CHARMM Fluctuating Charge Force Field for Proteins: I Parameterization and Application to Bulk Organic Liquid Simulations

SANDEEP PATEL, CHARLES L. BROOKS III
Department of Molecular Biology (TPC-6), The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, California 92037

Received 28 April 2003; Accepted 8 July 2003

Abstract: A first-generation fluctuating charge (FQ) force field to be ultimately applied for protein simulations is presented. The electrostatic model parameters, the atomic hardnesses, and electronegativities, are parameterized by fitting to DFT-based charge responses of small molecules perturbed by a dipolar probe mimicking a water dipole. The nonbonded parameters for atoms based on the CHARMM atom-typing scheme are determined via simultaneously optimizing vacuum water-solute geometries and energies (for a set of small organic molecules) and condensed phase properties (densities and vaporization enthalpies) for pure bulk liquids. Vacuum solute-water geometries, specifically hydrogen bond distances, are fit to 0.19 Å r.m.s. error, while dimerization energies are fit to 0.98 kcal/mol r.m.s. error. Properties of the liquids studied include bulk liquid structure and polarization. The FQ model does indeed show a condensed phase effect in the shifting of molecular dipole moments to higher values relative to the gas phase. The FQ liquids also appear to be more strongly associated, in the case of hydrogen bonding liquids, due to the enhanced dipolar interactions as evidenced by shifts toward lower energies in pair energy distributions. We present results from a short simulation of NMA in bulk TIP4P-FQ water as a step towards simulating solvated peptide/protein systems. As expected, there is a nontrivial dipole moment enhancement of the NMA (although the quantitative accuracy is difficult to assess). Furthermore, the distribution of dipole moments of water molecules in the vicinity of the solutes is shifted towards larger values by 0.1–0.2 Debye in keeping with previously reported work.

 J Comput Chem 25: 1–15, 2004

Key words: molecular simulations; polarizable force field; charge equilibration; density functional theory

Introduction

Classical molecular simulations are an integral tool in the theoretical study of wide-ranging physico-chemical systems, including biomacromolecular systems such as proteins and DNA/RNA.[1–3] Underlying the application of classical statistical mechanical methods such as molecular dynamics and Monte Carlo is the force field, the partitioning of the quantum mechanical interactions into classical terms such as bond stretching, angle bending, dihedral rotations, out-of-plane bending, and electrostatic and dispersive interactions, which are drastically simplified as effective two-body potentials, in part, for purposes of computational efficiency.[3]

Until recently, force fields have included electrostatic interactions via a mean field approach where partial charges are assigned to sites defined by the force field and allowed to interact via a Coulomb interaction. Electrostatic parameters (generally charges assigned to relevant sites) are determined from *ab inito* gas-phase calculations (using sophisticated approximations and levels of theory) on model compounds representing desired chemical environments; charge sets are extracted from the computed electronic densities via fits to electrostatic potentials.[4–6] A final scaling procedure may be applied, in a mean-field spirit, to correct for polarization (or in general many-body) effects arising in the condensed phase.[4,7,8] Certainly the application of an averaged polarization effect is deficient, particularly in interfacial regions (or more generally in cases involving strong anisotropy); this could be the liquid–vapor interface for a neat liquid or the interfacial region between protein and solvent. In this article we address a means to account in part for many-body polarization effects in classical molecular simulations, the fluctuating charge method.

The importance of accurately treating polarization in various physico-chemical systems has been documented in the literature.

Correspondence to: C. L. Brooks; e-mail: brooks@scripps.edu

Contract/grant sponsor: The NIH and the Division of Research Resources (the Pittsburgh Supercomputing Center's Biomedical Research Resource); contract/grant number: RR06009

Brooks and Foresman showed the deviation from two-body additivity of the association energy of a chloride ion with clusters of water molecules due to the polarization of the ion.[9] Dang et al. implemented polarizable models for small molecules for studying liquid–vapor interfacial systems, showing evidence of surface states (free energy minima) as halides pass through the aqueous vapor–liquid interface.[10] Tobias has presented similar evidence for the importance of polarization in the solvation behavior of metal ions near interfaces, specifically relating to the surface segregation of heavy halide anions in water clusters.[11]

Methods of treating polarization, which have found widespread application, have been the point dipole, fluctuating charge, and Drude shell (dispersion oscillator) models. The point dipole approach has found applications ranging from the study of bulk and interfacial properties of neat liquids to the calculation of free energies of ion transport across aqueous/organic interfaces.[10,12–16] The Drude model has been applied to study bulk water[17,18] and ion solvation;[19] currently, there is ongoing effort to parameterize a shell model for application to proteins and peptides.[20] The fluctuating charge method has found comparably widespread use, particularly within the last few years. Once again, it has been extensively applied to the study of various properties of bulk water, aqueous solvation of amides, hydration of the chloride ion, and polyanionic systems.[21–34] Interestingly, the fluctuating charge model has been applied to implicitly solvated systems (solute treated with explicit fluctuating charges)[25] as well as coupled with a QM/MM approach in which the classical region is allowed a polarization response via the fluctuating charge formalism.[35] In this contribution we present our initial progress towards a fluctuating charge force field for protein simulations; the following discussion addresses the model, its parameterization, and application to small molecule systems as test cases. In an ensuing article, we plan to present the application to explicitly solvated proteins.

Model

The fluctuating charge model has been applied to various systems over the last decade.[21,23–25,27–30,33,35–43] The formalism is well documented, and the reader is directed to the relevant literature for details of the theory. Here we present an abbreviated discussion of the method and the points relevant to its parametrization and implementation within CHARMM.

The fluctuating charge method derives from the density functional theory of atoms in molecules as formulated by Yang and Parr.[44] More fundamentally, the method is founded on Sanderson's idea of electronegativity equalization.[45,46] In the density functional sense, electronegativity equalization amounts to the equalization of the chemical potential in space. In a molecule, this translates to the redistribution of charge amongst constituent atoms so as to equalize the electronegativity (chemical potential) at each point.

The electrostatic energy of an N-atom molecule in vacuum is:

$$E_{\text{electrostatic}} = \sum_{i=1}^{N} \chi_i^0 Q_i + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \eta_{ij} Q_i Q_j + \sum_{i=1}^{N} \Phi_i Q_i \qquad (1)$$

The χs are atom electronegativities (formally the Mulliken electronegativities) and are related to the chemical potential of an electron gas surrounding a nucleus by,[24]

$$\mu_i = \frac{\partial E}{\partial N} = -\chi_i^0 = -e \frac{\partial E}{\partial q_i}. \qquad (2)$$

The ηs are the atom hardnesses. These values represent a "resistance" to electron flow to/from an atom. The last term in eq. (1) represents contributions from an external potential with Φ_i being an external electrostatic potential at site i. We note here that in the parameterization scheme, the electronegativities and hardnesses are taken to be adjustable parameters; there is no *a priori* requirement that they equate to "experimental" values, an approach that has been successfully demonstrated by Berne, Friesner, and coworkers in their development efforts.[39,42]

Homogeneous hardness values are parameterized as discussed below. Heterogeneous elements are derived from the atom type values based on the combining rule:[47]

$$\eta_{ij}(R_{ij}, \eta_i, \eta_j) = \frac{\frac{1}{2}(\eta_i + \eta_j)}{\sqrt{1 + \frac{1}{4}(\eta_i + \eta_j)^2 R_{ij}^2}}, \qquad (3)$$

where R_{ij} is the separation between atoms (or more generally *sites*) i and j, and the atomic hardness parameters are the ηs.This local screened Coulomb potential has the correct limiting behavior as $1/r$ for separations greater than about 2.5 Å. This interaction is computed for 1–2, 1–3, and 1–4 sites (sites included in bonds, angles, and dihedrals). Sites separated by four or more bonds interact via a Coulomb interaction; in the case of interacting molecules, the interaction between sites on different molecules is again of the Coulomb form. We note here that an alternate model for the shielding interaction is the Coulomb overlap integral between atom-centered Slater orbitals.[24,25,36] This approximation is computationally tractable for rigid systems where the interaction matrix elements can be computed once at the start of a simulation and then used for the duration; however, for flexible systems, calculation of the overlap integral leads to nontrivial cost.

In the present case, allowing the 1–4 shielded interaction to be coupled with the functional form of the hardness (explicit dependence on nuclear coordinates) is expected to incorporate a conformational dependence of the charge distribution, a behavior that one can expect to be important in macromolecular simulations. The fluctuating charge model, in contrast to dipole polarizable models, by including explicit charge–charge interactions, incorporates higher order multipoles that are neglected in the former case.[25] However, in certain circumstances, the symmetry of the polarizability tensor is compromised because the model is restricted to having charges at particular sites. In the case of planar molecules or molecular fragments (aromatic rings, water, etc.), out-of-plane polarization is neglected in the current implementation; however, this can be remedied by the careful addition of extra off-atom charge carrying sites (e.g., lone pairs).

With respect to charge dynamics, an extended Lagrangian formalism is used to propagate the charges in time with some

general charge constraint, thus strictly providing for electronegativity equalization at each dynamics step.[48–51] For the present work, molecular charge conservation is imposed; however, this can be modified to allow for charge transfer between molecular entities. In the case of proteins, or polymeric systems in general, normalization can be over physically relevant units (e.g., amino acid residues).[24]

For a system of M molecules, with N_i atoms on molecule i, the total energy is,

$$E(Q, r) = \sum_{i=1}^{M} \sum_{\alpha=1}^{N_i} \chi^o_{i\alpha} Q_{i\alpha} + \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \left(\sum_{\alpha=1}^{N_i} \sum_{\beta=1}^{N_j} \eta_{i\alpha j\beta} Q_{i\alpha} Q_{j\beta} + V(r_{i\alpha j\beta}) \right), \quad (4)$$

where $r_{i\alpha j\beta}$ is the distance between sites α and β on molecules i and j, respectively. The first two terms in this expression represent the electrostatic contribution, and the nonelectrostatic contributions are included in the last term with $V(r_{i\alpha j\beta})$ including intramolecular potential terms (bond stretching, angle bending, dihedrals, etc.) and dispersion interactions of the Lennard–Jones type. The system Lagrangian with the associated constraint on total molecular charge for each molecule is then,

$$L = \sum_{i=1}^{M} \sum_{\alpha=1}^{N_i} \frac{1}{2} m_{i\alpha} \dot{r}^2_{i\alpha} + \sum_{i=1}^{M} \sum_{\alpha=1}^{N_i} \frac{1}{2} m_{Q,i\alpha} \dot{Q}^2_{i\alpha} - E(Q, r) - \sum_{i=1}^{M} \lambda_i \sum_{\alpha=1}^{N_i} Q_{i\alpha} \quad (5)$$

where the first two terms represent the nuclear and charge kinetic energies, the third term is the total potential energy, and the fourth term is the molecular charge neutrality constraint with λ_i the Lagrange multiplier for each molecule, i. The fictitious charge dynamics, analogous to the fictitious wavefunction dynamics in Car-Parrinello (CP) type methods,[51,52] is determined with a fictitious charge "mass" (adiabaticity parameter in CP dynamics). The units for this mass are (energy·time^2/charge^2). The equations of motion for the electronic degrees of freedom are given by,

$$m_{Q,i\alpha} \ddot{Q}_{i\alpha} = - \frac{\partial E(\bar{Q}, \bar{r})}{\partial Q_{i\alpha}} - \lambda_i. \quad (6)$$

One solves for the Lagrange multipliers by substituting the previous equation into the constraint condition for total molecular charge,

$$\sum_{i=1}^{M} \sum_{\alpha=1}^{N_i} \ddot{Q}_{i\alpha} = 0, \quad (7)$$

which gives for the multiplier,

$$\lambda_i = - \frac{\sum_{\alpha=1}^{N_i} \frac{\partial E(\bar{Q}, \bar{r})}{\partial Q_{i\alpha}}}{N_i} = - \frac{\sum_{\alpha=1}^{N_i} (\tilde{\chi}_{i\alpha})}{N_i}. \quad (8)$$

The force on a charge site then becomes the difference between the average electronegativity of the molecule and the instantaneous electronegativity at the site,

$$m_{Q,i\alpha} \ddot{Q}_{i\alpha} = - \frac{\sum_{\beta=1}^{N_i} (\tilde{\chi}_{i\alpha} - \tilde{\chi}_{i\beta})}{N_i}. \quad (9)$$

In terms of implementation, the charge constraint is applied by normalizing the force over the sites within a unit; that is, the instantaneous average mass-weighted force is subtracted from the instantaneous total force on each site.

Parameterization of the Fluctuating Charge Model

Establishing Atomic Hardness and Electronegativity Parameters

We reiterate that the present approach to implementing the fluctuating charge model treats the electronegativities and hardnesses as empirical parameters. There is no attempt to associate the electronegativities to Pauling values (or to even relate the present parameters with each other via some electronegativity scale). One should bear in mind, also, that the lone atom electronegativities as defined in this formalism are not true lone atom values, i.e., they are not strictly determined for single atoms in the gas phase from the atomic ionization potential and electron affinity.[43]

To separate the fitting of the atom type electronegativities and hardnesses, the fluctuating charge model is reformulated in terms of a linear response model.[37,42,53] Equation (1) gives the electrostatic energy of a molecule in the presence of some external potential. In the absence of any perturbing external field, the last term vanishes to give

$$E^0_{\text{electrostatic}} = \sum_{i=1}^{N} \chi^0_i Q^0_i + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \eta_{ij} Q^0_i Q^0_j, \quad (10)$$

where the Q^0_i are site charges in the absence of an external field.

In each case, at equilibrium, the condition that

$$\frac{\partial E_{\text{electrostatic}}}{\partial Q_i} = 0, \quad i = 1, N \quad (11a)$$

and

$$\frac{\partial E_{\text{electrostatic}}}{\partial Q^0_i} = 0, \quad i = 1, N \quad (11b)$$

Table 1. Comparison of Fitted and Predicted Gas-Phase Dipole Moments for Selected Organic Molecules.[a]

Fitting set	Fit	DFT	Experimental	Test set	FQ (minimized)	Experimental
Liquid	μ	μ	μ	Liquid	μ	μ
NMA	3.94	3.85	3.76	Acetone	2.78	2.88
Methanol	1.69	1.68	1.7	Acetic acid	1.95	1.74
Ethanol	1.61	1.60	1.69	Benzene	0.0	0.0
Methyl acetate	2.01	2.02	1.71	2-Propanol	1.56	1.66
Methyl amine	1.43	1.40	1.31	Diethylether	1.0	1.25
Ethanethiol	1.76	1.76	1.6	Formic acid	1.7	1.4
Formamide	4.02	4.02	3.73	Dimethylsulfide	1.7	1.554
Acetamide	3.99	3.99	3.76	2-Butanone	2.69	2.78
Ethylbenzene	0.39	0.39	0.59	Phenol	1.6	1.45
Ethylimidazole	4.15	4.15		Toluene	0.49	0.375
ALAD, $C7_{eq}$	2.1	2.34		Methyl formate	2.26	1.77
ALAD, $C7_{ax}$	3.18	3.59		Propanoic acid	2.44	1.75
ALAD, C_5	2.69	3.44				
ALAD, α_R	7.76	7.53				
ALAD, α_L	7.67	7.36				
ALAD, β	4.09	4.15				
ALAD, β_2	4.88	5.16				
TIP4P-FQ	1.85	1.856	1.85			

[a]The set used for fitting is on the left; that used as a test set on the right. Dipole moments in Debye. The last row shows the TIP4P-FQ dipole moment for reference.

must be satisfied leads to the following expressions for the cases with and without an external potential,

$$\bar{\bar{\eta}}\bar{Q} = -(\bar{\chi} + \bar{\phi}) \tag{12a}$$

and

$$\bar{\bar{\eta}}\bar{Q}^0 = -\bar{\chi}. \tag{12b}$$

Taking the difference of eqs. (12a) and (12b) yields an expression for the response of the molecular charge distribution to the external probing field

$$\bar{\bar{\eta}}\Delta\bar{Q} = -\bar{\phi} \Rightarrow \Delta\bar{Q} = -\bar{\bar{\eta}}^{-1}\bar{\phi}. \tag{13}$$

We now have a relationship between the atomic hardness and the charge response due to the external potential.

An objective function, measuring the deviation from the parameterized model and the target response (determined from density functional theory (DFT)-based quantum chemistry) becomes

$$\varepsilon = \|(\Delta Q^{DFT} - \Delta Q^{FQ})\|. \tag{14}$$

To achieve optimal model parameters, we minimize the norm of the sum of squared differences between the DFT and FQ charge responses.

The parameterization proceeds as follows. Because we desire to maintain the CHARMM atom-typing scheme, we select a range of model compounds inclusive of the chemical functionalities encountered in biological systems. For a given molecule, sets of 30 perturbing fields are generated. This is done by placing a dipolar probe (consisting of two oppositely charged point charges of 0.7815e separated by 0.5744 Å, thus reproducing the dipole moment of water) at random locations around the molecule (greater than a distance of 2.0 Å from any heavy atom of the molecule of interest).[42]

For each set of perturbations, the hardness parameters for relevant atom types are fitted by minimizing eq. (14); the final hardness is calculated as the average from all sets, with the contribution from each set being weighted by the error of the fit for the set. The optimized DFT gas-phase geometry is maintained for the perturbation calculations. A simple simulated annealing search in parameter space is used for the parameter optimization. No explicit constraint function for the total charge is implemented since the CHelpG-based charge differences are normalized by design. Finally, we note that there appears to be some dependence of the "convergence" of parameters on the number of perturbations (dipole probe configurations) used per set; after about 30 perturbations, values appear to converge to a narrow region of parameter space (based on test calculations for fitting to a water model).

Density functional calculations using the B3LYP exchange-correlation functional are used to generate the electrostatic potential. The basis set, 6-31++G(2d+,p), is chosen, as it was shown to be able to reproduce the gas phase dipole moment for water (1.856 D);[54] furthermore, Table 1 demonstrates the robustness of this level of theory for reproducing gas-phase dipole moments of a variety of small molecules. This basis set adds additional diffuse polarization functions with exponents one-fourth as large as for the previous set for atoms with lone pairs (oxygen and nitrogen).[54] Charges for each configuration of external potential are determined by the CHelpG fit to the DFT electrostatic potential.[5,6] Zero-field charges are determined for optimized geometries for each mole-

Table 2. Comparison of Molecular Plarizabilities: FQ vs. Experiment.

Molecular polarizabilities from FQ and experiment			
Molecule	1/3 Tr (σ) (FQ)	1/3 Tr (σ) (Experimental)[b]	%Error[a]
NMA	8.08	7.82	3.3
Methanol	2.70	3.26, 3.20, 3.23 (3.23)	−16.4
Ethanol	5.14	5.07, 4.94 (5.005)	2.70
Acetone	6.51	6.4, 6.33, 6.39 (6.37)	2.1
Benzene	7.94	10.39, 10.32, 10.42 (10.38)	−23.5
Formamide	3.23	4.08, 3.88, 4.08 (4.01)	−19.5
N-methylformamide	5.39	5.89, 5.91 (5.90)	−8.6
Acetamide	5.41	5.39, 5.67 (5.53)	−2.2
Ethanethiol	8.06	7.38	9.2
Methylformate	4.89	5.05	−3.2
Acetic acid	5.26	5.15, 5.05 (5.10)	3.1
Ethane	3.96	4.47	−11.4
n-Propane	8.23	8.12	1.4
n-Butane	6.36	6.29	1.1
Imidazole	6.04	7.19	−16.0
Water	1.1 (TIP4P-FQ)	1.45	−24.1
Avg. absolute %Error			8.3[c]

[a]Errors in last column calculated based on the average experimental value.
[b]Average of experimental values shown in parentheses in third column. Experimental data from Miller et al.[81]
[c]Excluding TIP4P-FQ data.

cule. All DFT and electrostatic potential fitting calculations were performed using GAUSSIAN98.[55]

It should be noted that the hardness parameters are fitted in an hierarchical manner. First, the alkyl group is parameterized (methyl carbons, methyl hydrogens). These parameters are then carried over to the fitting of higher functional groups (alcohols, ketones, aldehydes, amides, etc.). The CHARMM22 atom typing scheme was followed without introducing any new types,[56] although this constraint can be lifted for future refinement and development.

We mention here that the dipole polarizability tensor derives simply from the hardness matrix elements as

$$\bar{\bar{\alpha}} = \Delta\bar{r}\bar{\bar{\eta}}^{-1}\Delta\bar{r}^{T}, \tag{15}$$

where the $\Delta\bar{r}$ are atomic coordinates relative to the center of geometry. This can be used to determine a set of hardnesses by fitting to the quantum mechanically derived gas phase or experimental (gas phase or solution) polarizabilities However, because the atom-centered charge representation adopted currently does not allow for anisotropic polarizability (i.e., the out of plane polarization of aromatic rings) without the introduction of extra sites, we opt not to include this property in the fitting scheme at this time. Table 2 shows values of the average molecular polarizability, calculated as one-third of the trace of the molecular dipole polarizability tensor given in eq. (15). The global average absolute error is 8.3%, with a range falling between 1.1 and 23.5%. It appears that, as expected, polarizabilities of the planar molecules are difficult to reproduce exactly. Furthermore, these data suggest that the ability of the model to reproduce condensed-phase properties may not necessarily be linked to its ability to reproduce vacuum polarizabilities; the case of the TIP4P-FQ model is a good example, because this model, which reproduces a highly accurate gas-phase dipole moment, condensed-phase density and vaporization enthalpy, and liquid dielectric constant, fails to faithfully reproduce the vacuum polarizability.

The parameterization of the atom electronegativities then proceeds by fitting to charge distributions of the optimized geometry in the gas phase (zero-field). The dipole moments are added to the objective function; this is found to be necessary to accurately reproduce the gas-phase moments, as these are sensitive to minor changes in the charge distribution. The electronegativities for the alkyl atom types are fitted to the ethane, propane, and butane DFT charge distributions. For the protein force field, the backbone atom electronegativities are fitted to reproduce the dipole moments and CHARMM22 charge distributions for the conformations of the blocked alanine residue, or alanine dipeptide (ALAD). Thus, the backbone carbonyl carbon, carbonyl oxygen, amide nitrogen, and amide hydrogen electronegativities are determined in this way. The electronegativities for the remaining atom types are determined by fitting to gas-phase charges and dipole moments for model compounds as listed in Table 1. To address the transferability of the electrostatic parameters, the dipole moments of several test molecules are computed and compared to experimental values. Table 1 shows the results. In general, the test cases overestimate the vacuum dipole moment; this appears to be a general trend observed in other parametrizations of the fluctuating charge model.[25,57] Nevertheless, the overall agreement is satisfactory. Table 3 shows the current set of atomic hardness and electronegativity parameters. We note that in the development of the protein force field, certain of the electronegativity parameters, are allowed

Table 3. Fitted Hardness and Electronegativity for Selected CHARMM Atom Types.

Atom type	η[a]	χ[b]	Atom type	η[a]	χ[b]
CT3[c]	240.34	319.65	NH3	159.9	252.73
CT1	196.88	306.79	NH2	260.0	325.53
C_α[d]	240.34	319.65			
CT2	208.92	313.36	HC	517.26	276.49
C	214.44	360.67	HA		
HA3	501.42	315.56	HS	473.4	292.64
HA2	501.42	318.27	S	243.4	358.52
HA1	501.42	319.83	CC	214.44	329.95
H	517.26	263.19	CY	155.6	323.36
NH1	260.0	341.31	CA	225.48	332.37
O	230.06	454.64	HP	454.14	319.46
OH1	307.2	364.85	NY	218.30	302.89
NR1	152.24	338.34	CPT	191.21	320.67
NR2	246.94	396.73	HR1	531.16	314.08
CPH1	242.68	349.68	HR3	528.2	312.53
CP1	196.00	321.67	OT	371.6	68.49
CP2	228.30	327.8	HT	353.0	0.0
CP3	224.03	324.16	OT-HT	286.4	
			HT-HT	203.6	

[a]Hardness is in units of kcal/mol-e^2.
[b]Electronegativity in units of kcal/mol-e.
[c]CT3 atom type is taken as a global reference value for electronegativities as dictated by the charge constraint conditions.
[d]Corresponds to C_α in peptide backbone.
[e]TIP4P-FQ parameters are shown for reference (OT is water oxygen, HT is water hydrogen, OT-HT and HT-HT correspond to off-diagonal intramolecular elements of the hardness matrix).

to differ between different residues as has been the approach in fixing charges for the CHARMM22 force field.

Optimizing Nonelectrostatic Parameters: Nonbonded Parameters

Following parameterization of the fluctuating charge model, the nonbonded contributions to the total potential are optimized in a manner that attempts to balance solute–solute and solute–solvent interactions.[56] The approach is comprised of two steps: (1) optimizing gas phase dimer geometries of model compounds (representing the range of chemical functionalities encountered in protein simulations) with water, and (2) performing condensed-phase simulations of select model compounds to monitor liquid density and enthalpy of vaporization, in the case of liquids, or crystal cell parameters and heats of sublimation in the case of crystals, as has been done by Yin and MacKerell in the development of the CHARMM force field[58] and Jorgensen and coworkers in the development of the OPLS force fields.[59–61] The first part tries to optimize the functional group–water interactions. For the present, we employ the TIP4P-FQ fluctuating point charge model;[24] this model accurately reproduces the dielectric constant of liquid water at ambient conditions and is thus satisfying to implement as a dielectric medium surrounding proteins. The second part attempts to optimize the self interactions of the particular functional group. On the whole, one hopes to balance the solute–solute and solute–water interactions, the balance being of prime importance in maintaining the proper energetic and structural integrity of the model with respect to experimental data. Properties targeted for parameterization include the bulk liquid densities (molecular volumes) and enthalpies of vaporization, crystal cell parameters (a ,b, c) and angles (α, β, γ) and heats of sublimation. For analysis, we compute vacuum versus condensed phase dipole moment distributions and structure (via radial distribution functions when experimental data is available). We emphasize that properties at one state point are targeted; thus, one might argue that this effectively limits the physical range of applicability. However, most applications to proteins initially are expected to be in the vicinity of ambient conditions under which most of the liquids are studied.

The parameterization protocol is an iterative process. For a given iteration, parameters giving reasonable agreement to vacuum dimer energies and geometries are determined. Condensed phase simulations follow; the results then indicate the next direction of change for the nonbonded parameters.

The nonbonded interaction is of the Lennard–Jones type (standard in the CHARMM implementation);

$$E_{LJ} = \sum_{ij} \varepsilon_{ij}\left(\frac{R_{\min,ij}^{12}}{r_{ij}^{12}} - 2\frac{R_{\min,ij}^{6}}{r_{ij}^{6}}\right) \tag{16}$$

where the summation runs over all nonbonded atom pairs.

Table 4 shows gas-phase water-solute geometries and energies; Figure 1 shows the corresponding geometries. The results indicate a satisfactory balance between solute–solute and solute–solvent interactions has been achieved. In general, the solute–water dimer energies are reproduced sufficiently accurately; in certain cases, such as the *cis* form of NMA, there appears to be more discrepancy. However, these conformations are not overly relevant to the structure of the peptide linkage, in which the amide hydrogen and carbonyl oxygen are *trans* to each other. In terms of geometry, hydrogen bond distances show an rmsd of 0.19 Å. The hydrogen bond angles tend to be more linear than the DFT predictions.

Bulk liquid simulations are performed on a series of liquids as listed in Table 5. Systems examined consist of 216 molecules in a cubic periodic box simulated at constant temperature and pressure; constant temperature is maintained by a Hoover thermostat,[62,63] and constant pressure via the Langevin piston method.[62] Electrostatics are treated fully via Particle Mesh Ewald.[64,65] We emphasize that the implementation of the FQ method as outlined here necessitates full accounting of electrostatic interactions; that is, the standard approach of electrostatic cutoffs used with fixed charge models is inappropriate. This arises from the charge normalization constraint—truncating electrostatic interactions of the sites in a defined normalization unit differently leads to incorrect forces on the charge degrees of freedom, leading to charges taking on unfeasible values. This has also been noted for dipole polarizable and earlier FQ models.[66,67] Crystal simulations for indole and imidazole are performed using the CHARMM crystal facility with a full unit cell (four molecules in both cases) as the primary simulation unit. The crystal parameters are determined from the literature (fractional coordinates for indole[68]) and the Cambridge

Table 4. Vacuum Dimer Geometries and Energies.

	Solute-water dimer energy[a]		Hydrogen bond distance (Å)			Hydrogen bond angle (degrees)		
Geometry	FQ[b]	DFT	FQ[c]	H-bond	DFT	FQ	Angle	DFT
NMAE1	8.74	9.18	2.15	(CO—H)	1.833	140.0	(OHO)	155.2
			2.15	(NH—O)	2.074	142.7	(NHO)	143.1
NMAE2	6.2	6.97	1.92	(CO—H)	1.85	177.7	(OHO)	167.6
NMAZ1	6.38	6.68	1.926	(CO—H)	1.875	175.4	(OHO)	174.8
NMAZ2	6.14	6.82	1.927	(CO—H)	1.853	179.3	(OHO)	168.8
AAZ1	8.14	9.46	1.878	(OH—O)	1.793	159.6	(OHO)	157.2
			2.46	(CO—H)	1.946	128.4	(OHO)	140.3
AAZ2	5.21	5.31	2.11	(C═O—H)	1.936	165.4	(OHO)	161.9
AAZ3	4.88	2.91	1.84	(C—O—H)	2.085	173.8	(OHO)	155.8
FAE1	7.56	7.73	1.83	(OH—O)	1.814	171.7	(OHO)	178.6
FAE2	4.74	4.91	2.1	(C═O—H)	1.991	178.3	(OHO)	148.7
FAZ1	8.15	9.51	1.855	(OH—O)	1.775	162.1	(OHO)	157.1
			2.61	(C═O—H)	2.016	119.4	(OHO)	136.5
FAZ2	4.62	4.43	2.08	(C═O—H)	1.992	178.8	(OHO)	151.4
FAZ3	3.50	2.44	1.86	(C—O—H)	2.212	170.4	(OHO)	132.2
FM1	9.32	8.64	2.04	(CO—H)	1.893	151.1	(OHO)	151.6
			1.97	(NH—O)	2.078	146.0	(NHO)	137.3
FM2	6.33	5.94	1.93	(CO—H)	1.892	178.8	(OHO)	160.7
FM3	6.28	4.60	1.84	(NH—O)	2.040	177.2	(NHO)	179.8
MAE1	5.16	5.16	2.085	(C═O—H)	1.976	178.5	(OHO)	176.7
MAE2	5.49	6.01	2.096	(C═O—H)	1.901	165.5	(OHO)	162.7
MAZ1	6.75	5.28	2.049	(C═O—H)	1.938	167.8	(OHO)	170.8
MAZ2	5.75	5.55	2.062	(C═O—H)	1.916	169.0	(OHO)	164.5
MT1	2.92	3.64	2.47	(CS—H)	2.449	168.0	(SHO)	159.6
MT2	3.35	1.84	2.08	(SH—O)	2.288	174.5	(SHO)	177.6
NMFE1	8.56	8.95	2.06	(CO—H)	1.864	147.4	(OHO)	152.8
			2.23	(NH—O)	2.074	137.8	(NHO)	140.1
NMFE2	6.31	6.34	1.886	(C═O—H)	1.874	179.5	(OHO)	162.7
NMFZ1	6.78	6.35	1.887	(CO—H)	1.892	178.1	(OHO)	171.9
NMFZ2	6.91	4.29	1.98	(NH—O)	2.058	172.0	(NHO)	176.9
M1	4.65	4.59	1.89	(CO—H)	1.911	180.0	(OHO)	170.6
M2	6.07	5.09	1.86	(OH—O)	1.959	167.6	(OHO)	171.6
IM1	6.22	6.85	2.0	(CN—H)	1.936	177.5	(NHO)	178.6
IM2	5.77	5.2	2.03	(NH—O)	1.994	176.6	(NHO)	176.3

[a]Energies in kcal/mol.
[b]Dimer energy RMSD is 0.98 kcal/mol.
[c]Hydrogen bond RMSD is 0.19 Å.

Structural Database for small molecules (in the case of imidazole).[69] Charge degrees of freedom are coupled to a temperature thermostat at 1.0 K using the Nose–Hoover method;[63] the thermostat mass is 0.005 (kcal/mol/ps^2). The charges are given masses ranging between 0.000069–0.00009 (kcal/mol/ps^2); this combination of thermostat and charge masses allows a stable trajectory with a 0.5-fs time step. The time step can be increased with a multiple-time step algorithm as has been implemented for CPMD and classical MD.[70–73] This is ongoing in our lab at present. We note that for each system, all atoms are assigned the same mass, although the same atom type can be assigned a different mass among different systems. For the current work, we choose a rather small timestep to remain consistent with the selected fictitious charge masses which determine the drift from the Born–Oppenheimer surface. Currently, FQ dynamics runs require 10% more resources relative to the CHARMM22 (consistent with the original application to water by Rick et al.) and minimizations require 3% more resources; however, due to the fact that the time step with FQ is limited to 0.5 to 1.0 fs vs. 2.0 fs for CHARMM22, the dynamics are effectively 2–4.5 times slower. This is still competitive with point dipole models under development, which report computational requirements increasing by factors up to 20 times as much as the underlying nonpolarizable models. As the extended Lagrangian method adopted here is analogous to the "on the fly" *ab initio* or Car–Parrinello molecular dynamics methods, one must deal with maintaining a meta-stable state with the fictitious charge kinetic energy at a much lower temperature than the nuclear (ionic) degrees of freedom; this allows the system to remain on the Born–Oppenheimer (BO) surface where the electronic state is the equilibrium ground state for any nuclear configuration. For the

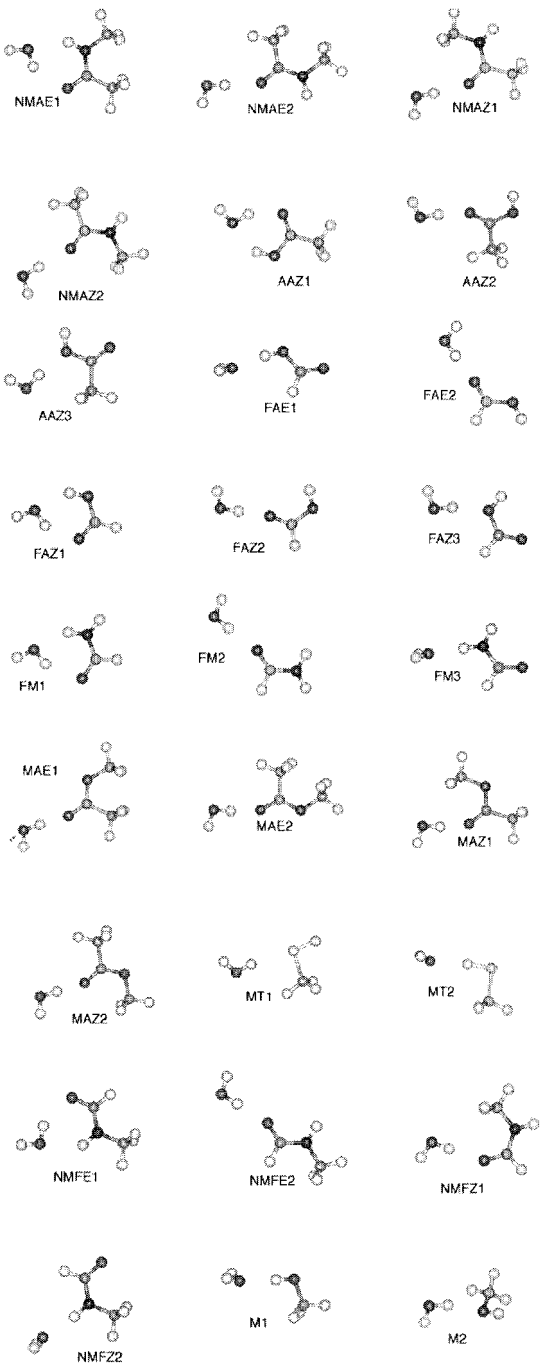

Figure 1. Water-solute geometries used for the vacuum dimer calculations (after Rablen et al.[54]).

present implementation, strong coupling to thermostats is required to keep the charges "cold" (~1.0 K); dynamics without temperature control sees the charge kinetic energy increase to unreasonable levels quite rapidly, not unlike that which is observed in CPMD for metallic systems, where the vanishing bandgap leads to little or no separation in the frequency spectra of the nuclear and electronic degrees of freedom.[74,75] Previously reported work on the application of fluctuating charge models to small molecules consistently claims weak coupling of the nuclear and electronic degrees of freedom.[24,27] This is perhaps due to the more or less rigid nature of the molecular systems with few internal degrees of freedom to which the charge degrees can couple; the present form of the shielded Coulomb interaction explicitly introduces the coupling to nuclear degrees of freedom, and the need for strict temperature control should not be surprising. The "motions" of the charge degrees of freedom can span a large range of frequencies, and overlap with the nuclear motions typically associated with classical force fields. Figure 2a shows the frequency spectra for the nuclei and charges of a single *N*-methylacetamide molecule in vacuum. Up to 2000 cm^{-1} are the modes arising from the bond, angle, and torsional force constants of the underlying potential; note, however, that there are a few electronic modes in this regime as well, as NMA is a fairly small molecule, with the overlap not being too severe. Also shown is the regime of higher frequency modes for the charges. Figure 2b shows the same information for the alanine dipeptide in vacuum. In this case, we see the increase of nuclear modes and the concomitant increase in overlap between nuclear and electronic modes. In practice, this behavior necessitates stricter temperature control for the charges as system size increases.

In the next section, we present results from bulk liquid simulations of select small molecule liquids to demonstrate the feasibility of the fluctuating charge model for condensed phase simulations in capturing many-body effects.

Results and Discussion

Bulk Liquid Structure

Figure 3 shows a set of representative radial distribution functions for relevant atomic pairs computed from dynamics trajectories. In general, the distributions represent liquid-like behavior, although due to lack of scattering data for most of the complex fluids, it is difficult to comment on the accuracy. At present, we compare to the structure determined from simulations based on the nonpolarizable CHARMM22 potentials, but as will be evident, the structures observed to give thermodynamic properties close to experimental values at times differ significantly from force field to force field. Of course, this is one manifestation of the multiple-minima nature of force field parametrization. Overall, the polarizable model predicts more structured condensed phases through sharper first peaks in the case of hydrogen bonding systems. This has been noted in previous applications of polarizable potentials, and is attributed to the enhanced charges allowing for stronger interactions. In the present case, the shift of condensed phase dipole moments (to be discussed below) to higher values is consistent with this behavior.

Methanethiol and ethanethiol (Fig. 3a and b) indicate a more structured liquid relative to the fixed charge force field. In fact, the current FQ model is more structured than the OPPE potential for the thiols,[76] which gives similar density and vaporization enthalpy predictions for the bulk liquids. However, compared to the OPLS united-atom model, the current potential shows less order and no strong first peak near 2 Å as found with OPLS.[77] The polarizable

Table 5. Molecular Volumes and Vaporization Enthalpies for Test Liquids.[a]

	Molecular volume (Å^3)			Enthalpy of vaporization (kcal/mol)		
Liquid	FQ	CHARMM22	Experimental	FQ	CHARMM22	Experimental
Ethane	91.3 (0.01)	91.5	91.5	3.6	3.57	3.52
Propane	125.1 (0.02)	126.0	126.0	4.02 (0.08)	4.51	4.49
Butane	159.2 (0.02)	162.9	160.3	5.21 (0.013)	5.26	5.35
N-Methylacetamide	139.9 (0.14)	133.7	135.9	13.91 (0.025)	13.85	13.3; 14.2
Methanol	69.4 (0.04)	69.75	67.3	8.91 (0.015)	9.13	9.0
Ethanol	101.9 (0.08)	96.71	96.76	10.78 (0.026)	10.71	10.108
Formamide	70.2 (0.09)	67.4	66.17	16.35 (0.03)	13.3	14.36
Methylamine	74.9 (0.02)	73.2	78.61	6.44 (0.0171)	5.82	6.12
Acetone	123.7 (0.04)	121.3	121.89	7.91 (0.016)	8.33	7.4
Methanethiol	88.4 (0.04)	90.1	89.95	7.2 (0.01)	5.38	5.86
Ethanethiol	118.5 (0.05)	120.4	119.82	7.8 (0.011)	6.86	6.82
Benzene	146.8 (0.06)	153.05	147.8	7.5 (0.04)	7.73	8.09
Phenol	153.59 (0.05)	155.9	149.77	13.78 (0.017)	12.99	13.24
Methyl formate	114.65 (0.12)	105.3	102.0	7.98 (0.03)	8.08	6.776
Imidazole (crystal)	a = 8.38 (0.06)	a = 7.45	a = 7.75	19.83 (0.09)	18.0	19.3–19.84
	b = 5.76 (0.03)	b = 4.75	b = 5.45	($\Delta H_{sub.}$)		
	c = 9.82 (0.02)	c = 9.91	c = 9.75			
	$\alpha = \gamma$;	$\alpha = \gamma$;	$\alpha = \gamma$;			
	β = 106.0 (1.04)	β = 102.1	β = 117.5			
Indole (crystal)	a = 7.37 (0.03)	a = 8.086	a = 7.86	18.33 (0.06)	15.14	17.9–18.59
	b = 5.70 (0.02)	b = 5.70	b = 5.66	($\Delta H_{sub.}$)		
	c = 14.53 (0.03)	c = 14.49	c = 14.89			
	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$			

[a]Simulation temperatures are as follow: Ethane, 184.5 K; Propane, 231.08 K; Butane, 272.65 K; methanethiol, 279.11 K; ethanethiol, 298 K; NMA, 373 K; methanol, 298 K; ethanol, 293 K; formamide, 293 K; acetone, 298 K; methyl formate 295 K; benzene 298 K; methylamine, 266.8 K; phenol, 333.15 K; imidazole, 298 K; indole, 298 K.

liquid structure does, however, resemble that of the nonpolarizable CHARMM22 force field, although the first peak for the FQ case is slightly stronger; this is probably due to a greater first-shell association arising from the stronger electrostatic interaction between higher charges in the FQ model relative to the fixed-charge model. The methanol (Fig. 3f) and NMA (Fig. 3e) radial distrubition functions are comparable to the CHARMM22 additive results, thus indicating again that the current model does not disrupt bulk liquid structure. The distributions are also similar to those reported by Kollman using atom centered polarizabilities and reduced fixed charges (relative to the OPLS potential).[13]

Bulk Liquid Energies, Densities, and Polarization Effects

Table 5 shows the enthalpies of vaporization and molecular volumes of the test liquids under the conditions indicated in the caption (generally ambient conditions or at the normal boiling points; crystal simulations are performed at 298 K and atmospheric pressure); enthalpies are computed using

$$\Delta H_{\text{vaporization}} = \langle E_{\text{liquid}} \rangle - \langle E_{\text{gas}} \rangle + RT, \tag{17}$$

where the angle brackets denote ensemble averages over the dynamics trajectories of the liquid and gas-phase potential energies (per molecule). Gas-phase potential energies are calculated from short dynamics runs (5–10 ps) of a single molecule. For the solids at ambient conditions, heats of sublimation are calculated; this requires dynamics simulations of crystals of the relevant small molecule unit cells. The nonbonded parameters are fitted in a hierarchical manner (much like the atomic hardnesses described above). The aliphatic hydrogens and carbons (sp3-hybridized) are determined via fits to the alkanes. These parameters are than held fixed for the iterations over, for example, the alcohols, to determine the relevant parameters for the hydroxy oxygen and hydrogen. This procedure is carried out to determine a first set of parameters for the atom types found in the CHARMM22 protein force field. Table 6 lists the current and fixed charge nonbonded parameters for relevant atom types.

Overall, the polarizable model captures the bulk liquid energies satisfactorily, and is comparable to the fixed charge force field. Although this is only a cursory test of the applicability of the force field, these initial tests indicate the transferability of the model from vacuum to condensed phases. Specifically, the chemical environments are faithfully modeled using the atom types defined in the CHARMM force field. The liquid densities are reproduced well, with the average absolute error for the FQ model being 2.6 vs. 2.0% for the CHARMM22 nonpolarizable force field. This suggests that the current FQ parameterization is just slightly less accurate than the CHARMM22 model over the range of chemical

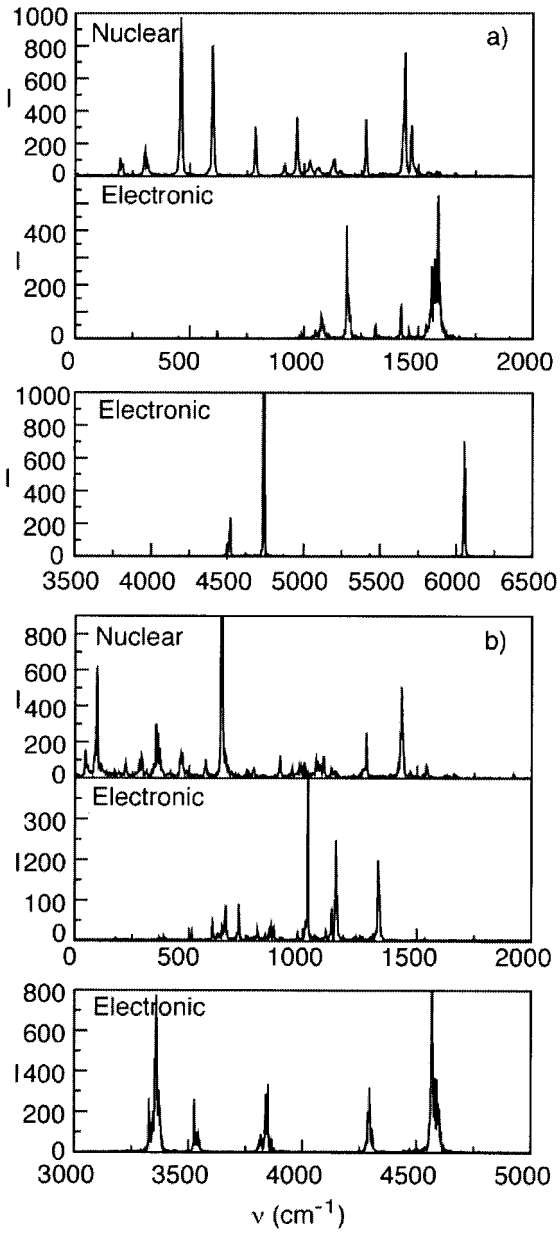

Figure 2. Power spectra for nuclear and electronic degrees of freedom (a) of *N*-methylacetamide (NMA), and (b) the alanine dipeptides from short a vacuum simulations.

types studied. We mention that the range of absolute errors for the FQ model are between 0.22 and 12.4% (0.22 and 6.0 % excluding the methyl formate results) compared to the CHARMM22 range of 0.0 to 6.89%. It is important to keep in mind that the relatively low errors for the alkanes in the CHARMM22 case are in large part due to the fact that these results are from a refinement of the original parameterization considering only four alkanes—there was no effort made to ensure that the alkane parameters were adequate to reproduce properties of higher functional groups to high accuracy, as considered in this work. Finally, we note that the current parameterization performs comparably in reproducing bulk liquid energies (as well as gas-phase solute–solvent energies) over a wide spectrum of chemical types. Vaporization enthalpies are within 1 kcal/mol, with some extreme cases in error by about 2 kcal/mol. In terms of parametrization, we note that there is a general trend towards a reduction in the contribution of the dispersion interactions associated with the fluctuating charge implementation. Before moving on to a discussion of other properties, we note that our current objective is to generate a *protein* force field consistent with CHARMM, that is, with the constraint of utilizing only the CHARMM-defined atom types. This being the case, the expectation is to achieve performance with regard to small molecule bulk liquid densities and energies commensurate to that of CHARMM22. In fact, if so desired, quite accurate small-molecule

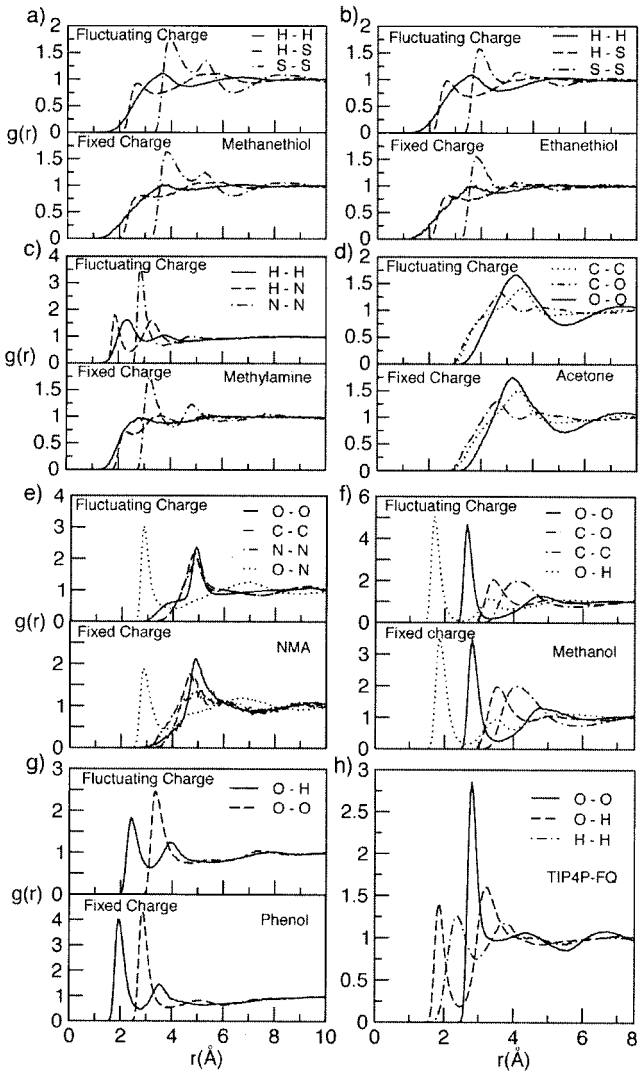

Figure 3. Comparison of radial distribution functions from FQ and fixed-charge CHARMM22 force fields for (a) methanethiol, (b) ethanethiol, (c) methylamine, (d) acetone, (e) *N*-Methylacetamide, (f) methanol, (g) phenol, and (h) TIP4P-FQ water (data shown for reference).

Table 6. Fluctuating Charge (FQ) and CHARMM22 (C22) Nonbonded Parameters.

Atom type	R^{min}/ε (FQ)[a]	R^{min}/ε (C22)	Atom type	R^{min}/ε (FQ)	R^{min}/ε (C22)
CT1 (aliphatic sp3 carbon, CH)	2.275/0.02	2.275/0.02	O	1.9/0.11	1.7/0.12
CT2 (aliphatic sp3 carbon, CH2)	2.01/0.06	2.01/0.056	OB	1.95/0.24	1.7/0.12
CT3 (aliphatic sp3 carbon, CH3)	2.02/0.073	2.04/0.078	OC	1.95/0.24	1.7/0.12
HA1 (hydrogen attached to CT1)	1.32/0.02	1.32/0.022	OH1 (aliphatic)	1.73/0.1521	1.77/0.1521
			OH1 (aromatic)	1.95/0.3771	—
HA2 (hydrogen attached to CT2)	1.34/0.032	1.34/0.028	S (sulphur)	2.1/0.47	2.0/0.45
HA3 (hydrogen attached to CT3)	1.32/0.023	1.34/0.024	HS (thiol hydrogen)	0.45/0.12	0.45/0.1
HR1	0.9/0.046	0.90/0.046	CD	1.8/0.2	2.0/0.07
HR3	1.468/0.0078	1.468/0.0078	CC	1.8/0.17	2.0/0.07
H	0.2245/0.046	0.2245/0.046	NH1	1.85/0.3	1.85/0.2
C (carbonyl C, peptide backbone)	2.0/0.11	2.0/0.11	NH2	1.77/0.15	1.85/0.2
CPH1 (Histidine)	2.2/0.05	1.8/0.05	NR1	1.95/0.35	1.85/0.2
CPH2 (Histidine)	2.2/0.05	1.8/0.05	NR2	1.85/0.35	1.85/0.2
CA (aromatic ring carbon, sp2)	1.95/0.07	1.9924/0.07	CPT	1.87/0.3	1.8/0.09
CY	1.9924/0.35	1.9924/0.07	NY	1.9/0.3	1.85/0.2
HP (aromatic ring hydrogen)	1.3082/0.03	1.3582/0.03			
OS	1.1/0.002	1.7682/0.1521	OT (σ/ε)	3.159/0.2862	
			HT (σ/ε)	—	

[a] R^{min} in Å and ε in kcal/mol. TIP4P-FQ values shown for reference, radius is σ in Å as opposed to R^{min} for the CHARMM implementation.

properties can be achieved by lifting this constraint as evidenced by literature reports on studies of *individual* molecules or classes of molecules.

The extent of polarization of the liquid relative to the vacuum is addressed via condensed phase dipole moment distributions. Figure 4 shows a set of molecular dipole moment distributions for the test liquids. As in the case for the SPC-FQ and TIP4P-FQ models, there is a nontrivial polarization, of varying extent, for each liquid. At this time, there is no firm experimental evidence for the magnitudes of liquid dipole moments of the liquids tested; in fact, there still appears to be controversy as to the correct value of the liquid state dipole moment of water. Interestingly, in the case of water, there is evidence indicating that a simulation will predict the correct value for the static dielectric constant if the force field used allows a liquid-state dipole moment averaging around 2.6 Debye[17] (see Fig. 4k; this TIP4P-FQ model predicts a static dielectric constant of 79). In light of this observation, the current model allows for an increase in the molecular dipole moment relative to the gas phase for all the polar molecules. Furthermore, the polarization shift is much smaller, on the order of 0.2 Debye for the nonpolar aliphatics tested—propane and butane (Fig. 4h and j)—as opposed to up to as much as 1 Debye in the case of methanol (Fig. 4b), for instance. Although not done here, it would be interesting to note how the shift in the distribution maximum correlates with the molecular polarizability, which is determined in this model solely by the hardness values. This polarization effect has also been observed in earlier simulations of non-aqueous systems (i.e., chlorinated alkanes), and appears to be a general property of models of this type.[29]

One can address the idea of enhanced interactions leading to more structure in the condensed phase by considering the distribution of pair energies in the liquid. Figure 5a shows the total pair energy distribution for methanol liquid computed from the fluctuating charge and CHARMM22 nonpolarizable force field simulations. The enhanced pair interaction for the polarizable liquid is quite evident in the peak around −8.0 kcal/mol compared to that for the CHARMM22 distribution at −6.0 kcal/mol. Figure 5b shows the electrostatic component of the total energy, and indeed, much of the interaction arises from the electrostatics; this behavior suggests the need to at some point determine the accuracy with which the condensed phase polarization is captured by the force field.

The accuracy of the computed self-diffusion coefficients shown in Table 7 varies considerably from liquid to liquid. There does not appear to be a systematic trend towards higher or lower values. The values for methanol and NMA are comparable in accuracy to those predicted using the point dipole model of Kollman and Cieplak.[13] Keeping in mind that diffusion constants were not included in the optimization scheme, one should not expect to achieve too high an accuracy. All in all, the computed values are acceptable, at least relative to current fixed charge potentials.

Towards Solvated Peptide Systems: N-*Methylacetamide in Water*

As a simple test of the FQ force field for solvated systems, we consider a dilute solution of *N*-methylacetamide in water. We perform a short (500 ps) molecular dynamics simulation of a single NMA molecule solvated in a box of 216 TIP4P-FQ water molecules. The computations are performed at constant temperature and pressure as previously described for the bulk liquid simulations. The charges are thermostatted with a Hoover bath; in this case, we find that coupling the solvent and solute electronic degrees of freedom to independent baths allows a more rigorous temperature control for the solute charges. Figure 6 shows radial distribution functions between the water and NMA, indicating the nature of

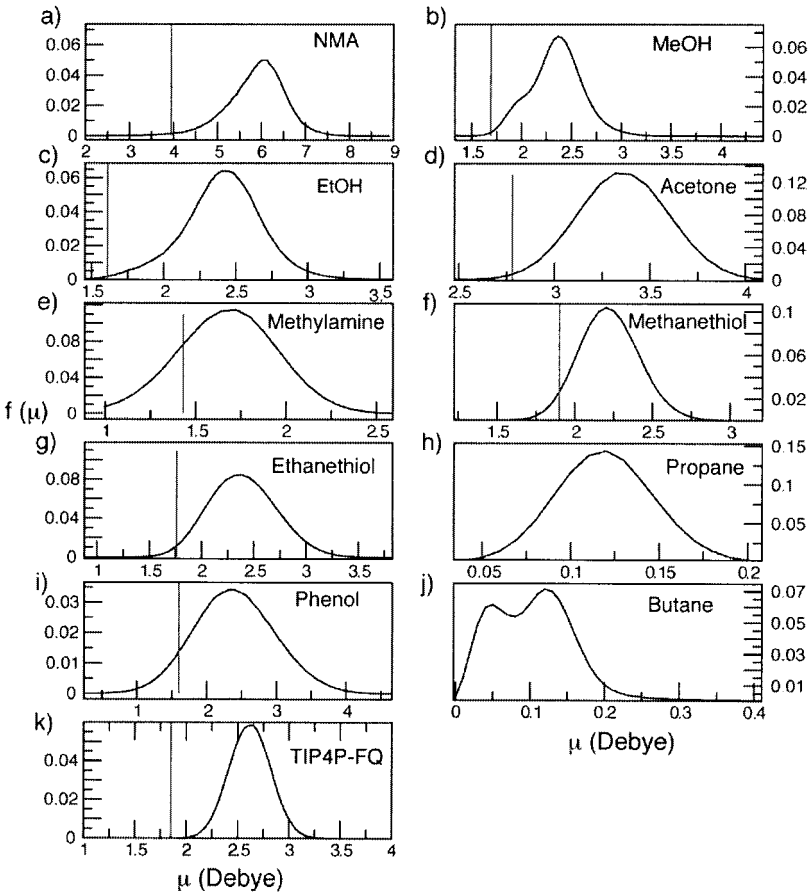

Figure 4. Dipole moment distributions for bulk liquids. The impulse represents the force field vacuum dipole moment. Dipole moments in units of Debye.

solvent structure around the amide. These are quite similar to those computed by Rick et al.[25] using the rigid molecule model with constant intramolecular shielded interactions based on overlaps of atom-centered Slater orbitals. Figure 6a indicates hydrogen bonding between the carboxyl oxygen and water, and as observed by Rick et al., the amide–water interaction is stronger than the water–water hydrogen-bonded interaction as suggested by the higher peak heights in the carbonyl oxygen–water distribution functions. Our model, however, seems to indicate a stronger amide hydrogen–water interaction compared to that of Rick et al.; this is likely due to a higher induced polarization using the current parameterization, which we discuss next.

Condensed-phase effects were again probed by considering the polarization of molecules. Figure 7 shows a trajectory profile of the instantaneous dipole moment of NMA. First, there is a clearly enhanced dipole moment relative to vacuum and the bulk neat liquid, averaging around 7.2 Debye. This prediction significantly differs from values computed from electronic structure calculations using continuum dielectric models and an *ab initio* study using explicit TIP3P water, the former giving 5.33 Debye[78] and the later 5.18 Debye.[79] This suggests the need to perhaps include solvation energies in the fitting procedure at some stage. Of par-

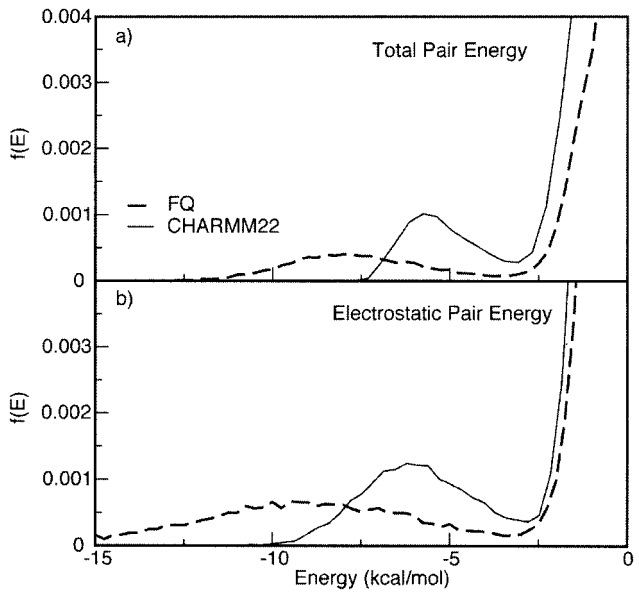

Figure 5. Pair energy distributions for methanol. The upper figure shows the distributions of the total interaction energy, the lower figure shows the electrostatic component.

Table 7. Self-Diffusion Coefficients.

	Diffusion coefficient (10^{-5} cm^2/s)		
Liquid	FQ	C22	Experimental
N-Methylacetamide	1.93 (0.09)	2.04	1.2
Methanol	2.01 (0.047)	1.99	2.2
Ethanol	1.66 (0.11)	0.82	1.2
Formamide	0.71 (0.07)	0.98	
Methylamine	3.99 (0.13)	3.43	
Acetone	3.31 (0.07)	2.14	
Methanethiol	2.91 (0.13)	3.85	
Ethanethiol	2.75 (0.098)	2.63	
Benzene	2.41 (0.26)	1.75	
Methylformate	4.32 (0.18)	2.487	
Phenol	1.18 (0.01)	0.72	
TIP4P-FQ	1.43	1.9[a]	2.3

[a]The TIP4P-FQ result is for constant pressure and temperature at 298 K and average pressure of 0.017 kbar. This pressure is higher than the reported NVT value of −0.16 kbar,[24] and the self-diffusion constant is correspondingly lower due to the higher density.

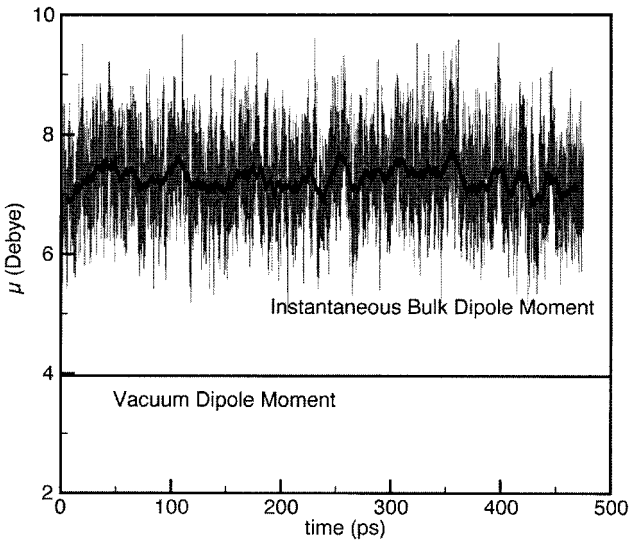

Figure 7. Trajectory of instantaneous dipole moment for NMA in water.

ticular interest in the application of polarizable models to peptide/protein simulations is the polarization of the solvent in the vicinity of the solute. That structural and dynamical properties of vicinal water are different from those of the bulk is well documented. Here, we show the polarization response for vicinal and bulk water molecules. Figure 8 shows the molecular dipole moment distributions of water molecules in a shell between 1.0 and 2.9 Å from the center of mass of the solute and in the bulk. As one would anticipate, the molecular dipole moments in the inner shell are shifted by about 0.1–0.2 Debye to higher values. Rick et al. reported a higher average dipole moment of 2.72 Debye for water molecules nearest to the carbonyl oxygen relative to the bulk value of 2.62, a shift that is similar in magnitude to that observed in the present simulations. It is this mutual enhancement of dipoles that leads to the strong amide–water interaction as evidenced in the relevant radial distribution functions (Fig. 6). We note that Rick et al. report a suppressed dipole moment for water molecules in the vicinity of atoms other than the carbonyl, consistent with the less structured solvent distribution around the methyl and amide hy-

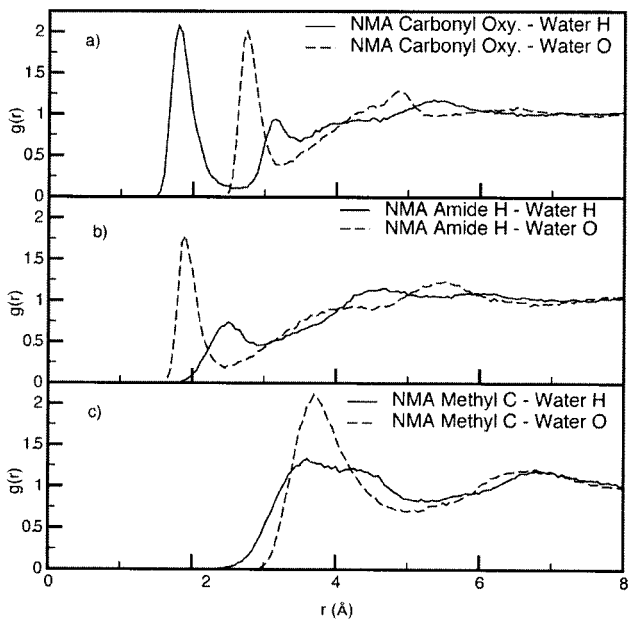

Figure 6. NMA–water radial distribution functions.

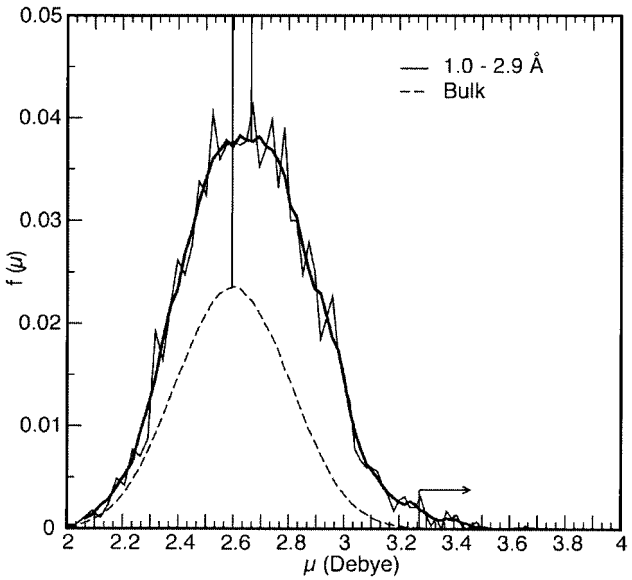

Figure 8. TIP4P-FQ water molecular dipole moment distributions. Solid curve represents distribution for molecules near the solute, and broken curve that for bulk-like water.

drogen atoms. The present results seem to indicate that there is dipole enhancement for all waters near the solute as evidenced by the amide hydrogen interaction with water.

Conclusions

This contribution presents the results of our current and ongoing efforts towards developing and implementing a fluctuating charge force field for use with the CHARMM molecular modeling package.[80] The fluctuating charge approach has been applied widely, and is no longer a novelty in some respects. Our results show, interestingly, that differing approaches to parametrizing the model yield similar results at a qualitative level. For instance, our results for bulk liquid NMA (structure and thermodynamics) are similar to those of Caldwell and Kollmann[13] as well as Rick and Berne.[25] Furthermore, the approach can be applied to a spectrum of chemical environments such as fairly nonpolar aliphatics (alkanes) to highly polar systems such as water, methanol, formamide, and the like; in this respect, this is moving towards the goal of a transferable polarizable force field as opposed to earlier polarizable models developed for specific molecular systems or functional groups. The model is balanced with respect to solute–water and solute–solute interactions, which are important in modeling solvated biomacromolecular systems. The present model also captures the condensed phase polarization, as evidenced by the shift in dipole moment distributions of Figure 4, although the accuracy of such predicted shifts is difficult to assess at this time. The parameterization also appears to be transferable, as evidenced by the prediction of vacuum dipole moments of small molecules distinct from the fitting set. We are currently applying the force field to simulations of small proteins in explicit solvent to understand the differences nonadditivity introduced in this manner leads to; this will be reported in ensuing publications.

Acknowledgments

We would like to thank Professor Alex MacKerrel for his inspirational prodding and the expertise in force field development he shared with us during his sabbatical stay at Scripps. An additional debt of gratitude is owed to Drs. Joe Zhou and Tom Cleveland for their early efforts in helping define the approach used in this work.

References

1. Karplus, M.; McCammon, J. A. Nat Struct Biol 2002, 9, 646.
2. McCammon, J. A.; Harvey, S. C. Dynamics of Proteins and Nucleic Acids; Cambridge University Press: Cambridge, UK, 1987.
3. Brooks, C. L., III; Karplus, M.; Pettitt, B. M. Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics; John Wiley & Sons: New York, 1988, vol. LXXI.
4. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. J Phys Chem 1993, 97, 10269.
5. Chirlian, L. E.; Francl, M. M. J Comp Chem 1987, 8, 894.
6. Breneman, C. M.; Wiberg, K. B. J Comp Chem 1990, 11, 361.
7. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. J Am Chem Soc 1993, 115, 9620.
8. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. J Am Chem Soc 1995, 117, 5179.
9. Foresman, J. B.; Brooks, C. L., III. J Chem Phys 1987, 87, 5892.
10. Dang, L. X. J Chem Phys 1992, 97, 2659.
11. Jungwirth, P.; Tobias, D. J. J Phys Chem B 2002, 106, 6361.
12. Dang, L. X. J Phys Chem B 2001, 105, 804.
13. Caldwell, J. W.; Kollman, P. A. J Phys Chem 1995, 99, 6208.
14. Dang, L. X.; Chang, T.-M. J Chem Phys 1997, 106, 8149.
15. Chang, T.-M.; Dang, L. X.; Peterson, K. A. J Phys Chem 1997, 101, 3413.
16. Mannfors, B.; Mirkin, N. G.; Palmo, K.; Krimm, S. J Comp Chem 2001, 22, 1933.
17. Sprik, M. J Chem Phys 1991, 95, 6762.
18. Sprik, M.; Klein, M. L. J Chem Phys 1988, 89, 7556.
19. Sprik, M.; Klein, M. L.; Watanabe, K. J Phys Chem 1990, 94, 6483.
20. MacKerell, A. D., Jr., Personal Communication.
21. Ando, K. J Chem Phys 2001, 115, 5228.
22. Krishnan, M.; Verma, A.; Balasubramanian, S. Proc Indian Acad Sci (Chem Sci) 2001, 113, 579.
23. Chen, B.; Xing, J.; Siepmann, J. I. J Phys Chem 2000, 104, 2391.
24. Rick, S. W.; Stuart, S. J.; Berne, B. J. J Chem Phys 1994, 101, 6141.
25. Rick, S. W.; Berne, B. J. J Am Chem Soc 1996, 118, 672.
26. Stuart, S. J.; Berne, B. J. J Phys Chem 1996, 100, 11934.
27. Rick, S. W.; Stuart, S. J.; Bader, J. S.; Berne, B. J. J Mol Liq 1995, 65/66, 31.
28. Ribeiro, M. C. C.; Almeida, L. C. J. J Chem Phys 1999, 110, 11445.
29. Llanta, E.; Ando, K.; Rey, R. J Phys Chem B 2001, 105, 7783.
30. Rick, S. W.; Berne, B. J. J Phys Chem B 1997, 101, 10488.
31. Yoshii, N.; Miura, S.; Okazaki, S. Chem Phys Lett 2001, 345, 195.
32. Yoshii, N.; Yoshie, H.; Miura, S.; Okazaki, S. J Chem Phys 1998, 109, 4873.
33. Rick, S. W. J Chem Phys 2001, 114, 2276.
34. Stuart, S. J.; Berne, B. J. J Phys Chem 1999, 103, 10300.
35. Bryce, R. A.; Vincent, M. A.; Malcolm, N. O. J.; Hillier, I. H.; Burton, N. A. J Chem Phys 1998, 109, 3077.
36. Rappe, A. K.; Goddard, W. A., III. J Phys Chem 1991, 95, 3358.
37. Stern, H. A.; Rittner, F.; Berne, B. J.; Friesner, R. J Chem Phys 2001, 115, 2237.
38. Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X.; Murphy, R. B.; Zhou, R.; Halgren, T. A. J Comp Chem 2002, 23, 1515.
39. Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B. J.; Friesner, R. A. J Phys Chem B 1999, 103, 4730.
40. Yoshii, N.; Miyauchi, R.; Miura, S.; Okazaki, S. Chem Phys Lett 2000, 317, 414.
41. Chen, B.; Siepman, J. I. Theo Chem Accts 1999.
42. Banks, J. L.; Kaminski, G. A.; Zhou, R.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. J Chem Phys 1999, 110, 741.
43. Chelli, R.; Procacci, P. J Chem Phys 2002, 117, 9175.
44. Parr, R. G.; Yang, W. Density-Functional Theory of Atoms and Molecules; Oxford University Press: Oxford, 1989.
45. Sanderson, R. T. Science 1951, 114, 670.
46. Sanderson, R. T. Chemical Bonds and Bond Energy; Academic: New York, 1976, 2nd ed.
47. Nalewajski, R. F.; Korchowiec, J.; Zhou, Z. Int J Quantum Chem Quantum Chem Symp 1988, 22, 349.
48. Nose, S. Mol Phys 1984, 52, 255.
49. Andersen, H. C. J Chem Phys 1980, 72, 2384.
50. Parrinello, M.; Rahman, A. Phys Rev Lett 1980, 45, 1196.
51. Car, R.; Parrinello, M. Phys Rev Lett 1985, 55, 2471.
52. Pastore, G.; Smargiassi, E.; Buda, F. Phys Rev A 1991, 44, 6334.

53. Liu, Y.-P.; Kim, K.; Berne, B. J.; Friesner, R. A.; Rick, S. W. J Chem Phys 1998, 108, 4739.
54. Rablen, P. R.; Lockman, J. W.; Jorgensen, W. L. J Phys Chem A 1998, 102, 3782.
55. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr. ; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head–Gordon, M.; Replogle, E. S.; Pople, J. A.; Revision A. H. Gaussian; Gaussian, Inc.: Pittsburgh, PA, 2001, 1st ed.
56. MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; 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., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz–Kuczera, J.; Yin, D.; Karplus, M. J Phys Chem B 1998, 102, 3586.
57. Nakano, T.; Kaminuma, T.; Uebayasi, M.; Nakata, Y. Chem-Bio Inform J 2001, 1, 35.
58. Yin, D.; MacKerell, A. D., Jr. J Comp Chem 1998, 19, 334.
59. Jorgensen, W. L.; Maxwell, D. S.; Tirado–Rives, J. J Am Chem Soc 1996, 118, 11225.
60. Jorgensen, W. L.; Tirado–Rives, J. J Am Chem Soc 1988, 1657.
61. Jorgensen, W. L., Briggs, J. M.; Contreras, L. J Phys Chem 1990, 1683.
62. Allen, M. P.; Tildesley, D. J. Computer Simulation of Liquids; Clarendon Press: Oxford, 1987.
63. Martyna, G. J.; Klein, M. L.; Tuckerman, M. J Chem Phys 1992, 97, 2635.
64. Darden, T.; York, D.; Pederson, L. J Chem Phys 1993, 98, 10089.
65. Darden, T.; Essman, U.; Lee, H.; Perera, L.; Berkowitz, M. L.; Pederson, L. J Chem Phys 1995, 103, 8577.
66. Rick, S. W. In Simulation and Theory of Electrostatic Interactions in Solution: Computational Chemistry, Biophysics, and Aqueous Solutions; Pratt, L. R., Hummer, G., Eds.; American Institute of Physics: Santa Fe, NM, 1999.
67. Nymand, T. M.; Linse, P. J Chem Phys 2000, 112, 6386.
68. Roychowdhury, P.; Basak, B. S. Acta Crystallogr B 1975, 31, 1559.
69. Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R. Acta Crystallogr B 2002, 58, 389.
70. Tuckerman, M. E.; Parrinello, M. J Chem Phys 1994, 101, 1302.
71. Tuckerman, M. E.; Parrinello, M. J Chem Phys 1994, 101, 1316.
72. Izaguirre, J. A.; Reich, S.; Skeel, R. D. J Chem Phys 1999, 110, 9853.
73. Tuckerman, M. E.; Berne, B. J. J Chem Phys 1990, 94, 1465.
74. Blochl, P. E.; Parrinello, M. Phys Rev B 1992, 45, 9413.
75. Blochl, P. E. Phys Rev B 2001, 65, 104303.
76. Delhommelle, J.; Tschirwitz; C.; Ungerer, P.; Granucci, G.; Millie, P.; Pattou, D.; Fuchs, A. H. J Phys Chem B 2000, 104, 4745.
77. Jorgensen, W. L.; Maxwell, D. S.; Tirado–Rives, J. J Am Chem Soc 1996, 118, 11225.
78. Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M.; Goddard, W. A., III; Honig, B. J Am Chem Soc 1994, 116, 11875.
79. Gao, J.; Xia, X. Science 1992, 258, 631.
80. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J Comp Chem 1983, 4, 187.
81. Miller, K. J. J Am Chem Soc 1990, 112, 8533.