JCTC Journal of Chemical Theory and Computation

Development of a Polarizable Force Field Using Multiple Fluctuating Charges per Atom

Dong-Xia Zhao,[†] Cui Liu,[†] Fang-Fang Wang,[†] Chun-Yang Yu,[†] Li-Dong Gong,[†] Shu-Bin Liu,[‡] and Zhong-Zhi Yang*,[†]

Chemistry and Chemical Engineering Faculty, Liaoning Normal University, Dalian, 116029, China and Research Computing Center, University of North Carolina, Chapel Hill, North Carolina 27599-3420

Received December 10, 2009

Abstract: A polarizable force field (PFF) using multiple fluctuating charges per atom, ABEEM$\sigma\pi$ PFF, is presented in this work. The fluctuating partial charges are obtained from the electronegativity equalization principle applied to the decomposition scheme of atom-bond regions into multiple charge sites: atomic, lone-pair electron, and σ and π bond regions. These multiple partial charges per atom should better account for the polarization effect than single atomic charge in other PFFs. To evaluate the PFF, structural and energetic properties for some organic and biochemical systems, including rotational barriers; binding energies of base pairs; a base−base interaction in a B-DNA decamer; and interaction energies of ten stationary conformers of a water dimer, peptides, and bases with water molecules, have been calculated and compared with the experimental data or *ab initio* MP2 results. Molecular dynamics simulations using the PFF have been performed for crambin and BPTI protein systems. Better performances in modeling root-mean-square deviations of backbone bond lengths, bond angles, key dihedral angles, the coordinate root-mean-square shift of atoms, and the distribution of hydrogen bonds have been observed in comparison with other PFFs. These results indicate that the fluctuating charge force field, ABEEM$\sigma\pi$/MM, is accurate and reliable and can be applied to wide ranges of organic and biomolecular systems.

Introduction

Molecular dynamics (MD) simulations using force fields (FF) are still an important tool in understanding structure, dynamics, and function properties for biological systems. Although quantum calculations and *ab initio* MD modeling have tremendously advanced in recent decades, they are limited to relatively small systems.[1,2] In a primitive force field, molecules are represented by a collection of atom-centered interaction sites with fixed partial charges. The electrostatic energy is determined by the Coulombic interaction between the partial charges without any inclusion of the polarization effect. In view of the significance of the polarization effect, much has been achieved since the 1970s in the development and employment of the polarizable force field (PFF) for biomolecule simulation, such as CHARMM,[3−8] OPLS,[9−16] AMBER,[17−20] NEMO,[21] AMOEBA,[22,23] QMPFF,[24−26] and so forth. The first physically consistent microscopic study of dielectric effects in nonpolar environments was reported by Warshel and Levitt.[27] Accounts and reviews of PFF are available elsewhere.[27,28] Two approaches have been employed to account for the polarization effect. One is through the induced dipole or multipole (including Drude charge oscillator and induced dipole mixed fluctuating charge models), and the other is to use fluctuating partial charge.

In the induced dipole (multipole) PFF, the potential energy function is augmented by an inductive term from the induced dipole. The contribution of the electrostatic interaction comes from both permanent charges and induced dipole moments obtained from the atomic polarizabilities through an iterative procedure.[29] Karlström et al.[21] have demonstrated the

* Corresponding author e-mail: zzyang@lnnu.edu.cn.
[†] Liaoning Normal University.
[‡] University of North Carolina.

10.1021/ct9006647
Published on Web 02/04/2010

importance of the quadrupole moment. Higher order multipoles were included in the AMOEBA PFF.[22] A new PFF of this category, the X-POL potential,[30] was recently proposed by Xie and co-workers. The published CHARMM all-atom parameters for nucleic acids[4] and proteins[5] provide a consistent set for condensed-phase simulations of a wide variety of biological molecules. Xie et al. developed a polarizable intermolecular function for liquid amides and alkanes by employing the "standard" CHARMM force field plus a polarizable term.[8] The OPLS series of force fields have been developed for more than 20 years and have proved to be highly successful in computing liquid state thermodynamic properties[15] and in dipeptide, protein, and protein−ligand modeling.[11−13] The AMBER force field has simulated the structures, conformational energies, and interaction energies of proteins, nucleic acids, and many related organic molecules in condensed phases.[18] Its point-charge all-atom force field for proteins and united-atom force field for simulations involving highly demanding conformational sampling such as protein folding and protein−protein binding have been advanced.[20] The QMPFF[24−26] was fitted solely to QM data at the MP2/aTZ(-hp) level and had demonstrated high accuracy and transferability in crystal and liquid simulations as a result of its strong physical basis and explicit polarizability. As is well-known, there are many other well-developed useful force fields, such as MM1-MM4,[31] MMFF,[32] ECEPP,[33] CFF,[34] Tripos,[35] GROMOS,[36] etc.

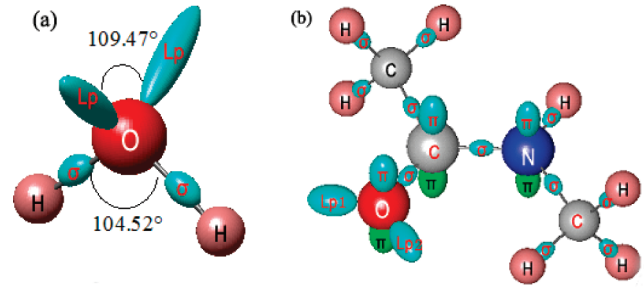
In fluctuating charge models, whose basis is the electronegativity equalization method (EEM) in density functional theory (DFT),[37,38] atomic partial charges of a molecular system are allowed to change with geometry and ambient environment. In the usual EEM scheme, using the atomic partial charge and two characteristic parameters per atom, the electrostatic energy is written as

$$E_{es} = \sum_a (\chi_a^* q_a + \eta_a^* q_a^2) + \sum_{a<b} k_{ab} \frac{q_a q_b}{R_{ab}} \qquad (1)$$

where q_a is the atomic partial charge at the a region (atom and/or specified one), χ_a^* and η_a^* are the valence-state electronegativity and valence-state hardness of region a, R_{ab} is the separation between regions a and b, and the summation for both a and b is over all sites. In DFT, the effective electronegativity χ_a of a site a is equal to the partial derivative of the electrostatic energy with respect to the partial charge q_a of site a. k_{ab} is a correction factor of the Coulombic interaction energy between the partial charges q_a and q_b, which stems from the reality that q_a and q_b involve the electron clouds rather than the ideal point charges.[39−42] According to the electronegativity equalization principle based on DFT, at the equilibrium state, the effective electronegativities of all sites are equal to the global molecular electronegativity χ_{mol}, which constitutes the electronegativity equalization equation. By solving these equations with the charge constraint for a molecular system, the partial charges of all sites in the system are directly and quickly obtained.

There are several fine implementations of EEM to allow rapid calculations of the partial charge distribution in molecules. Mortier et al. proposed a systematic formulation[43] to directly calculate atomic partial charges of a large molecular system. York and Yang[44] established a chemical potential equalization model. Rappé and Goddard presented the charge equilibration method[39] for MD simulations. Stern et al. formulated a combined fluctuating charge and polarizable dipole force field for water, amino acids, and peptides based on *ab initio* data.[45] Recently, Chelli and Procacci developed a transferable polarizable electrostatic force field[46] by generalizing Mortier's method to include atom-based dipolar distributions as done in the chemical potential equalization model by York and Yang. A CHARMM fluctuating charge force field[7] for proteins has been applied to MD simulations for bulk organic liquid and small proteins. Notice that, in all the existing methods, only one partial charge per atomic site is employed.

Chart 1. The Sketch of All Regions in (a) Water and (b) NMA Molecules Defined in the ABEEM$\sigma\pi$ Model



As recently pointed out by Jorgensen, broadly applicable PFFs have yet to emerge.[47] The development of PFFs remains to be a frontier challenge in molecular modeling. Key to the development of a reliable PFF is to effectively account for the polarization effect. In this work, we present a PFF with multiple fluctuating partial charges per atom, ABEEM$\sigma\pi$/MM, based on the atom-bond electronegativity equalization method (ABEEM) that we have recently developed.[48−54]

Model

ABEEM$\sigma\pi$ with Multiple Partial Charges per Atom. Conventional EEM models partition a molecular system into individual atomic regions, each having one partial charge only. As is well-known, besides nuclei, electron densities also concentrate around chemical bonds and lone-pair regions to some extent. Henceforth, to better account for the polarization effect, we suggested that a molecular system is decomposed into atomic regions, lone-pair regions, and σ and π bond regions, and each region is assigned by a partial charge.[46,49,51] Better elucidation of the polarization effect from this partition scheme comes from the greater freedom and larger flexibility in calculating the fluctuating partial charge associated with each of the regions. In principle, if the total number of charge sites approaches infinity, the partial charge q_i will resemble the continuous charge density, the fundamental variable of DFT. This is the essence and foundation of our ABEEM$\sigma\pi$ model.

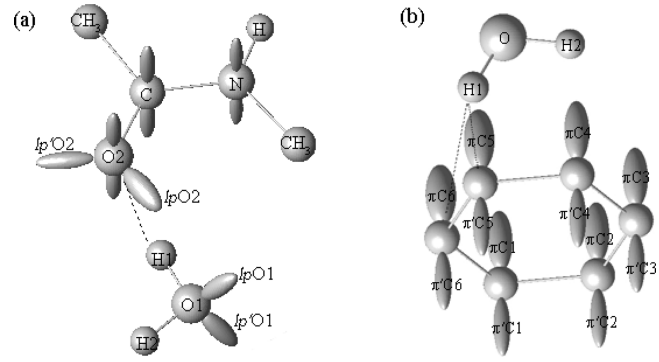
In Chart 1, as examples, we draw the charts of all regions in water and N-methylacetamide (NMA) molecules in the

ABEEM$\sigma\pi$ model. In addition to atomic sites, some new virtual sites including lone pair and σ and π bond sites are placed according to the physical meaning. A water molecule contains seven electron cloud regions in the ABEEM$\sigma\pi$ model: three atoms, two σ bonds whose angle is 104.5°, and two lone pairs whose angle is 109.47°. As is known, there are four electron pairs around the oxygen atom, which are spread so as to point roughly toward the apexes of a tetrahedron. Every atomic charge is placed in the position of the corresponding atom. The σ bond charge is assumed to locate on the point that partitions the bond length according to the ratio of covalent radii of two bonded atoms, and the lone pair sites are placed on the points which are 0.74 Å from the oxygen nucleus with an intervening angle of 109.47° between two lone pairs on oxygen atom. Thus, there are seven partial charge sites for a water molecule in the ABEEM$\sigma\pi$ model.

Consider the NMA molecule, as shown in Chart 1b. The C atom of carbonyl connects two single bonds and one double bond. The geometry around this C atom is trigonal planar. The geometry around the O atom is trigonal planar too, because the O atom of carbonyl connects one double bond and two lone pairs. All the angles between the two lone pairs and C═O bond are 120°. The oxygen atom involves six partial charges, namely, one centered on the oxygen nucleus, one σ region, two π separate upper and lower regions, and two lone pairs. The σ bond partial charge shared by the oxygen and carbon atoms is on the bond at the point that partitions the bond length according to the ratio of covalent radii between O and C atoms; the π bond partial charges are placed above and below the O atom at the covalent radius (0.74 Å) of the O atom perpendicular to the plane formed by the σ bonds and may have different values depending on the environment; the two lone pair partial charges are placed in the covalent radius of the oxygen atom (0.74 Å). An electron pair of the N atom can be used to make a delocalized π bond with carbonyl. There are also similar π bond partial charges for the nitrogen atom and carbon atom of carbonyl in NMA. So the nitrogen atom involves six partial charges, including one atom and three σ and two π regions. As a whole, besides 12 atomic sites, a NMA molecule has an additional 19 sites: 11 σ bond sites, 6 π bond sites, and 2 lone pair sites.

For a molecule, the electrostatic energy is written as eq 1 in the ABEEM$\sigma\pi$ model which contains the regions or sites of the atoms, bonds, and lone pairs, as shown in Chart 1. On the basis of DFT, the effective electronegativity χ_a of every site a can be also expressed as the partial derivative of the electrostatic energy with respect to the partial charge q_a of site a. According to the EEM, at the equilibrium state the effective electronegativities of all sites, including atoms, bonds, and lone pairs are equal to the global molecular electronegativity χ_{mol} for every molecule, which constitutes the electronegativity equalization equation, i.e., $\chi_a = \chi_b = \cdots = \chi_{\mathrm{mol}}$. It can be shown that the number of the equations of EEM is equal to the number of the sites of molecules. These equations, together with the charge constraint and given parameters (valence-state electronegativity χ_a^* and valence-state hardness η_a^* of region a), can be explicitly and quickly solved to give the global molecular electronegativity χ_{mol} and the partial charge q_i on each site i. The detailed formulation of ABEEM$\sigma\pi$ is available in the Supporting Information (SI1).

Chart 2. Sketch of (a) the Cluster NMA+H_2O and (b) the Cluster Benzene+H_2O, Where Hydrogen Atoms in Aromatic Ring and σ Bond Sites Are Omitted



In NMA, the atomic partial charge plus four or five negative partial charges around a non-hydrogen atom, like the C, N, or O atom, may fluctuate, giving a proper response to the geometry and the environmental change. For example, charges of *lp*O2 and *lp′*O2 of NMA are −0.1689 and −0.1639, respectively. In cluster NMA+H_2O (Chart 2a), the charges located on *lp* of O2 all become more negative than that on *lp* electrons of O2 in NMA due to the intermolecular HB. Because the H1 points to the *lp*O2, the charge located on *lp*O2 (−0.1748) is more negative than that on *lp′*O2 (−0.1691). All the π bond charges of all carbon atoms in benzene are −0.0135. In cluster benzene+H_2O, as shown in Chart 2b, the largest polarized regions are πC5 and πC6 partial charge regions, whose charges are −0.0141. All the charges on πC1∼πC4 partial charge regions, which are slightly polarized, are −0.0139; the charges on π'C1∼π'C6 partial charge regions are all −0.0134.

ABEEM$\sigma\pi$ Polarizable Force Field, ABEEM$\sigma\pi$/MM. The energy function $E_{\mathrm{ABEEM}\sigma\pi}$ of the ABEEM$\sigma\pi$ polarizable force field can be written as the sum of following terms:

$$E_{\mathrm{ABEEM}\sigma\pi} = E_{\mathrm{b}} + E_\theta + E_\phi + E_{\mathrm{imptors}} + E_{\mathrm{vdw}} + E_{\mathrm{elec}} \tag{2}$$

where E_{b} and E_θ are the usual energy terms from bond stretching and angle bending contributions modeled as harmonic oscillators (or Morse function for water), respectively,

$$E_{\mathrm{b}}(r) = \sum_{\mathrm{bonds}} k_r(r - r_{\mathrm{eq}})^2 \tag{3}$$

$$E_\theta(\theta) = \sum_{\mathrm{angles}} k_\theta(\theta - \theta_{\mathrm{eq}})^2 \tag{4}$$

where k_r and k_θ represent the force constants of the stretching and bending, r and θ are the actual values of the bond length and bond angle, and r_{eq} and θ_{eq} denote the equilibrium values of the bond length and bond angle, respectively. E_ϕ is the torsional energy for a bond rotation. The torsional term takes the following form:

$$E_{\phi}(\phi) = \sum_{\text{torsions}} \left[\frac{V_1}{2}(1 + \cos\phi) + \frac{V_2}{2}(1 - \cos 2\phi) + \frac{V_3}{2}(1 + \cos 3\phi) \right] \quad (5)$$

and the improper dihedral angle term is written as

$$E_{\text{imptors}} = \sum_{\text{imptors}} v(1 - \cos 2\phi) \quad (6)$$

where V_1, V_2, V_3, and v are the dihedral angle and improper dihedral angle force constants, respectively. E_{vdw} describes the van der Waals nonbonded atom−atom interaction:

$$E_{\text{vdw}} = \sum_{i<j} 4 f_{ij} \varepsilon_{ij} (\sigma_{ij}^{12}/r_{ij}^{12} - \sigma_{ij}^{6}/r_{ij}^{6}) \quad (7)$$

Geometric combining rules for the Lennard-Jones coefficients are employed: $\sigma_{ij} = (\sigma_{ii}\sigma_{jj})^{1/2}$ and $\varepsilon_{ij} = (\varepsilon_{ii}\varepsilon_{jj})^{1/2}$. The summation runs over all of the pairs of unbonded atoms i and j. If i and j are intramolecular, the coefficient $f_{ij} = 0.0$ for any $i-j$ pair connected by a valence bond (1−2 pairs) or a valence bond angle (1−3 pairs), $f_{ij} = 0.5$ for 1,4 interactions (atoms separated by three bonds), and $f_{ij} = 1.0$ for all other intramolecular and intermolecular cases.

The key difference between the ABEEM$\sigma\pi$ PFF and other force fields is the treatment of the electrostatic interaction energy. The electrostatic interaction energy, E_{elec}, is actually expressed as

$$E_{\text{elec}} = \sum_{i<j} k_{ij} q_i q_j / r_{ij} \quad (8)$$

where r_{ij} is the distance between sites i and j. q_i and q_j are the partial charges of regions or sites i and j, which are calculated by the ABEEM$\sigma\pi$ method. In reality, either q_i or q_j involves or represents some sort of electron clouds. Therefore, when we model them as point charges in EEM or ABEEM, their electrostatic interaction energy is expressed as $k_{ij}q_iq_j/r_{ij}$ rather than the pure Coulombic form q_iq_j/r_{ij}. The introduced parameter k_{ij} may be said to be a result of considering the exchange, penetration, and shielding effect in the interaction between the two pieces of electron clouds i and j. There is a similar parameter, the electrostatic interaction factor (modified Coulomb interaction), $J_{ij}(r_{ij})$, such as in the models of the charge equilibration of Rappé and Goddard,[39] Bakowies and Theil,[40] as well as Dias and co-workers,[41,42] which was taken into account in their electrostatic energy expressions and is related to the local hardnesses of atoms i and j.

In our ABEEM or MEEM model, k_{ij} is optimized and set to be 0.57 empirically, except for hydrogen-bond regions. In the hydrogen bond interaction region,[51,53] k_{ij} is replaced by a $k_{\text{HB}}(r_{ij})$ function to describe the electrostatic interaction between the hydrogen atom and the lone-pair electron.

In molecular simulations, we use the ABEEM$\sigma\pi$ method to calculate partial charges of all regions, namely, atoms, lone pairs, σ bonds, and π bonds, and then employ eq 2 to compute the total energy of the system. If there is a geometrical change, i.e., a change of bond length, angle, dihedral angle, or relative position between molecules, we will recalculate the partial charges using the ABEEM$\sigma\pi$ method, and then the total energy. In this manner, a systematic way to account for the polarization effect by allowing partial charge fluctuations in accordance with the changing molecular environment is provided by the ABEEM$\sigma\pi$ force field, termed in short as ABEEM$\sigma\pi$/MM.

ABEEM$\sigma\pi$/MM Parameters' Determination and Calibration. The systematic determination and calibration of parameters in developing any force field is tedious, involving tremendous amount of tasks in testing, calibrations, and analyses. In our case, more than 2000 organic and biological molecules were chosen. A large amount of efforts have been invested to optimize the ABEEM$\sigma\pi$/MM parameters to make them consistent, reliable, and transferable in accurately reproducing structural, energetic, and dynamic properties.

As is known, FF involves a series of parameters, such as bond stretching parameters, angle bending parameters, torsional parameters, van der Waals parameters, as well as the valence-state electronegativities and valence-state hardnesses in the electrostatic energy term etc. Strenuous and time-consuming work is needed for calibration of the parameters for development of a good FF.

The parameters χ^* and η^* are fitted through a regression and least-squares optimization procedure and are listed in Table S1 (Supporting Information), where the scaling Pauling electronegativity unit is used. The ABEEM$\sigma\pi$ labels are listed in Chart S1 (Supporting Information). The parameters are fitted not only to reproduce the charges of the *ab initio* calculation but also to fit the dipole moments, structures, and dimer binding energies given by experimental measurements and *ab initio* calculations. The fitting function k_{HB} is extracted from k in eq 8 to describe the electrostatic interaction between the H atom and *lp* of the acceptor. A general formulation of the k_{HB} function is available in SI1 (Supporting Information).

It is well-known that parameters for the hard degrees of freedom (bond stretching and angle bending) can be transferred from one FF to another without modification. For example, for some bonds and angles in the amide system, the same force constants, equilibrium bond lengths, and bond angles are used in both OPLS-AA[10] and AMBER[18] FFs. So we take the parameters of bond stretching and angle bending of protein directly from OPLS-AA FF and those of DNA from AMBER FF. The torsional terms are often regarded as "soft" degrees of freedom, in which most of the variations in structure and relative energy are due to the complex interplay between the torsional and nonbonded contributions. In ABEEM$\sigma\pi$ FF, we take the torsional and improper torsional parameters of OPLS-AA or AMBER as a reference and refit them through the least-squares optimization procedure to make the conformation energies and the key dihedral angles' root of mean square deviation (rmsd) of the model molecules be in good agreement with those calculated by the *ab initio* method. In addition, the Lennard-Jones parameters are determined by fitting *ab initio* conformational energies, dimer binding energies, dipole moments, and so on, using a regression and least-squares method. All the parameters are summarized in Table S1 (Supporting Information).

It should be pointed out that the parameters we have used are transferable over a large amount of chemical and

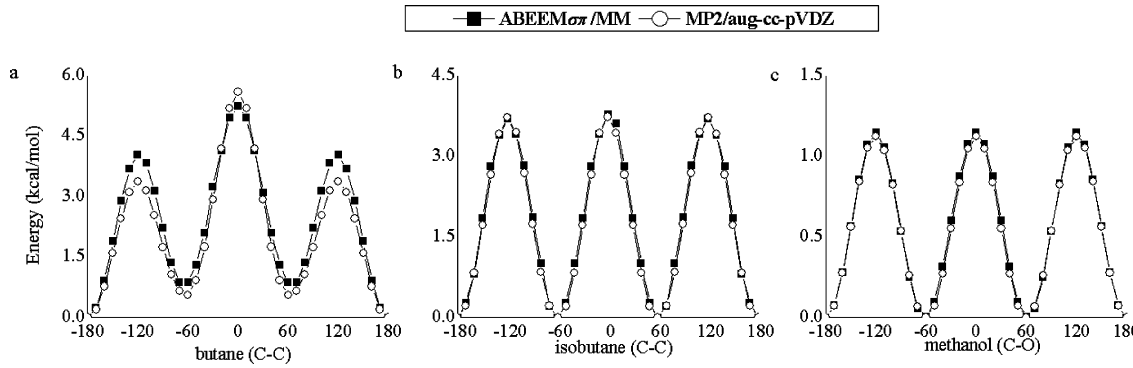


Figure 1. Rotational energy profiles (a) for the central C−C bond rotation of butane, (b) for the C−C bond rotation of isobutane, and (c) for the C−O rotation of methanol from MP2/aug-cc-pVDZ calculations and the ABEEM$\sigma\pi$ force field.

biological species rather than needing to be reparameterized for each molecule. The charge parameters, such as valence electronegativity and hardness for a type of site or region, are used for all molecules rather than only for one molecule in the ABEEM$\sigma\pi$ model. A site type in the ABEEM$\sigma\pi$ model is properly defined according to the type of its surrounding chemical environment, so that the parameters are transferable, only depending on the site types as in the usual force fields.

For example, a single water molecule has 7 total sites: three atomic sites, two σ bond sites, and two lone pair sites. Each site has two charge parameters (one valence electronegativity and one valence hardness). Besides the charge parameters, a single water molecule has 9 other parameters which contain r_{OH}, θ_{HOH}, k_{HOH}, ε_O, ε_H, σ_O, and σ_H, as well as α and D, because the stretch vibration of every O−H bond is described by the Morse potential function.

The ABEEM$\sigma\pi$/MM parameters and their calibration results are available in SI1 (Supporting Information).

Computational Details of Molecular Dynamics. The structures of Crambin, BPTI, and Trypsin from the protein data bank were used as the initial geometries of MD simulations. The MD simulations were performed using the modified TINKER program in the NVT ensemble with Berendsen thermostats, the velocity Verlet integrator, and a time step of 1 fs. The systems were initially heated over 5 ps to 285 K. The cutoff radius for nonbonding interactions was 10.0 Å with the minimum image convention if the periodic boundary condition was used. For all simulations, 0.5 ns of a MD run for equilibration was performed, followed by 9.5 ns of simulations for the calculation of various properties. We recomputed the partial charges of all sites using the ABEEM$\sigma\pi$ method every 0.1 ps. The charges are placed on some virtual sites according to the physical meaning, including the sites of σ and π bonds and lone pairs, for better calculating the electrostatic energy. When the MD simulation was performed, the force was only acting on the atoms with redistribution of the partial charges of bonds and lone pairs to the connected atoms. The partial charge of one σ bond is averagely assigned to two bonded atoms. The charges of π regions and lone pairs are all assigned to the connected atoms.

We performed simulations of 1crn by ABEEM$\sigma\pi$ (1792 sites) and standard AMBER (642 sites) force fields, respectively. These simulations have a time step of 1 fs, a simulation time of 1 ns, and a cutoff radius of 10.0 Å. With the same computational power, the AMBER fixed charge FF expends 10851.52 s. If the partial charges were fixed, the ABEEM$\sigma\pi$ FF expends 12093.66 s, and if the partial charges were recomputed every 0.1 ps, ABEEM$\sigma\pi$ FF uses 149116.20 s.

In what follows, we present results for representative systems obtained by the ABEEM$\sigma\pi$/MM and compare them with experimental data or results from the *ab initio* method at the MP2 level.

Results and Discussion

Rotational Barriers. Rotational energy profiles for the central C−C bond of butane and isobutane, and for the C−O rotation of methanol obtained at the ABEEM$\sigma\pi$/MM and MP2/aug-cc-pVDZ levels of theory, are shown in Figure 1. One can see that ABEEM$\sigma\pi$/MM accurately reproduces the rotational energy profiles. The experimental syn rotation and trans−gauche barriers of butane are 4.56[55] and 3.30 kcal/mol,[56] respectively. The corresponding results from ABEEM$\sigma\pi$/MM are 4.38 and 3.18 kcal/mol, whereas MP2 gives 4.95 and 2.72 kcal/mol, and AMBER[57] FF yields 5.31 and 3.53 kcal/mol, respectively. The experimental energy difference between trans and gauche isomers is 0.497−0.89 kcal/mol.[58] The result of ABEEM$\sigma\pi$/MM, MP2, and AMBER FF is 0.86, 0.65, and 0.79 kcal/mol, respectively. For isobutene, ABEEM$\sigma\pi$/MM yields 3.78 kcal/mol, comparable to the experimental[59] result of 3.90 kcal/mol and the AMBER[57] and MP2/aug-cc-pVDZ results, which are 3.48 kcal/mol and 3.61 kcal/mol, respectively. The experimental[60] barrier height of methanol is 1.07 kcal/mol, and the corresponding value from ABEEM$\sigma\pi$/MM and MP2/aug-cc-pVDZ is 1.12 and 1.14 kcal/mol, better than the results from MMFF94[61] and MM3[62] FF, which are 1.23 and 0.78 kcal/mol, respectively.

Structures of Peptides. Alanine dipeptide is a model compound of the protein backbone, containing two peptide linkages. We use the *ab initio* result at the level of MP2/6-311++G(d,p)//HF/6-31G(d,p) as the reference. The same as for the *ab initio* result, ABEEM$\sigma\pi$/MM is able to find all six of its stable conformers, especially β_2 and α_L, which are difficult to locate because they are shallow minima. Table S3 (Supporting Information) shows conformational energies

Table 1. Conformational Energies (in kcal/mol), RMSD of Conformational Energies, and the Key Dihedral Angles (in deg) for Alanine Tetrapeptide Relative to the *ab Initio* Data[a]

	ab initio[b]	ABEEM$\sigma\pi$	*ab initio*[c]	ABEEM[d]	OPLS-AA/L[e]	FQ-Dipole[f]	OPLS-FQ[g]	AMBER[h]	CHARMM-FQ[i]
1	4.36	4.85/2.8	2.71	2.88/5.6	3.19/4.4	2.88/-	3.03/-	3.52/-	5.15/-
2	3.76	2.54/3.5	2.84	1.94/8.3	3.19/6.5	1.84/-	3.97/-	3.74/-	3.42/-
3	0.00	0.00/3.7	0.00	0.00/7.8	−0.32/8.4	0.22/-	0.26/-	0.00/-	−0.70/-
4	4.67	4.22/5.6	4.13	2.90/8.2	4.40/5.8	3.69/-	2.34/-	4.32/-	5.36/-
5	4.64	4.15/5.1	3.88	4.81/7.3	3.14/9.3	3.70/-	4.62/-	3.64/-	4.21/-
6	0.66	0.67/4.0	2.20	1.80/8.8	0.96/12.7	1.45/-	1.67/-	2.91/-	2.40/-
7	4.13	4.51/9.0	5.77	5.83/11.7	5.82/6.6	5.48/-	6.18/-	4.54/-	5.75/-
8	4.16	4.16/12.7	4.16	4.10/12.7	4.83/18.8	5.38/-	4.39/-	5.91/-	3.40/-
9	6.13	5.76/3.4	6.92	7.21/5.9	7.14/8.2	6.74/-	5.33/-	4.93/-	6.10/-
10	5.10	4.99/6.0	6.99	8.03/4.3	7.25/14.2	8.21/-	7.82/-	8.53/-	4.55/-
rmsd		0.50/6.3[j]		0.67/8.4[k]	0.56/10.4[k]	0.71[k]	0.94[k]	1.14[k]	1.25[k]

[a] The values before "/" are the conformational energies, and those after "/" are the rmsd of the key dihedral angles relative to HF/6-31G(d,p) level geometries, which are listed in Table S2 (Supporting Information). The FQ-dipole, OPLS-FQ, AMBER, and CHARM-FQ methods did not provide the definite dihedral information. [b] The conformational energies at the MP2/6-311++G(d,p)//HF/6-31G(d,p) level calculated in this work. [c] The conformational energies at the LMP2/cc-pVTZ(-f)//HF/6-31G(d,p) level are cited from ref 5. [d] Ref 53, ABEEM FF only involves the partial charge sites of atoms and bonds without adding lone pair sites and π charge sites. [e] Ref 12. [f] Ref 45. [g] Ref 11. [h] Ref 19. [i] Ref 7. [j] The lowest rmsd relative to the results of the MP2/6-311++G(d,p) calculation. [k] The lowest rmsd relative to the results of the LMP2/cc-pVTZ(-f) calculation.

and rmsd's of key dihedral angles for alanine dipeptide from ABEEM$\sigma\pi$/MM, ABEEM/MM,[53] OPLS-AA/L,[12] OPLS-PFF,[13] FQ-dipole,[45] and OPLS-FQ[11] methods in comparison with *ab initio* results. It is seen from the table that the ABEEM$\sigma\pi$/MM method best reproduces both conformational energies and the dihedral angles for alanine dipeptide, giving the lowest rmsd for both quantities. Its rmsd value relative to the *ab initio* MP2/cc-pVTZ//MP2/6-31G(d,p) result[63] in conformational energies is merely 0.20 kcal/mol, and that of the key dihedral angles is only 5.9°. In comparison with ABEEM/MM, whose only difference from ABEEM$\sigma\pi$/MM is without partial charge sites for lone pairs and π bonds, a marked difference in rmsd values is also apparent.

For alanine tetrapeptide, Table 1 displays a rmsd in conformational energies and key dihedral angles obtained from ABEEM$\sigma\pi$/MM and other methods, compared to *ab initio* results. ABEEM$\sigma\pi$/MM gives a rmsd of 0.50 kcal/mol in conformational energies and 6.3° in the key dihedral angles relative to the MP2/6-311++G(d,p)//HF/6-31G(d,p) data, whereas those quantities are 0.67 kcal/mol and 8.4° for ABEEM/MM,[53] 0.56 kcal/mol and 10.4° for OPLS-AA/L,[12] and 0.69 kcal/mol and 19.1° for OPLS/PFF,[13] respectively. For FQ-Dipole,[45] OPLS-FQ,[11] AMBER,[19] and CHARMM-FQ,[7] the rmsd in conformational energies is 0.71, 0.94, 1.14, and 1.25 kcal/mol, respectively, in comparison with the LMP2/cc-pVTZ(-f)//HF/6-31G(d,p) result. No dihedral angle rmsd result is available for those latter approaches for this system.

Table 2 summarizes rmsd values in conformational energies and key dihedral angles obtained for the ABEEM$\sigma\pi$/MM, OPLS-AA/L, and OPLS/PFF methods in comparison with the *ab initio* result for various neutral dipeptides and tetrapeptide. The average rmsd in conformational energies and key dihedral angles of neutral peptides are 0.36 kcal/mol and 4.3° for ABEEM$\sigma\pi$/MM, respectively, but 0.47 kcal/mol and 10.1° for OPLS-AA/L[12] and 0.43 kcal/mol and 10.5° for OPLS/PFF, respectively.[13] For phenylalanine, tryptophan, tyrosine, and histidine dipeptides containing at least one aromatic ring, because ABEEM$\sigma\pi$/MM explicitly adds partial charges to π bonds, significantly lower rmsd values in angles are obtained, 3.4° on average for the four dipeptides from ABEEM$\sigma\pi$/MM, whereas the corresponding result from OPLS-AA/L[12] and OPLS/PFF[13] is 14.6° and 14.0°, respectively. Hydrogen bonding is closely related to the orientation of lone pairs of polar atoms. For asparagine and glutamine dipeptides with more hydrogen bonds, since ABEEM$\sigma\pi$/MM explicitly augments partial charges sites for lone pairs, noticeable improvements in rmsd are observed, 0.00 and 0.85 kcal/mol in conformation energies and 6.8° and 4.8° in dihedral angles for the two peptides, respectively, compared to 0.16 and 0.96 kcal/mol and 19.5° and 13.9° from OPLS-AA/L[12] and 0.02 and 0.92 kcal/mol and 8.7° and 18.0° from OPLS/PFF[13] for them. ABEEM$\sigma\pi$/MM results agree well with OPLS-AA/L and OPLS/PFF results for aliphatic amino acid dipeptides, such as valine, leucine,

Table 2. RMSDs in Conformational Energies (in kcal/mol) and the Key Dihedral Angles (in deg) for Different Peptides Relative to the *ab Initio* Result

molecule	ABEEM$\sigma\pi$	OPLS-AA/L[a]	OPLS/PFF[b]
Di-Ala	0.20/5.9	0.27/6.5	0.35/7.1
Tetra-Ala	0.50/6.3	0.56/10.4	0.69/19.1
Di-Phe	0.00/2.7	0.15/7.5	0.02/9.5
Di-Trp	0.68/4.2	0.50/24.2	0.49/19.4
Di-Tyr	0.27/3.4	0.39/8.1	0.27/8.9
Di-His	1.01/3.5	0.85/18.7	0.83/18.2
Di-Asn	0.00/6.8	0.16/19.5	0.02/8.7
Di-Gln	0.85/4.8	0.96/13.9	0.92/18.0
Di-Val	0.00/1.9	0.08/8.4	0.01/5.1
Di-Leu	0.30/3.0	0.34/6.1	0.35/5.1
Di-Ile	0.59/5.3	0.38/5.5	0.88/11.8
Di-Ser	0.33/5.4	0.44/4.9	0.34/8.1
Di-Cys	0.11/3.3	0.35/5.8	0.27/4.8
Di-Met	0.24/3.2	0.59/5.2	0.53/5.4
Di-Thr	0.37/5.3	0.87/7.1	0.75/8.9
average	0.36/4.3	0.47/10.1	0.43/10.5
Di-Asp	0.15/4.1	0.16	0.77
Di-Glu	1.29/5.0	1.53	1.47
Di-Lys	0.58/2.9	0.88	0.59
Di-Pro his	0.70/5.9	0.97	0.97
Di-Arg	1.14/4.1	1.15	0.79
average	0.77/4.4	0.94	0.92

[a] Ref 12. [b] Ref 13.

Table 3. RMSD and Linear Regression Analyses of Interaction Energies for H-Bonded Nucleic Acid Bases Relative to MP2/6-31G(d)(0.25) Values

force field	rmsd	R^a	SD^b	A^c	B^c	AAD^d
ABEEM$\sigma\pi$	0.90	0.98	0.77	−0.87	0.97	0.7
CHARMM27[e]	1.16	0.96	1.15	−0.34	0.96	0.9
AMBER4.1[f]	1.21	0.98	0.93	−0.59	1.10	0.9
CVFF[f]	4.83	0.88	1.30	0.79	0.62	4.4
CFF95[f]	1.72	0.95	1.05	1.64	0.80	1.2
OPLS[f]	2.65	0.95	1.76	−1.69	0.95	2.4

[a] Correlation coefficient. [b] Standard deviation (in kcal/mol). [c] The equation is $Y = A + Bx$, where A is intercept, B is slope, and x represents the MP2/6-31G(d)(0.25) values. [d] Average absolute deviation (in kcal/mol). [e] Ref 65. [f] Ref 64.

and isoleucine dipeptides, and polar amino acid dipeptides, including serine, threonine, cysteine, and methionine dipeptides. There is nevertheless an important difference, where OPLS-AA/L and OPLS/PFF performed structure optimizations with all the key dihedral angles constrained to their *ab initio* structure positions for the charged dipeptides, but in ABEEM$\sigma\pi$/MM calculations no such constraints were applied. The average rmsd value in conformational energy was 0.77 kcal/mol for ABEEM$\sigma\pi$/MM, whereas that quantity was 0.94 and 0.92 kcal/mol for OPLS-AA/L and OPLS/PFF, respectively. Furthermore, ABEEM$\sigma\pi$/MM obtained a rather small average angular rmsd of 4.4°. Proline is a special case, where disulfide bridges stabilize polypeptides and proteins. We considered several conformations for them whose calculated RMSDs were also found to be small, as listed in Table S3 (Supporting Information).

Interactions in Base Pairs and Water Dimer, Dipeptide−Water and Base−Water Complexes. We computed the interaction energies with ABEEM$\sigma\pi$/MM and compared them with other FFs and MP2/6-31G(d)(0.25) results of 26 H-bonded base pairs from Hobza et al.[64] Table 3 exhibits their rmsd's and linear regression analyses. The rmsd of ABEEM$\sigma\pi$/MM is 0.90 kcal/mol, smaller than that of CHARMM27,[65] AMBER4.1,[64] CFF95,[64] and OPLS[64] force fields. The correlation coefficient R is 0.98, and the standard deviation is 0.77 kcal/mol, with the intercept A and slope B in the fitted linear function of $Y = A + Bx$ being −0.87 and 0.97, respectively. The average absolute deviation is 0.7 kcal/mol, which is also the smallest among the methods tested. The good agreement between ABEEM$\sigma\pi$/MM results and high-level *ab initio* data indicates that ABEEM$\sigma\pi$/MM can correctly predict the interaction energy between H-bonded nucleic acid bases. The detailed information of the above calculations is contained in Table S4 (Supporting Information).

The stacked structures of A and T (denoted as ATs) and C and G (denoted as CGs) base pairs are what we investigated next. Their initial geometries (shown in Figure 2) are from ref 66. We calculated interaction energies of the stacked bases as a function of the distance between two planes formed by two base pairs. Single point calculations were done with MP2/6-31G(d), MP2/6-311++G(d,p), and ABEEM$\sigma\pi$/MM methods. The vertical separation between A and T base planes and between C and G planes was varied from 2.8 Å to 10.0 Å. The potential energy profiles of ATs and CGs are shown in Figure 2a and b, respectively. One

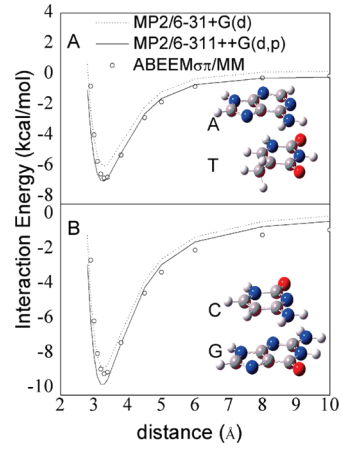


Figure 2. Potential energy profiles of stacked nucleic acid base pairs. (a) ATs. (b) CGs.

Table 4. Lowest Interaction Energies of Stacked Nucleic Acid Base Pairs (kcal/mol)

method	ATs	CGs
MP2/6-31+G(d)	−6.0	−8.9
MP2/6-311++G(d,p)	−7.0	−9.9
ABEEM$\sigma\pi$/MM	−6.8	−9.3
PMP[a]	−9.7	−10.9
AMBER[a]	−8.7	−13.0

[a] Ref 66.

local minimum is found with a vertical separation of 3.3 Å for each profile. Table 4 summarizes the lowest interaction energies of ATs and CGs from ABEEM$\sigma\pi$/MM, Nakagawa's polarizable model potential function,[66] AMBER[66] force field, MP2/6-31+G(d), and MP2/6-311++G(d,p) calculations. The interaction energies by Nakagawa's model and AMBER are underestimated. The lowest ABEEM$\sigma\pi$/MM interaction energies for ATs and CGs are −6.8 and −9.3 kcal/mol, respectively, in good agreement with *ab intio* results of −7.0 and −9.9 kcal/mol at the MP2/6-311++G(d,p) level of theory, indicating that ABEEM$\sigma\pi$/MM is suitable to predict the interaction energy for stacked nucleic acid base pairs.

We next considered the monoclinic B-DNA decamer $(CCAACGTTGG)_2$, which has five crystallographically different base pairs. Only nucleic acid bases are considered, and sugar−phosphate units are omitted. The interaction energy for H-bonded pairs, intrastrand stacking pairs, and interstrand stacking pairs from the MP2/6-31G(d) (0.25) (BSSE corrected) calculation, ABEEM$\sigma\pi$, CHARMM27,[65] and AMBER force fields[64] are reported in Table S5 (Supporting Information). At the bottom of Table S5 is the sum of different types of interacting pairs, including 10 H-bonded interactions (ΣH), 18 intrastrand stacking interactions (ΣS), 18 interstrand stacking interactions (ΣI), and the total interaction energies (ΣH + S + I). The sum of H-bonded base pair interaction energy from the ABEEM$\sigma\pi$ force field is −212.23 kcal/mol, which is very close to the −209.4 kcal/mol from the MP2 calculation. The intrastrand stacking interaction energy from ABEEM$\sigma\pi$/MM is −52.76 kcal/mol, higher than that from the MP2 calculation (−71.0 kcal/mol). For the interstrand stacked base pairs, the dominant contributions of electrostatic and Lennard-Jones term are varied. The MP2 interaction energy of interstrand stacked base pairs is

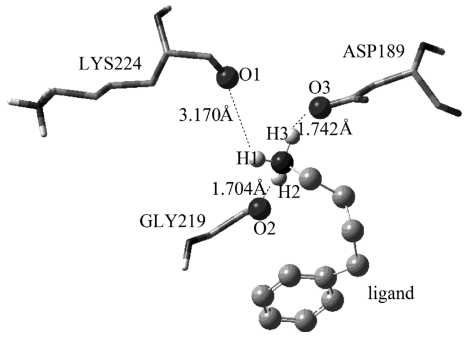


Figure 3. The configuration of the active pocket of 1tni.

well reproduced by ABEEM$\sigma\pi$/MM with a difference of less than 1.2 kcal/mol. The above analyses suggest that there exits a balance between H-binding and stacking interactions in the system, and electrostatic and Lennard-Jones contributions in base−base interactions can satisfactorily be treated by ABEEM$\sigma\pi$/MM.

The treatments of the most stable conformer of water clusters have been given elsewhere.[51] Here, we present the detailed study results for ten water dimer stationary conformers. The ABEEM$\sigma\pi$ force field can obtain all ten lowest conformers of the water dimer and excellently reproduce the structures obtained from counterpoise corrected MP2/aug-cc-pVDZ calculations obtained in this work (Table S6, Supporting Information). The binding energies of the ABEEM$\sigma\pi$ force field are comparable with those results of high level *ab initio* CCSD(T)/6-311++G(3df,3pd) calculations (Table S7, Supporting Information).

We have further investigated other systems of dipeptide−water and base−water clusters, such as (1) alanine dipeptide with $(H_2O)_{1-4}$ by ABEEM$\sigma\pi$/MM and MP2/6-31+G(d)//B3LYP/6-31G(d) methods and (2) bases with $(H_2O)_{1-3}$ by ABEEM$\sigma\pi$/MM and MP2/6-311++G(d,p)//B3LYP/6-311++G(d,p) methods, whose geometries, hydrogen bond energies, and cooperative properties are presented in Table S4 (Supporting Information). Results from all these systems in SI1 (Supporting Information) confirm that reliable predictions of interaction energies can be quantitatively obtained from ABEEM$\sigma\pi$/MM.

Illustration of Partial Charges in a Protein/Ligand Docking System. Protein−ligand recognition is a complicated process, but mandatory for the structure-based drug discovery and design. One quantity that is fundamentally important and essential in the process is to reliably and accurately calculate partial charges for ligands and protein, especially for those atoms near the active site and the interface between the donor and acceptor. For example, to design inhibitors for serine proteinase (PDB ID: 1tni), three hydrogen atoms (H1, H2, and H3) of the amine group at the tail of the ligand are found to form hydrogen bonds and one salt-bridge (between positively charged protonated amine and negatively charged aspartic acid) with three receptor atoms (see Figure 3), O1(LYS224), O2(GLY219), and O3(ASP189). The partial charges on those three hydrogen atoms are 0.415, 0.535, and 0.536 at the *ab initio* HF/6-31G level,[67] respectively. When evaluated by ABEEM$\sigma\pi$/MM, which uses a global scale factor k for the electrostatics energy calculations, their charges are 0.443, 0.519, and 0.566, close to the *ab initio* result and with the same trend. If OPLS-AA is used, each of these three hydrogen atoms has a fixed charge of 0.330. In a recent calculation by Cho et al. using a QM/MM model,[68] their charges are 0.340, 0.360, and 0.440, much smaller than the *ab initio* values. These results reveal that ABEEM$\sigma\pi$/MM not only takes good care of the polarization effect but it can also satisfactorily treat the salt-bridge effect.

MD Simulations for Proteins. To demonstrate the reliability of ABEEM$\sigma\pi$ PFF in molecular dynamics simulations, we performed 10 ns MD simulations for crambin (PDB ID: 1crn) and bovine pancreatic trypsin (PDB ID: 5pti), examined the obtained protein backbone geometries, and then compared them with those from the CHARMM force fields. The average rmsd's of the two proteins in backbone bond lengths are 0.037 Å and 0.040 Å, and the rmsd's in backbone bond angles are 4.16° and 4.73°, respectively. The rmsd value in backbone dihedral angles, φ, Ψ, and ω, from ABEEM$\sigma\pi$/MM is smaller than the result from CHARMM: 25.0°, 23.3°, and 7.6° for crambin and 29.4°, 30.2°, and 7.8° for trypsin from ABEEM$\sigma\pi$/MM and 26.0°, 29.6°, and 6.9° for crambin and 40.8°, 36.2°, and 8.7° for trypsin from CHARMM, respectively. These results demonstrate that ABEEM$\sigma\pi$/MM

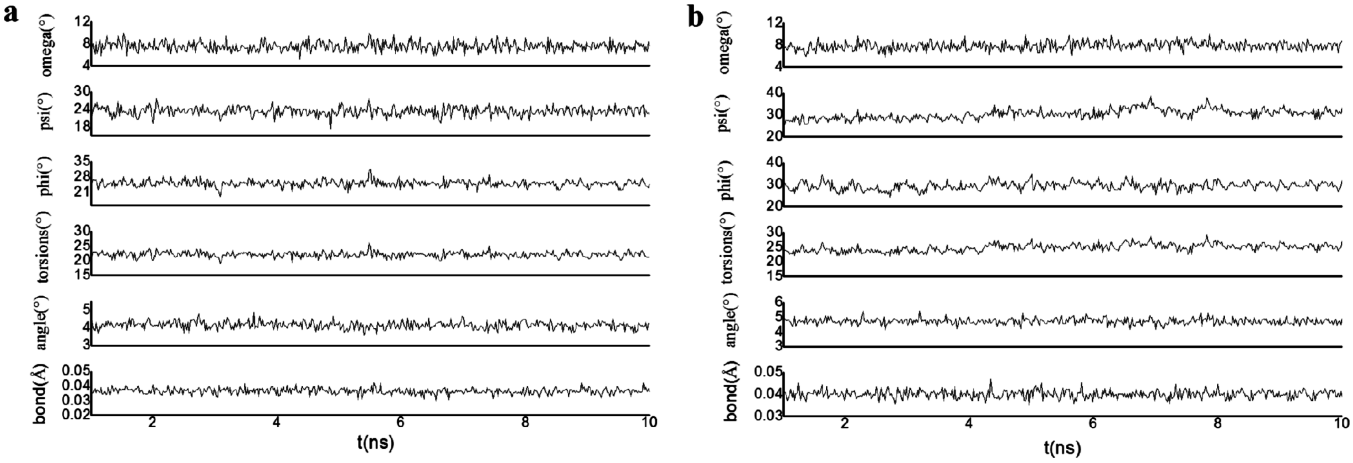


Figure 4. The rmsd in bond lengths, bond angles, and key dihedral angles (φ, Ψ, and ω) between X-ray structures and the ABEEM$\sigma\pi$/MM structures as a function of simulation time for 1crn and 5pti.

can reliably reproduce the protein backbone from the X-ray structure. Figure 4 details the coordinate rmsd of different groups of non-hydrogen atoms, indicating that the structures during the course of MD simulations were stable and the simulations were in the equilibrium state. The rmsd averages over C_α, backbone heavy atoms, all heavy atoms, side chain heavy atoms between experimental structures, and MD simulations using ABEEM$\sigma\pi$ are 2.075, 2.022, 2.222, and 2.502 Å for crambin and 1.786, 1.911, 2.800, and 3.461 Å for trypsin; the rmsd averages over backbone heavy atoms, all heavy atoms, side chain heavy atoms between experimental structures, and MD simulations using CHARMM[5] are 1.700, 1.910, and 2.160 Å for crambin and 2.580, 3.190, and 3.730 Å for trypsin. Average hydrogen bond distances of N···O pairs and O···O pairs of proteins are presented in Table S8 (Supporting Information). ABEEM$\sigma\pi$/MM can accurately describe hydrogen bond interactions for the protein systems. Overall, the results of ABEEM$\sigma\pi$/MM are better than those from CHARMM.[5] This is because, in ABEEM$\sigma\pi$/MM, there is more than one partial charge per atom site, providing more flexibility to account for the polarization effect by allowing different partial charges to adjust their values in accordance with changing geometry and environment during the course of dynamic simulation processes.

Conclusions

A polarizable force field, ABEEM$\sigma\pi$/MM, with multiple fluctuating partial charges per atom site is proposed and evaluated in the present work. The main difference between ABEEM$\sigma\pi$/MM and other force fields is the treatment of the electrostatic interaction energy. In ABEEM$\sigma\pi$ FF, partial charges are evaluated and updated by the ABEEM$\sigma\pi$ method based on the electronegativity equalization principle. The ABEEM$\sigma\pi$ method explicitly considers contributions from lone pairs, σ and π bond regions. This addition increases the flexibility to better account for the polarization effect for both intra- and intermolecular processes.

With high-level *ab initio* [MP2, MP4, and/or CCSD(T)] calculation results or experimental data as the reference, we tested and evaluated ABEEM$\sigma\pi$/MM with a number of biomolecular systems. We found that it can satisfactorily reproduce structural and energetic properties, such as rotational energy profiles, dihedral angles for peptides, conformational energies, partial charge distribution inside an active site, and protein structures in MD simulations. The interaction energies of H-bonded and stacked nucleic base pairs are in good agreement with the available experimental data and *ab initio* results. The ABEEM$\sigma\pi$ polarizable force field also provides reliable binding energies for ten stationary conformers of water dimer, dipeptide−water, and base−water clusters. These results indicate that the fluctuating charge force field, ABEEM$\sigma\pi$/MM, is accurate and reliable and can be applied to wider ranges of biomolecular systems, including aqueous solutions.

Acknowledgment. Helpful comments and suggestions from Professors Robert G. Parr and Lee G. Pedersen of University of North Carolina are gratefully acknowledged. We greatly thank Professor Jay William Ponder for providing the Tinker program; thanks are also given to the support of the National Science Foundation of China (NSFC, Nos. 20633050, 20873055 and 20703022), and the Department of Education of Liaoning Province (Nos. 2008S133, 2009T057, LNET RC0503, and 20060494).

Supporting Information Available: ABEEM$\sigma\pi$ for calculating charge distribution in molecules; ABEEM$\sigma\pi$ force field parameter determination, calibration, and evaluation; training set; interactions in water dimer, dipeptide−water, and base−water complexes. This information is available free of charge via the Internet at http://pubs.acs.org/.

References

(1) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.

(2) Grossman, J. C.; Schwegler, E.; Draeger, E. W.; Gygi, F.; Galli, G. *J. Chem. Phys.* **2004**, *120*, 300–311.

(3) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.

(4) MacKerell, A. D., Jr.; Wiórkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946–11975.

(5) 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.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(6) Patel, S.; Brooks, C. L., III. *J. Comput. Chem.* **2004**, *25*, 1–15.

(7) Patel, S.; MacKerell, A. D., Jr.; Brooks, C. L., III. *J. Comput. Chem.* **2004**, *25*, 1504–1514.

(8) Xie, W.; Pu, J.; MacKerell, A. D., Jr.; Gao, J. *J. Chem. Theory Comput.* **2007**, *3*, 1878–1889.

(9) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.

(10) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(11) Banks, J. L.; Kaminski, G. A.; Zhou, R.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **1999**, *110*, 741–754.

(12) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.

(13) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X. X.; Murphy, R. B.; Zhou, R. H.; Halgren, T. A. *J. Comput. Chem.* **2002**, *23*, 1515–1531.

(14) Winn, P. J.; Ferenczy, G. G.; Reynolds, C. A. *J. Comput. Chem.* **1999**, *20*, 704–712.

(15) Jorgensen, W. L.; Severance, D. L. *J. Am. Chem. Soc.* **1990**, *112*, 4768–4774.

(16) Essex, J. W.; Severance, D. L.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **1997**, *101*, 9663–9669.

(17) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.

(18) 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*, 5179–5197.

(19) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999–2012.

(20) Yang, L. J.; Tan, C.-H.; Hsieh, M.-J.; M., W. J.; Duan, Y.; Cieplak, P.; Caldwell, J.; Kollman, P. A.; Luo, R. *J. Phys. Chem. B* **2006**, *110*, 13166–13176.

(21) Engkvist, O.; Åstrand, P.-O.; Karlström, G. *Chem. Rev.* **2000**, *100*, 4087–4108.

(22) Kaminský, J.; Jensen, F. *J. Chem. Theory Comput.* **2007**, *3*, 1774–1788.

(23) Penev, E.; Ireta, J.; Shea, J.-E. *J. Phys. Chem. B* **2008**, *112*, 6872–6877.

(24) Donchev, A. G.; Galkin, N. G.; Illarionov, A. A.; Khoruzhii, O. V.; Olevanov, M. A.; Ozrin, V. D.; Subbotin, M. V.; Tarasov, V. I. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8613–8617.

(25) Khoruzhii, O.; Donchev, A. G.; Galkin, N.; Illarionov, A.; Olevanov, M.; Ozrin, V.; Queen, C.; Tarasov, V. I. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 10378–10383.

(26) Donchev, A. G.; Ozrin, V. D.; Subbotin, M. V.; Tarasov, O. V.; Tarasov, V. I. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7829–7834.

(27) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227–249.

(28) Warshel, A.; Kato, M.; Pisliakov, A. V. *J. Chem. Theory Comput.* **2007**, *3*, 2034–2045.

(29) Rasmussen, T. D.; Ren, P.; Ponder, J. W.; Jensen, F. *Int. J. Quantum Chem.* **2007**, *107*, 1390–1395.

(30) Xie, W.; Gao, J. *J. Chem. Theory Comput.* **2007**, *3*, 1890–1900.

(31) Wertz, D. H.; Allinger, N. L. *Tetrahedron* **1974**, *30*, 1579–1586.

(32) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490–519.

(33) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361–2381.

(34) Lifson, S.; Hagler, A. T.; Dauber, P. *J. Am. Chem. Soc.* **1979**, *101*, 5111–5121.

(35) Clark, M.; Cramer, R. D., III; van Opdenhosch, N. *J. Comput. Chem.* **1989**, *10*, 982–1012.

(36) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596–3607.

(37) Parr, R. G.; Yang, W. Chemical Potential Derivatives. In *Density Functional Theory of Atom and Molecules*; Oxford University Press and Clarendon Press: New York, 1989; pp 90−95.

(38) Gerrlings, P.; De Proft, F.; Langenaeker, W. *Chem. Rev.* **2003**, *103*, 1793–1873.

(39) Rappé, A. K.; Goddard, W. A., III. *J. Phys. Chem.* **1991**, *95*, 3358–3363.

(40) Bakowies, D.; Theil, W. *J. Comput. Chem.* **1996**, *17*, 87–108.

(41) Shimizu, K.; Chaimovich, H.; Farah, J. P. S.; Dias, L. G. *J. Phys. Chem. B* **2004**, *108*, 4171–4177.

(42) Menegou, G.; Shimizu, K.; Farah, J. P. S.; Dias, L. G.; Chaimovich, H. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5933–5936.

(43) Mortier, W. J.; Ghosh, S. K.; Shankar, S. *J. Am. Chem. Soc.* **1986**, *108*, 4315–4320.

(44) York, D. M.; Yang, W. *J. Chem. Phys.* **1996**, *104*, 159–172.

(45) Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. B* **1999**, *103*, 4730–4737.

(46) Chelli, R.; Procacci, P. *J. Chem. Phys.* **2002**, *117*, 9175–9189.

(47) Jorgensen, W. L. *J. Chem. Theory Comput*, **2007**, *3*, foreword.

(48) Yang, Z. Z.; Cui, B. Q. *J. Chem. Theory Comput.* **2007**, *3*, 1561–1568.

(49) Yang, Z. Z.; Wang, C. S. *J. Phys. Chem. A* **1997**, *101*, 6315–6321.

(50) Cong, Y.; Yang, Z. Z. *Chem. Phys. Lett.* **2000**, *316*, 324–329.

(51) Yang, Z. Z.; Wu, Y.; Zhao, D. X. *J. Chem. Phys.* **2004**, *120*, 2541–2557.

(52) Wu, Y.; Yang, Z. Z. *J. Phys. Chem. A* **2004**, *108*, 7563–7576.

(53) Yang, Z. Z.; Zhang, Q. *J. Comput. Chem.* **2006**, *27*, 1–10.

(54) Yang, Z. Z.; Qian, P. *J. Chem. Phys.* **2006**, *125*, 064311.

(55) Comptom, D. A.; Montero, S. M.; Murphy, W. F. *J. Phys. Chem.* **1980**, *84*, 3587–3591.

(56) Pitzer, K. S. *J. Chem. Phys.* **1944**, *12*, 310–314.

(57) Hwang, M.-J.; Stockfisch, T. P.; Hagler, A. T. *J. Am. Chem. Soc.* **1994**, *116*, 2515–2525.

(58) Durig, J. R.; Comptom, D. A. C. *J. Phys. Chem.* **1979**, *83*, 265–268.

(59) Lide, D. R., Jr.; Mann, D. E. *J. Chem. Phys.* **1958**, *29*, 914–920.

(60) Lowe, J. P. *Prog. Phys. Org. Chem.* **1968**, *6*, 1–80.

(61) Halgren, T. A.; Nachbar, R. B. *J. Comput. Chem.* **1996**, *17*, 587–615.

(62) Allinger, N. L.; Rahman, M.; Lii, J.-H. *J. Am. Chem. Soc.* **1990**, *112*, 8293–8307.

(63) Wang, Z. X.; Duan, Y. *J. Comput. Chem.* **2004**, *25*, 1699–1716.

(64) Hobza, P.; Kabeláč, M.; Šponer, J.; Mejzlík, P.; Vondrášek, J. *J. Comput. Chem.* **1997**, *18*, 1136–1150.

(65) Foloppe, N.; Mackerell, A. D. *J. Comput. Chem.* **2000**, *21*, 86–104.

(66) Nakagawa, S. *J. Comput. Chem.* **2007**, *28*, 1538–1550.

(67) Guan, Q. M.; Yang, Z. Z. *J. Theory Comput. Chem.* **2007**, *6*, 731–744.

(68) Cho, A. E.; Guallar, V.; Berne, B. J.; Friesner, R. *J. Comput. Chem.* **2005**, *26*, 915–931.

CT9006647