eq. (9)) approximation to the parent AMOEBA model and corresponding error. The solute-water clusters were not done with Ewald, while the water clusters used standard Ewald with a real space cutoff of 7 Å was used.

System	ΔU_1^{ind}	ΔU_2^{ind}	ΔU_3^{ind}	$\sum_{i=1}^{3} \Delta U_i^{ind}$	ΔU_N^{ind}	***% Error***
$SO_4^{2-}(H_2O)_7$	0.0000	-57.1948	11.8837	-45.3111	-47.4281	-4.4636
$SO_4^{polar}(H_2O)_{216}$	-0.4925	-582.1794	-72.4687	-654.6481	-651.9152	0.4192
$SO_4^{hphobe}(H_2O)_{216}$	-0.4932	-581.1871	-72.6796	-653.8667	-651.1337	0.4197
$(H_2O)_{17}$	-0.6406	-80.9481	-15.8696	-96.8191	-97.0842	-0.2731
$(H_2O)_{216}$	-0.5050	-593.3109	-75.0113	-668.3223	-665.5851	0.4112
$(H_2O)_{365}$	-0.3573	-997.3359	-128.6859	-1126.0218	-1119.5548	0.5776
$(H_2O)_{430}$	-0.4345	-1102.0153	-141.8521	-1243.8674	-1239.7052	0.3357
$(H_2O)_{512}$	-0.6152	-1617.2707	-230.8362	-1848.1069	-1840.3689	0.4205
$(H_2O)_{1000}$	-2.2703	-1901.2646	-217.8616	-2119.1263	-2115.6351	0.1650

FIGURE CAPTIONS

Figure 1. *Mean signed errors in calculated solvation free energies for chemical components of amino side chain analogues.* Results using the HF/6-31G* charge model with optimized solute-water (TIP4P-Ew) van der Waals parameters (magneta) are compared with benchmark results for the AM1-BCC charge model with TIP3P water (light blue) and HF/6-31G* charge model with unoptimized solute-water (TIP4P-Ew) van der Waals parameters (purple). Reprinted with permission from ((35)). Copyright 2012 ACS.

Figure 2. *The Linearized Poisson Boltzmann energy approximated by direct (red), 2-body (blue), 3-body (magenta) and full mutual (green) polarization as a function of separation, a, for salt concentrations of (A) 0.001 M, (B) 0.05 M, and (C) 0.1 M.* The system is comprised of 4 spherical molecules of radius 21.8 Å and net charge -5e placed at the center of sphere, which is a simplification of the barstar protein molecule. The dielectric constant of the spheres and the solvent are $\varepsilon_{in} = 4$ and $\varepsilon_{out} = 78$, respectively. The multipole expansion has been calculated up to order p = 30 poles.

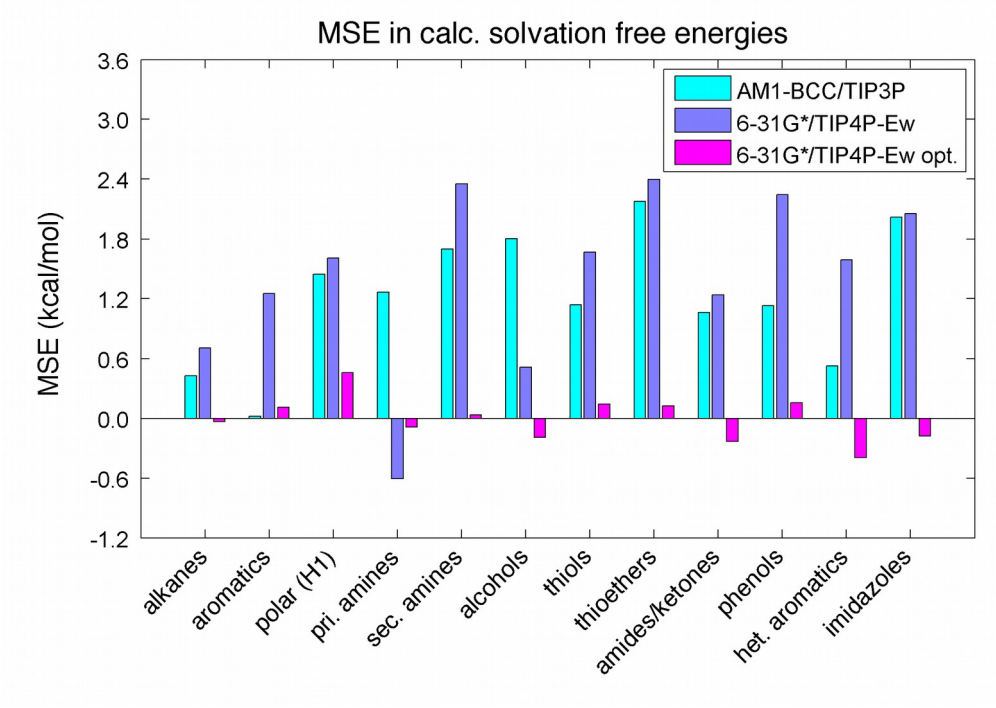
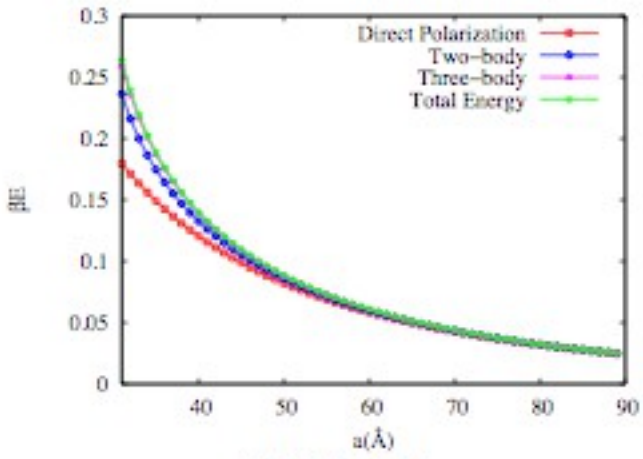
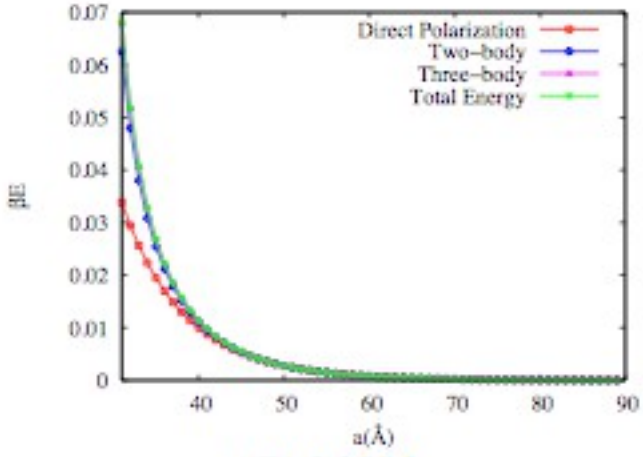


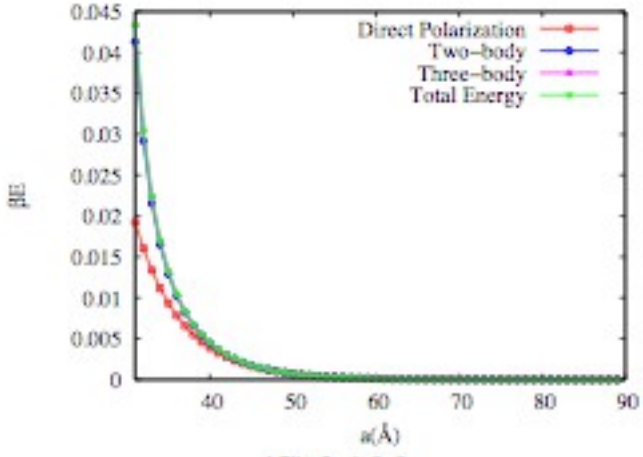
Figure 1. Demerdash and co-workers



(A) 0.001 M



(B) 0.05 M



(C) 0.1 M

Figure 2. Demerdash and co-workers

ACRONYMS

AMBER: Assisted Model Building with Energy Refinement

AMOEBA: Atomic Multipole Optimized Energetics for Biomolecular Applications

CHARMM: Chemistry at HARvard Molecular Mechanics.

DLVO: Derjaguin, Landau, Verwey and Overbeek

GROMACS: GROningen MAchine for Chemical Simulations

GROMOS: GROningen MOlecular Simulation

NEMO: Non-Empirical Molecular Orbitals

OPLS: Optimized Potential for Liquid Simulations

QMPFF: Quantum Mechanical Polarizable Force Field

SDFF: Spectroscopically Determined Force Field

SIBFA: Sum of Interactions Between Fragments *Ab Initio*

SPC: Simple Point Charge

TIP: Transferable Intermolecular Potential

SUMMARY POINTS

1. Many-body effects play a key role in solvation, phase properties, and intrinsic structural propensities of biological macromolecules. Mutual polarization is the most dominant of the many-body effects owing to its long-range behavior.

2. Due to the computational tractability of pairwise-additive empirical force fields, there has emerged a major effort to carefully appraise their shortcomings and improve their ability to describe bulk properties of water, solvation, and protein properties.

3. Aside from inclusion of many-body effects such as polarization, there exists a need for models that include off-center atom charges. Distributed multipole analysis is a straightforward method grounded in classical electrostatics that can be used to model off-atom charges.

4. Three main methods exist for modeling mutual polarization classically: inducible multipoles, fluctuating charge/charge equilibration, and the Drude oscillator method.

5. Continuum electrostatic models such as the linearized Poisson-Boltzmann equation are essential when modeling at the mesoscale. Recently developed methods for solving the Poisson-Boltzmann equation include taking into account mutual polarization, as well as analytical and semi-analytical methods that obviate the limitations of numerical methods.

6. Truncation of the many-body expansion at low order, both in the atomistic and continuum regimes, can be used to reduce the computational cost of full N-body mutual polarization. Preliminary results show that this approximation does not lead to large errors. Recent parameterization of water with only direct polarization (iAMOEBA) gave excellent agreement with a number of experimentally determined bulk properties.