JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

General Model for Treating Short-Range Electrostatic Penetration in a Molecular Mechanics Force Field

Qiantao Wang,[†,‡] Joshua A. Rackers,[§] Chenfeng He,[†] Rui Qi,[†] Christophe Narth,[∥] Louis Lagardere,[∥] Nohad Gresh,[∥] Jay W. Ponder,[⊥] Jean-Philip Piquemal,*[,∥] and Pengyu Ren*[,†]

[†]Department of Biomedical Engineering and [‡]Division of Medicinal Chemistry, College of Pharmacy, The University of Texas at Austin, Austin, Texas 78712, United States

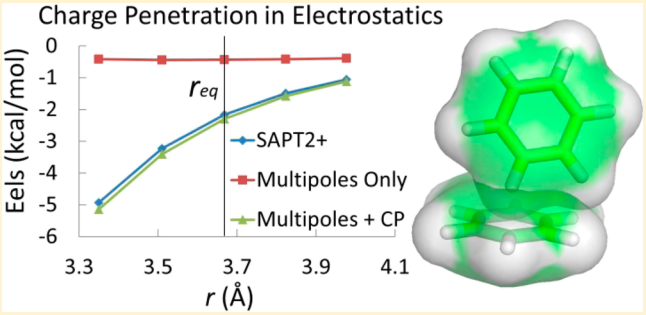
[§]Computational and Molecular Biophysics Program, Division of Biology & Biomedical Sciences, Washington University in St. Louis, St. Louis, Missouri 63110, United States

[∥]Laboratoire de Chimie Théorique, Sorbonne Universités, UPMC Paris 06, UMR 7616, Case Courrier 137, 4 Place Jussieu, F-75005 Paris, France

[⊥]Department of Chemistry, Washington University in St. Louis, St. Louis, Missouri 63130, United States

Supporting Information

ABSTRACT: Classical molecular mechanics force fields typically model interatomic electrostatic interactions with point charges or multipole expansions, which can fail for atoms in close contact due to the lack of a description of penetration effects between their electron clouds. These short-range penetration effects can be significant and are essential for accurate modeling of intermolecular interactions. In this work we report parametrization of an empirical charge−charge function previously reported (Piquemal, J.-P.; et al. *J. Phys. Chem. A* **2003**, 107, 10353) to correct for the missing penetration term in standard molecular mechanics force fields. For this purpose, we have developed a database (S101×7) of 101 unique molecular dimers, each at 7 different intermolecular distances. Electrostatic, induction/polarization, repulsion, and dispersion energies, as well as the total interaction energy for each complex in the database are calculated using the SAPT2+ method (Parker, T. M.; et al. *J. Chem. Phys.* **2014**, *140*, 094106). This empirical penetration model significantly improves agreement between point multipole and quantum mechanical electrostatic energies across the set of dimers and distances, while using only a limited set of parameters for each chemical element. Given the simplicity and effectiveness of the model, we expect the electrostatic penetration correction will become a standard component of future molecular mechanics force fields.



INTRODUCTION

Electrostatic interactions comprise one of the principle interatomic forces, along with exchange-repulsion, dispersion, and polarization or induction. The importance of electrostatic interactions is paramount at long range and for polar molecules. Much development effort has been focused on computational treatment of long-range electrostatics, e.g., the development of particle-meshed Ewald (PME) methods.[1−7] Electrostatic interactions at short range have received less consideration until recently. At close distances, a spherical approximation of atomic charge distributions is insufficiently accurate and use of atomic multipole expansions provides much greater flexibility in modeling complex electrostatic potentials near a molecular surface, an insight which inspired the development of the AMOEBA force field.[6,8−10] Nonetheless, at *very* close interatomic distances, when electron clouds overlap, a point multipole approximation becomes inadequate. The electrostatic potential within a spherical electron cloud no longer behaves as a simple $1/r$ interaction potential at small separation distances. Such deviation from a simple Coulomb potential is referred to as a penetration effect. While the charge penetration effect leads to a negative correction to energy at typical molecular interaction distances, where the electron−electron penetration is dominant, it can be repulsive at very short range.[11] A recent study by Lewis and co-workers reported the counterintuitive result that any ring substitutions of the benzene dimer (parallel) with electron-withdrawing or electron-donating groups yield more favorable electrostatic contributions than the unsubstituted benzene−benzene dimer itself.[12] This result is contrary to the conventional thought that such interactions are correlated with the ability to withdraw or donate electrons to the π cloud as described by the Hunter−Sanders rules.[13] Sherrill and co-workers suggested this is because the electrostatic interactions in such systems at the $\pi-\pi$ stacking distance exhibit a significant charge penetration effect. The multipole

Received: March 20, 2015
Published: April 28, 2015

DOI: 10.1021/acs.jctc.5b00267
J. Chem. Theory Comput. 2015, 11, 2609−2618

model, which the Hunter–Sanders rules are based upon, cannot correctly account for such effects.[14] Moreover, in a recent study of aromatic crystals, a charge penetration corrected AMOEBA-like model predicted better crystal properties than the uncorrected model.[15] It was shown that point atomic multipoles consistently predict positive (repulsive) electrostatic interactions between stacked or T-shaped benzene dimers while symmetry-adapted perturbation theory (SAPT)[16] suggests an opposite trend toward attractive interactions. The current AMOEBA force field seemingly compensates for penetration with a less repulsive van der Waals interaction, so the total interaction energy is reasonable at certain dimer configurations. However, explicit incorporation of the penetration effect provides much better anisotropy in crystal packing and makes the overall force field more transferable. In another study of organochlorine compounds using the AMOEBA model, it was found that the transferability of chlorine van der Waals parameters was unsatisfactory, likely due to lack of an explicit penetration correction.[17]

There have been previous attempts to incorporate the charge penetration effect into implicit solvent models,[18−20] multipole-based electrostatic models,[21−27] charge-density-based (including Gaussian multipole) models,[15,28−30] and combined quantum and molecular mechanics (QM/MM) models.[26,30−34] Generally, the charge penetration correction involves breaking the atom-centered point charge into an effective core and a valence electron density, as suggested by Gordon et al.[21] and Piquemal et al.[22] In this way, the electrostatic interaction between two atoms is described as a sum of interactions between core and valence charge densities, which can be modeled with empirical exponential functions. Alternatively, rigorous integration over the two charge densities can be used to model short-range electrostatic interactions,[15,28,29] with a significantly greater expenditure of computational effort. Others have explored incorporating charge penetration effects into the QM/MM framework, using either screened molecular mechanics (MM) charges[26,31−33] or simple empirical damping corrections.[34] Such screened MM charges are typically parametrized for QM/MM applications and may be not directly applicable in full MM calculations, e.g., to reproduce the attractive SAPT electrostatic energy in a stacked benzene–benzene conformation, unless an explicit term for interactions of the valence charge densities is included, as in the model recently proposed by Wang and Truhlar.[30]

In this study, the charge–charge penetration electrostatic model of Piquemal et al.[22] is revisited, implemented, and extensively tested in the context of the AMOEBA force field and using a new parametrization strategy. The charge penetration corrected AMOEBA point multipole model (multipoles + CP) is developed using a comprehensive set of small molecule complexes, and the parameters are determined for H, C, N, O, P, S, F, Cl, and Br to cover the elements commonly found in organic and biological molecules. To facilitate model development in this and future studies, a new database of SAPT2+[35,36] decomposed quantum mechanical energies constructed for 101 small molecule pairs, each at 7 different intermolecular distances (the S101×7 database), is presented.

METHODS

S101 and S101×7 Databases. In order to systematically examine the electrostatic and other components of intermolecular forces, the S101 and S101×7 databases of homo- and heterodimers of common organic molecules have been constructed. The S101 database contains 101 unique molecule pairs (Figure 1). The first 66 pairs, which cover the majority of

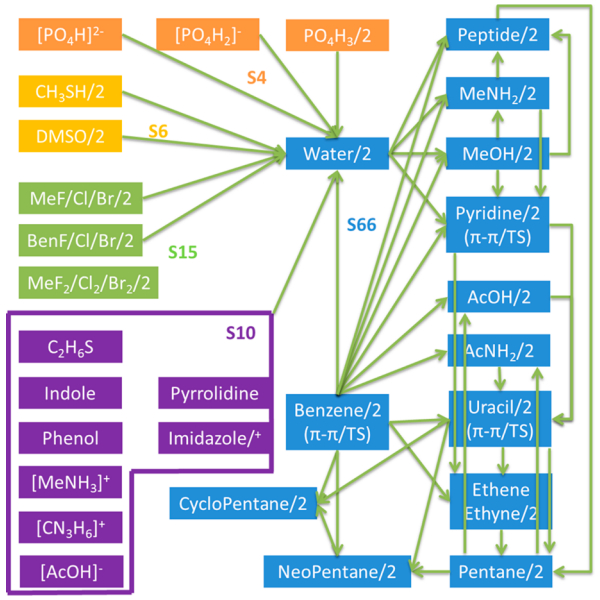


Figure 1. Schematic view of monomers and dimers in the S101 data set. The arrows connect two molecules that form a dimer; "/2" represents the existence of a homo dimer; "/+" indicates both neutral and ionized molecules are included. Different configurations of the same dimers, e.g., $MeNH_2$–water, phenol–water, benzene–benzene, and MeCl–MeCl, are included in the data set to take into account the orientational effect.

the typical organic interactions of H, C, N, and O atoms, are taken from the S66 database from Hobza et al.[37] In addition, 15 complexes containing halogen atoms (F, Cl, and Br), six complexes containing sulfur, and four complexes containing phosphorus have been added. Furthermore, 10 monomer–water complexes, which encompass amino acid side chain analogs (including the charged ones) missing in the S66 data set, have also been added, yielding a total of 101 pairs. To construct the S101×7 database, definitions of the intermolecular distance vectors from the S66×8 database of Hobza et al.[37] were used. Unlike S66×8, each of the 101 model complexes were placed at seven separation distances, corresponding to 0.70, 0.80, 0.90, 0.95, 1.00, 1.05, and 1.10 times the equilibrium intermolecular distances. Compared against the S66×8 database, the S101×7 set includes more dimer configurations at very short separations, which have been rarely investigated but are essential to the study of penetration effects and exchange–repulsion interactions. We have selected 0.7 times the equilibrium distance as the lower bound because the SAPT calculations below show that at this close distance the electrostatic energy is about 50% of the exchange–repulsion energy or higher; i.e., both electrostatic and van der Waals (vdW) components are important in the total interaction energy. As these short distances are being sampled in molecular dynamics simulations at room temperature and even more so at higher temperatures, their contributions to the simulated bulk properties are nonnegligible. Thus, it is essential to ensure the charge penetration model behaves correctly at these short distances.

DOI: 10.1021/acs.jctc.5b00267
J. Chem. Theory Comput. 2015, 11, 2609−2618

The newly added structures among the 101 complexes were optimized at the MP2/cc-pVTZ level of theory with counterpoise correction using the Gaussian09 program.[38] For each of the resulting 707 dimer configurations, the interaction energy has been decomposed using SAPT2+ analysis[35,36] provided by the PSI4 program.[39] The SAPT2+ calculation returns electrostatic, exchange−repulsion, induction, and dispersion energies, all to second order with respect to intramolecular electron correlation. Exact definitions of each component can be found in Figure 1 of Sherrill et al.[36] It should be noted that dispersion energy can only be separated from other effects in long range when two molecules do not overlap. Thus, at van der Waals distances, it may be more appropriately to refer to this as "dispersion-like" or "damped dispersion" energy. This should be kept in mind even though for simplicity the term "dispersion" is used throughout the discussion. All SAPT calculations were carried out using Dunning's correlation consistent basis sets[40,41] at both aug-cc-pVDZ and aug-cc-pVTZ levels. The complete basis set (CBS) limits of the SAPT2+ energies were also estimated. (Data can be found in the Supporting Information)

CBS Extrapolation Scheme. A two-point extrapolation strategy has been used to estimate the complete basis set limit of the exchange−repulsion and dispersion energy at the SAPT2+ level of theory. This is similar to Helgaker's scheme[42] but with an optimized p value (eq 1). Such a protocol has previously been applied to extrapolate the dispersion energy of DFT-SAPT calculation in earlier study.[43]

$$E_{\mathrm{CBS}} = \frac{E_{X+1}(X+1)^p - E_X X^p}{(X+1)^p - X^p} \tag{1}$$

Different p values, 3.0 for exchange−repulsion and 4.3 for dispersion energy, were obtained using the small pairs and subsequently applied for extrapolation over the full S101×7 database.

Scaling of the SAPT2+/CBS Dispersion Energy. Since the truncated terms in the SAPT2+ dispersion energy make a considerable contribution to the total interaction energy, dispersion energies obtained at the SAPT2+/CBS level are scaled by a factor f in order to match the SAPT total interaction energy to those obtained at the CCSD(T)/CBS level of theory (eq 2).

$$S = |E_{\mathrm{total}}^{\mathrm{CCSD(T)/CBS}} - (E_{\mathrm{non\text{-}disp}}^{\mathrm{SAPT2+/CBS}} + E_{\mathrm{disp}}^{\mathrm{SAPT2+/CBS}} f)| \tag{2}$$

By minimizing eq 2 using the 66 pairs in the S66 data set, a scale factor of $f = 0.89$ has been determined and used to construct the S101×7 database.

Modified Charge−Charge Interaction. In order to model the charge penetration effect, the method of Piquemal et al.[22] is revisited. Their original model corrects the charge−charge and charge−dipole interactions. Here, we propose to retain the charge−charge correction only. As a result, each atomic point charge is divided into an effective core and a damped valence electron distribution. Thus, the electrostatic energy between two atomic charges can be written as

$$\begin{aligned} E_{qq}(r) = [&Z_1Z_2 - Z_1(Z_2 - q_2)(1 - \exp(-\alpha_2 r)) \\ &- Z_2(Z_1 - q_1)(1 - \exp(-\alpha_1 r)) \\ &+ (Z_1 - q_1)(Z_2 - q_2)(1 - \exp(-\beta_1 r))(1 - \exp(-\beta_2 r))]/r \end{aligned} \tag{3}$$

where r is the interatomic distance; Z is the positive effective core charge, which is set to be equal to the number of valence electrons of each atom; q is the net charge of the atom, thus $(Z - q)$ can be considered as the magnitude of the (negatively charged) electron cloud; and α and β are two parameters controlling the magnitude of the damping of the electron cloud when the atom is interacting with the core and with electrons from other atoms, respectively. Thus, the total electrostatic energy between two atoms now involves three components, the core−core, core−electron, and electron−electron interactions.

Two methods have been explored to determine the α parameter values. The first method involves fitting the damped potential to the QM electrostatic potential at short range, near or within the molecular surface. By considering a probe charge of $+1e$ as a particle with an effective core charge of $+1e$ and having no valence electrons, $(Z_2 - q_2)$ becomes zero. Thus, the electrostatic potential can be written as

$$V_q(r) = [Z_1 - (Z_1 - q_1)(1 - \exp(-\alpha_1 r))]/r \tag{4}$$

Once Z and q are determined, α can be obtained easily by fitting eq 4 to the QM electrostatic potential.

In the second method, α is intuitively set to be the same as the number of valence electrons (except the hydrogen atom):

$$\alpha = \max\{Z, 2\} \tag{5}$$

When Z and β are fixed in eq 3, the electrostatic energy is more attractive when α is greater. This is in accordance with the intuition that atoms having a larger electron cloud may exhibit a stronger penetration effect. Although the final parameters for H, C, N, O, P, S, F, Cl, and Br were derived based on the second method, the performance of both methods is examined for H, C, N, and O containing molecules in later sections.

As the distance between two atoms increases, eq 3 will reduce to the classical Coulomb charge−charge interaction (q_1q_2/r). Thus, the electrostatic interaction at medium and long distances can still be accurately modeled via a multipole expansion, as the penetration correction diminishes rapidly with distance. As the data will show, the penetration correction is only significant when atomic separation is shorter than the sum of atomic van der Waals radii and thus does not affect the reciprocal space portion of an Ewald summation approach such as particle mesh Ewald (PME). In addition, to ensure the continuity between the real and reciprocal space, a switching function is used near the real space Ewald cutoff distance (typically 7 Å for atomic multipole PME) to ensure the penetration correction completely disappears:

$$f_{\mathrm{switch}}(r) = \begin{cases} 1, & r \le r_{\mathrm{l}} \\ 10\left(\dfrac{r_{\mathrm{u}} - r}{r_{\mathrm{u}} - r_{\mathrm{l}}}\right)^3 - 15\left(\dfrac{r_{\mathrm{u}} - r}{r_{\mathrm{u}} - r_{\mathrm{l}}}\right)^4 + 6\left(\dfrac{r_{\mathrm{u}} - r}{r_{\mathrm{u}} - r_{\mathrm{l}}}\right)^5, & r_{\mathrm{l}} < r < r_{\mathrm{u}} \\ 0, & r \ge r_{\mathrm{u}} \end{cases} \tag{6}$$

where r is the interatomic distance and r_{l} and r_{u} are the lower and upper bounds of the switching function.

Derivation of Atomic Point Multipole Moments. The permanent electrostatic energy in the AMOEBA force field includes through quadrupole−quadrupole interactions. Following previously detailed procedures,[9,10,44] an initial set of atomic multipole moments for each molecule was obtained from distributed multipole analysis (DMA)[45] at the MP2/6-311G** level of theory. Then the dipole and quadrupole moments were further optimized by fitting to the electrostatic potential

DOI: 10.1021/acs.jctc.5b00267
J. Chem. Theory Comput. 2015, 11, 2609−2618

calculated at the MP2/aug-cc-pVTZ level. This procedure is automated for organic molecules by the Poltype program.[44] The same strategy has been used in developing the AMOEBA force field for small molecules,[9] proteins,[10] and organochloroine compounds[17] previously. With the addition of the penetration correction described here, monopole−monopole (charge−charge) interactions are calculated using eq 3 while all other terms in the AMOEBA retain their original form. The SAPT and AMOEBA multipole based intermolecular interaction energies were compared on exactly the same dimer structures.

Parametrization of the Penetration Model. Since H, C, N, and O are the most common elements in organics and biomolecules, their parameters were determined first within the new charge penetration formalism. A training set of 35×7 molecule pairs was used, consisting of 15×7 pairs of hydrogen-bonded complexes, 12×7 pairs of dispersion-dominant complexes, and 8×7 pairs with mixed features of both. The initial parameter searching was done using a divide and conquer approach. For example, an initial set of parameters for sp^3 or sp^2 carbon, nonpolar hydrogen, and hydrogen attached to sp^2 C were obtained by selecting a smaller number of pairs (e.g., 8×7) from the dispersion-dominant and mixed complexes. Similarly, other parameters such as those for sp^3 oxygen and polar hydrogen were obtained initially from water dimers. Systematic scanning is feasible for determination of a small number of parameters. Then, one or more sets of initial parameters for all H, C, N, and O atom types were optimized together using the entire training set. Once the parameters for these four elements were finalized, further parametrization for P, S, F, Cl, and Br was carried out using the subsets in S101×7. In each of these subsets, an 80/20 ratio for the training and testing complexes was maintained to ensure a sufficient amount of data points for each atom type. An optimization program written in Python, using the quasi-Newton and Nelder−Mead simplex methods from the SciPy library, was applied to all of the parametrization work. The first derivative of the sum of unsigned errors with respect to each parameter was calculated numerically.

RESULTS

Convergence of the SAPT2+ Energy toward Basis Set Limit. In order to examine the convergence of the SAPT2+ energy toward the basis set limit, five small molecule pairs were selected and calculated using the aug-cc-pVXZ (X = D, T, Q, or 5, abbreviated as aXZ in the following paragraphs) basis sets. The five pairs are water−water, water−methanol, water−methylamine, ethyne−ethyne (T-shaped), and ethyne−water (CH···O). The energy difference between different aXZ (X = D, T, Q, and 5) basis sets for the total interaction energy and each energy component, including electrostatic, induction, exchange−repulsion, and dispersion energy, are compared. In general, a steady decrease in the energy gaps between aDZ−aTZ, aTZ−aQZ, and aQZ−a5Z can be observed (Table 1). The difference between aQZ and a5Z basis sets of all of the energy components as well as the total interaction energy are already well below 0.05 kcal/mol. In particular, the differences in electrostatic and induction energies are even smaller, at 0.003 and 0.005 kcal/mol, respectively. This implies that the difference between a5Z and a bigger basis set, e.g., a6Z, should be even smaller and negligible. Therefore, the results obtained using the a5Z basis set were used to approximate the complete basis set limit.

Table 1. Average Unsigned Differences between Energy Calculated at Different SAPT2+/aug-cc-pVXZ (X = D, T, Q, and 5) Levels of Theory

	MUE (kcal/mol)				
	electrostatic	induction	exch−repul	dispersion	total
aDZ − aTZ	0.076	0.010	0.213	0.237	0.382
aTZ − aQZ	0.016	0.005	0.091	0.047	0.140
aQZ − a5Z	0.003	0.005	0.036	0.011	0.043

Since SAPT2+ calculations are computationally expensive, the practical size of the basis set has been limited to aug-cc-pVTZ for most molecule pairs in the S101 database. To obtain an estimate of the SAPT2+ energy at the CBS limit, extrapolation thus is necessary. As shown in Figure 2, the electrostatic and induction energy components converge quickly to the CBS limit (approximated by a5Z results). The mean unsigned errors between aTZ and a5Z of five pairs are 0.018 and 0.010 kcal/mol for the two components, respectively. Therefore, for electrostatic and induction energies, the results obtained with the aTZ basis set are considered a reasonable approximation of the CBS limit. For exchange−repulsion and dispersion energy, a two-point scheme was applied to extrapolate the energy calculated at aDZ and aTZ to the CBS limit.

SAPT2+ Estimation of the CBS Limit. As mentioned in the previous section, different energy components converge at different rates with respect to the basis set size. Electrostatic and induction energies calculated using the aTZ basis set are sufficiently converged, while the exchange−repulsion and dispersion energy terms are not. Thus, a two-point extrapolation scheme is used to extrapolate the exchange−repulsion and dispersion energy at aTZ basis set to the complete basis set limit. The extrapolated dispersion energy is further scaled by a factor of 0.89 to compensate for higher order dispersion terms missing in the SAPT2+ approach (refer to Methods for details of the extrapolation and scale factor determination). Then the total SAPT2+ interaction energy is obtained by summing up the individual energy components. Finally, the quality of the SAPT2+ interaction energy at different basis set levels was examined by comparing with the CCSD(T)/CBS interaction energy of the S66 data set (Figure 3).[37] In general, the SAPT2+/CBS estimates with scaled dispersion energy (which will be referred to as SAPT2+/CBS/scaled to distinguish from the SAPT2+/CBS values without scaling) have the smallest mean unsigned error (MUE) among tested combinations, given MUE values of 0.16, 0.68, 0.47, and 0.17 kcal/mol (and RMSE of 0.25, 0.80, 0.56, and 0.21 kcal/mol) for SAPT2+/CBS/scaled, SAPT2+/CBS, SAPT2+/aTZ, and SAPT2+/aDZ, respectively. It is not surprising that the SAPT2+/aDZ combination yields very reasonable total interaction energy as this is consistent with a previous report.[36] However, the small error in the total interaction energy of SAPT2+/aDZ is due to the error cancellation of individual energy components. To best estimate individual energy components, SAPT2+/CBS/scaled results remain the most accurate choice in this study. However, for the larger systems where aug-cc-pVTZ calculations are not practical, SAPT2+/aDZ may be considered as an alternative. For all 707 pairs in the S101×7 database reported in this study, the same strategy (SAPT2+/CBS/scaled) is applied to calculate the individual energy components. All SAPT data can be accessed from Table S4 of the Supporting Information.

DOI: 10.1021/acs.jctc.5b00267
J. Chem. Theory Comput. 2015, 11, 2609−2618

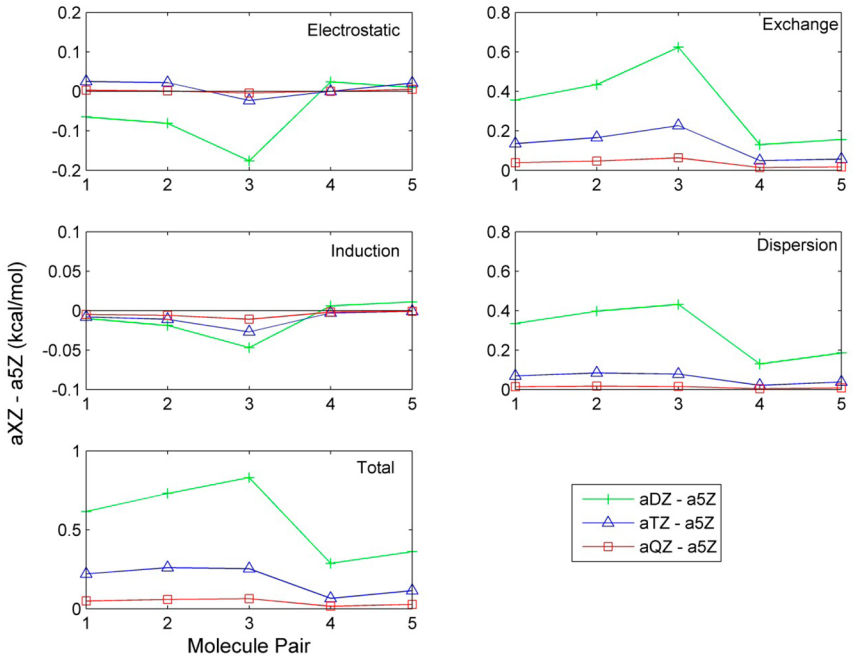


Figure 2. Difference between the SAPT2+ energy components calculated using aug-cc-pVXZ (X = D, T, and Q) basis set with the value obtained using aug-cc-pV5Z.

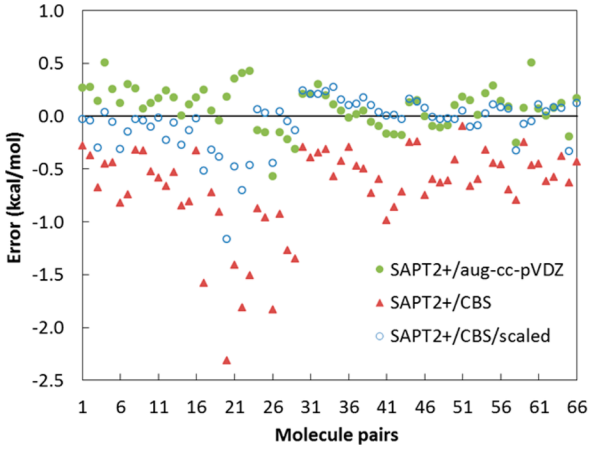


Figure 3. Errors of SAPT2+ interaction energy compared to CCSD(T)/CBS estimation[37] for dimers in the S66 data set. SAPT2+/aug-cc-pVDZ energy is shown in solid circles, SAPT2+/CBS is shown in triangles, and SAPT2+/CBS/scaled is shown in hollow circles.

Short-Range Electrostatic Interactions. The electrostatic interaction energy due to AMOEBA point multipoles as well as the charge penetration correction (multipoles + CP) are calculated and compared with SAPT2+ data for the S101×7 data set (excluding 7×7 complexes containing an ethyne molecule due to a lack of AMOEBA parameters) (Supporting Information Table S3). In our current model, the parameters Z and α for each atom are uniquely determined by the element type. Z is the number of valence electron. α is set equal to Z, or if Z is less than 2, then α is set to 2. The only parameter to be determined for the penetration correction is β in eq 3. For each of the H, C, N, and O elements, three atom types are used for β (Table 2). Hydrogen atoms are divided into nonpolar, aromatic, and polar hydrogens. Carbon, nitrogen, and oxygen all have three β values representing sp^3, sp^2, and aromatic cases.

Table 2. Two Sets of Charge Penetration Parameters for H, C, N, O, P, S, F, Cl, and Br

		valence-α set		fitted-α set	
atom type	Z	α (Å^{-1})	β (Å^{-1})	α (Å^{-1})	β (Å^{-1})
H (nonpolar)	1	2.0	1.999	3.3	2.924
H (aromatic)	1	2.0	2.010	3.3	3.064
H (polar, water)	1	2.0	2.004	3.3	3.178
C (sp^3)	4	4.0	2.646	3.8	2.934
C (aromatic)	4	4.0	2.708	3.8	2.764
C (sp^2)	4	4.0	2.685	3.8	2.673
N (sp^3)	5	5.0	3.097	3.1	2.790
N (aromatic)	5	5.0	3.072	3.1	2.784
N (sp^2)	5	5.0	3.054	3.1	2.761
O (sp^3, hydroxyl, water)	6	6.0	3.661	3.5	3.131
O (aromatic)	6	6.0	4.282	3.5	3.188
O (sp^2, carbonyl)	6	6.0	4.469	3.5	3.213
P (phosphate)	5	5.0	2.360	2.4	2.603
S (sulfide, e.g., R−SH)	6	6.0	2.770	2.6	2.382
S (sulfur IV, e.g., DMSO)	6	6.0	2.381	2.6	2.230
F (organofluorine)	7	7.0	4.275	4.2	4.030
Cl (organochloride)	7	7.0	2.830	3.0	2.594
Br (organobromine)	7	7.0	2.564	2.7	2.336

For sulfur, distinct β values are used for sulfide and sulfur IV, while P, F, Cl, and Br have only a single β value per element in current parametrization.

In general, after fitting of β parameters, the new electrostatic model with charge penetration correction shows excellent agreement with the SAPT2+ results (Figure 4). Taking the valence-α parameter set as an example, for dimers near the equilibrium distances (R_{min}), i.e., 0.90, 0.95, 1.00, 1.05, and 1.10 of R_{min}, the mean unsigned error (MUE) of the original point multipoles is 3.16 kcal/mol, which is reduced about 5-fold to 0.57 kcal/mol after inclusion of the charge penetration correction (Table 3). For the dimers at very short separation, i.e., 0.70 and 0.80 of R_{min}, the MUEs for the corrected and

DOI: 10.1021/acs.jctc.5b00267
J. Chem. Theory Comput. 2015, 11, 2609−2618

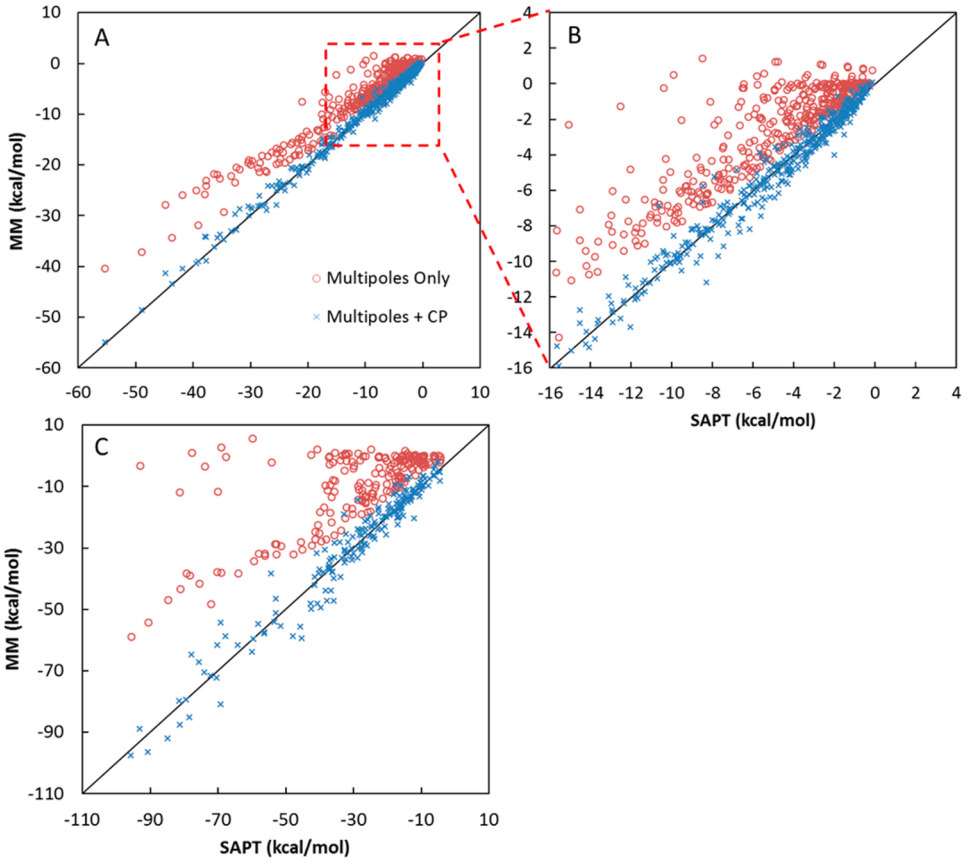


Figure 4. Plots of multipole electrostatic energy (kcal/mol) against the reference SAPT2+/aug-cc-pVTZ calculation for (A) near-equilibrium (0.90, 0.95, 1.00, 1.05, and 1.10) complexes taken from the S101×7 data set, (B) expanded plot of the boxed region in A, and (C) short-range (0.70 and 0.80) complexes in the S101×7 data set. The uncorrected AMOEBA point multipole energy (multipoles only) is shown in red circles, and the charge penetration corrected point multipole energies using the valence-α parameter set (multipoles + CP) are denoted by blue crosses.

Table 3. Differences between AMOEBA Electrostatic Energies, Either with or without Charge Penetration Correction, Compared against SAPT2+/CBS/Scaled Electrostatic Energies for the S101×7 Data Set

S101 set	statistics	multipoles only	multipoles + CP[a]	multipoles + CP[b]
R (0.90−1.10) (94×5 pairs)	MUE	3.16	0.57	0.72
	MSE	3.16	−0.04	−0.17
	RMSE	4.35	0.83	1.04
	% error	52.7%	13.6%	16.5%
R (0.70−0.80) (94×2 pairs)	MUE	19.16	3.28	2.84
	MSE	19.16	−0.15	0.46
	RMSE	24.17	4.63	4.36
	% error	69.3%	13.4%	10.8%
all distance (94×7 pairs)	MUE	7.73	1.35	1.33
	MSE	7.73	−0.07	0.01
	RMSE	13.43	2.57	2.49
	% error	57.4%	13.6%	14.9%

[a]Charge penetration corrected model using the valence-α parameter set. [b]Charge penetration corrected model using the fitted-α parameter set.

uncorrected electrostatic energy are 3.28 and 19.16 kcal/mol, respectively. As shown in Figure 4, it is striking that point-multipole-based electrostatic energy alone yields very large errors for dimers in close contact, and the simple charge penetration correction applied here is able to systematically improve agreement with SAPT-derived electrostatics. Based upon the mean unsigned errors, the charge penetration corrected model results in a percentage error of 13.6% and 13.4% at near-equilibrium and very short separations, respectively. In contrast, the uncorrected model has errors of 53% and 69% for these same two distance ranges. It is clear the charge penetration corrected model not only reduces the magnitude of absolute and relative errors compared to SAPT but also provides consistent performance over a range of distances. In the uncorrected model, the percentage of error at very short distances is larger than at near-equilibrium distances, due to the increased effect of short-ranged charge penetration.

For S66 dimers at near-equilibrium separations and using uncorrected AMOEBA multipoles, the hydrogen-bonded complexes exhibit the largest mean unsigned error of 4.41 kcal/mol, compared to MUEs of 3.08 and 2.08 kcal/mol for dispersion-dominant and mixed complexes (Table 4). This is not surprising since the hydrogen-bonded complexes generally have the strongest electrostatic interactions. However, in terms of relative errors, the dispersion-dominant complexes carry the largest error at 105%, while the hydrogen-bonded and mixed complexes have the mean percentage of errors of 30% and 58%, respectively. It is somewhat surprising the dispersion-dominant complexes have such absolute and relative errors, as they are normally considered to have the weakest electrostatic interaction among the three types.

DOI: 10.1021/acs.jctc.5b00267
J. Chem. Theory Comput. 2015, 11, 2609−2618

Table 4. Electrostatic Energy (kcal/mol) in Different Interaction Types of the S66 Complexes at the Near-Equilibrium (0.90−1.10) Distances in the S101×7 Data Set

	multipoles only		multipoles + CP[a]		multipoles + CP[b]	
S66 set	MUE	% error	MUE	% error	MUE	% error
hydrogen-bonded	4.41	30.5	0.50	3.5	0.47	3.3
dispersion-dominant	3.08	105.3	0.53	23.2	0.69	26.7
mixed	2.08	58.4	0.61	18.7	0.74	23.6

[a]Charge penetration corrected model using the valence-α parameter set. [b]Charge penetration corrected model using the fitted-α parameter set.

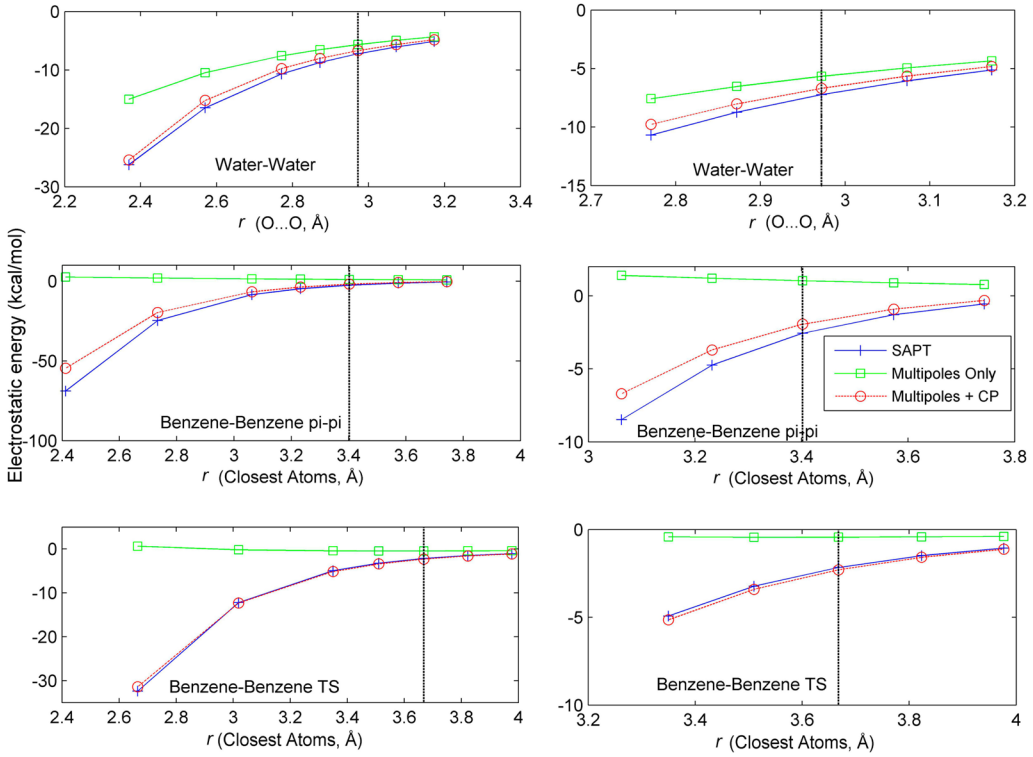


Figure 5. Plots of the electrostatic energy profiles of water−water and benzene−benzene dimer complexes. The valence-α parameter set was used in calculations of the charge penetration corrected model (multipoles + CP). The vertical line indicates the equilibrium distance.

To help understand why dispersion-dominant complexes have such large relative errors, the electrostatic energies of benzene dimers and $\pi-\pi$ stacked and T-shaped complexes, as well as hydrogen-bonded water dimers, are shown in Figure 5. For the uncorrected AMOEBA model, the calculated electrostatic energy is positive for the $\pi-\pi$ benzene pairs yet QM calculations suggest the interaction is attractive with a negative electrostatic energy. Taking the electrostatic energy for this pair at the equilibrium distance as an example, the SAPT2+/CBS/scaled calculation yields a value of −2.6 kcal/mol, while the uncorrected AMOEBA multipoles give +1.0 kcal/mol, an error of 3.6 kcal/mol or 138%. For the T-shaped benzene dimer, the SAPT2+/CBS/scaled and the uncorrected AMOEBA multipoles have values of −2.2 and −0.4 kcal/mol, respectively. The unsigned error is 1.8 kcal/mol or 82% of the SAPT values, both somewhat less than for the $\pi-\pi$ complex. These findings are consistent with the previous study by Tafipolsky and Engels.[15] In contrast, although the hydrogen-bonded water dimer has the larger electrostatic energy of −7.2 kcal/mol, the uncorrected model has an unsigned error of 1.6 kcal/mol and a relative error of only 22%. This trend is in accordance with the averaged errors reported in Table 4 and suggests the electrostatic interaction in dispersion-dominant complexes is the most charge penetration dependent. This might be explained by two effects. First, in the nonpolar molecules, the electron distribution is more "balanced"; i.e., there is more electron density on the hydrogen atoms, hence a stronger penetration effect for hydrogens. Second, in the stacked benzene dimer, interactions between heavier atoms, carbon−carbon for example, suffer stronger charge penetration effect, thus weight more in electrostatic energy. For hydrogen-bonded pairs, although the percentage of error is relatively low for the uncorrected atomic multipoles, the absolute error remains significant. Therefore, a correction is still necessary in order to achieve better accuracy in the force field. It is notable that, after the charge penetration correction, the mean unsigned errors of all three types of complexes are reduced to 0.5−0.6 kcal/mol near the equilibrium distances, which is approaching the possible error of the QM calculation itself.

Alternative Way To Derive the α Parameter. As mentioned in Methods, an alternative way to derive the α parameter is to fit the penetration-damped electrostatic potential (eq 4) to the target QM values. An attempt to use this fitting strategy has been also made, and the resulting parameters have been compared. The parametrization of α is restricted to a single unique value for each element type, as

DOI: 10.1021/acs.jctc.5b00267
J. Chem. Theory Comput. 2015, 11, 2609−2618

Table 5. Comparison of the Two Sets of Parameters, Valence-α and Fitted-α, for H, C, N, and O Containing Molecules in S66 Complexes[a]

S66 set	statistics (kcal/mol)	multipoles only	valence-α	fitted-α	valence-α; single β per element
R (0.90−1.10) (59×5 pairs)	MUE	3.34	0.54	0.62	0.65
	MSE	3.34	−0.04	−0.19	−0.00
	RMSE	4.45	0.74	0.83	0.88
R (0.70−0.80) (59×2 pairs)	MUE	21.34	2.72	2.39	3.52
	MSE	21.34	−0.10	0.41	−0.19
	RMSE	26.84	3.80	3.92	4.40
all distances (59×7 pairs)	MUE	8.48	1.16	1.12	1.47
	MSE	8.48	−0.06	−0.02	−0.06
	RMSE	14.83	2.13	2.21	2.47

[a]An additional set of parameters which has a unique β for each element is also presented.

before. A brute force scanning of the parameter using a grid size of 0.1 $Å^{-1}$ was used to search for the global minimum since the α parameter is less sensitive than β. All 13 monomers (excluding ethyne) in the S66 data set were used in fitting of the α for H, C, N, and O elements. Then β parameters were determined as before with α values fixed to their potential-fitted values. The penetration parameter set obtained this way will be referred to as the fitted-α set, while the parameter set with α based on eq 5 will be referred as the valence-α set. With the fitted-α parameters, the RMSE of the electrostatic potential of the 13 monomers calculated using eq 4 is greatly reduced to 0.07 kcal/mol, compared to an RMSE of 0.95 kcal/mol for the valence-α (see Supporting Information Table S2).

The performance of the two sets of parameters has been compared using the S101 data set. The overall performance of the two parameter sets is very similar to mean unsigned errors of 1.35 and 1.33 kcal/mol for the valence- and fitted-α sets, respectively (Table 3). For near-equilibrium pairs, the valence-α set has a marginally better MUE of 0.57 kcal/mol against 0.72 kcal/mol for the fitted-α set. For short-ranged pairs, the fitted-α set with a MUE of 2.84 kcal/mol yet is slightly better than a MUE of 3.28 kcal/mol of the valence-α set. Similar trends in RMSEs of the two sets of parameters are also observed. However, the valence-α parameter set tends to have more balanced performances for hydrogen-bonded, dispersion-dominant, and mixed complexes, giving the MUEs of 0.50, 0.53, and 0.61 kcal/mol for the three groups, respectively (Table 4). In contrast, the fitted-α set, with a MUE of 0.47 kcal/mol for the hydrogen-bonded complexes, shows subtly better agreement with SAPT results yet has slightly worse performances for the aromatic compounds. The MUEs of the dispersion-dominant and mixed complexes are 0.69 and 0.74 kcal/mol, respectively (electrostatic energy of individual pairs can be found in Table S3 in the Supporting Information). Nonetheless, the two sets of parameters all have excellent agreement with the SAPT results for the whole S101×7 database, while the fitted-α set yields better electrostatic potential than the valence-α set.

The charge penetration model also exhibited good transferability during the fitting of β parameters. Although three atom types are used for H, C, N, and O in the current parametrization, restriction to a single β for each element also results in reasonable accuracy. Simply applying the arithmetic mean of the three β parameters in valence-α parameter set for each element (Supporting Information Table S1) increases the MUE by only 0.1 to 0.65 kcal/mol for the near-equilibrium pairs in the S66 set (Table 5). For pairs with shorter distances, the MUE increases by 0.8 to 3.52 kcal/mol in the same set. Only marginal improvements in MUEs were found after optimizing the β parameters for each element starting from the averaged value. We believe this demonstrates the robustness and transferability of the charge penetration correction and the parametrization strategy. For the purpose of retaining flexibility, we recommend the use of three atom types for each of the H, C, N, and O elements in our final model.

CONCLUSION

The charge penetration effect is usually overlooked in molecular mechanical models and traditional force fields. Our results show that DMA-derived point multipoles systematically underestimate the SAPT electrostatic interaction energy at typical molecular interaction distances based on the 101×7 dimers studied here (see the Supporting Information). An exponential damping function providing a simple charge−charge penetration model suitable for force field incorporation has been revisited, along with a new parametrization strategy. The S101×7 SAPT-decomposed quantum mechanical energy database is developed as a reference for parameter training and for use in future force field comparison. The database is an extension of the S66 and S66×8 data set previously developed by Hobza and co-workers,[37] with additional prototype molecular complexes. The decomposed energies are calculated at the SAPT2+/aug-cc-pVTZ level of theory, with exchange−repulsion and dispersion components extrapolated to the complete basis set limit. The dispersion energy is further scaled to compensate for missing higher order terms in the SAPT2+ method. The total SAPT interaction energy is in excellent agreement with CCSD(T)/CBS results, which are currently considered to be the "gold standard" for estimation of intermolecular interactions, with a mean unsigned error of 0.16 kcal/mol for the S66 data set. Thus, the SAPT results should provide a reliable reference for force field development.

By replacing the idealized charge−charge (Coulomb) interaction with the charge penetration corrected model (eq 3) in the AMOEBA framework, the accuracy of calculated electrostatic energies for the S101×7 database is improved by 5-fold. For the five distance pairs near the equilibrium distances (i.e., 0.90−1.10 times the equilibrium distance), the mean unsigned error of the charge penetration corrected and uncorrected point multipole models are 0.57 and 3.16 kcal/mol, respectively; for the extremely close distance separations (i.e., 0.70 and 0.80 times the equilibrium distance), the mean unsigned errors of the two models are 3.28 and 19.16 kcal/mol, respectively. The improvement for the corrected model is significant and shows a consistent agreement with the quantum mechanics data at both long and short distances. The

DOI: 10.1021/acs.jctc.5b00267
J. Chem. Theory Comput. 2015, 11, 2609−2618

robustness and transferability of this model is also reflected in the use of very limited (element-based) parameters. The charge penetration correction is short-ranged and rapidly converges to the classical Coulomb interaction beyond 6−7 Å. Thus, it can be completely incorporated into the real space of Ewald summation without any additional computational cost in reciprocal space. Because simulations including penetration correction are clearly feasible, there is ongoing work dedicated to the optimization of parallel scaling the coupled penetration/smooth particle mesh Ewald approach. In addition, higher order penetration corrections (charge−dipole and charge−quadrupole penetration) are also possible and have been implemented in models such as SIBFA.[46] Nonetheless, the simple empirical charge penetration model presented in this work provides us with an efficient approach to achieve accurate electrostatic energy that is systematically modeled after SAPT quantum mechanical energy decomposition. The change in the electrostatic component requires re-examination of the van der Waals interaction to arrive at a balanced representation of the total energy. Overall we expect this improvement in the electrostatic component will alleviate the need for error compensation via other components and lead to more balanced and transferable potential energy functions in general. The comprehensive SAPT database we developed in this work will also be useful for many others who are interested in understanding intermolecular forces or evaluating different empirical models.

■ ASSOCIATED CONTENT

Supporting Information

Tables listing unique α and β for each element, RMSEs of the electrostatic potential on the grid compared to MP2/aug-cc-pVTZ calculation, interaction energies, and decomposed interaction energies. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00267.

■ AUTHOR INFORMATION

Corresponding Authors

*(P.R.) E-mail: pren@mail.utexas.edu.

*(J.-P.P.) E-mail: jpp@lct.jussieu.fr.

Funding

We are grateful for the support from the Robert A. Welch Foundation (Grant F-1691 to P.R.), the National Institutes of Health (Grants GM106137 and GM114237 to J.W.P. and P.R.), the National Science Foundation (Grant CHE1152823 to J.W.P.), and CPRIT (Grant RP110532 to Q.W.). The high performance computing resources were provided by TACC and XSEDE (Grant TG-MCB100057 to P.R.). This work was also supported in part by French state funds managed by CALSIMLAB and the ANR within the Investissements d'Avenir program under reference ANR-11-IDEX-0004-02 (to J.-P.P).

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

CBS, complete basis set; DMA, distributed multipole analysis; MSE, mean signed error; MUE, mean unsigned error; PME, particle mesh Ewald; QM/MM, combined quantum and molecular mechanics method; RMSE, root-mean-square error; SAPT, symmetry-adapted perturbation theory

■ REFERENCES

(1) Ewald, P. P. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys. (Berlin, Ger.)* **1921**, *369* (3), 253−287.

(2) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98* (12), 10089−10092.

(3) Darden, T.; Perera, L.; Li, L.; Pedersen, L. New tricks for modelers from the crystallography toolkit: The particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure* **1999**, *7* (3), R55−R60.

(4) Smith, W. Point multipoles in the Ewald summation (revisited). *CCP5 Newslett.* **1998**, *46*, 18−30.

(5) Nam, K.; Gao, J.; York, D. M. An Efficient Linear-Scaling Ewald Method for Long-Range Electrostatic Interactions in Combined QM/MM Calculations. *J. Chem. Theory Comput.* **2004**, *1* (1), 2−13.

(6) Ren, P.; Ponder, J. W. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **2003**, *107* (24), 5933−5947.

(7) Walker, R. C.; Crowley, M. F.; Case, D. A. The implementation of a fast and accurate QM/MM potential method in Amber. *J. Comput. Chem.* **2008**, *29* (7), 1019−1031.

(8) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114* (8), 2549−2564.

(9) Ren, P.; Wu, C.; Ponder, J. W. Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules. *J. Chem. Theory Comput.* **2011**, *7* (10), 3143−3161.

(10) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins. *J. Chem. Theory Comput.* **2013**, *9* (9), 4046−4063.

(11) Stone, A. J. *The Theory of Intermolecular Forces*. Oxford University Press: New York, 1996.

(12) Watt, M.; Hardebeck, L. K. E.; Kirkpatrick, C. C.; Lewis, M. Face-to-Face Arene−Arene Binding Energies: Dominated by Dispersion but Predicted by Electrostatic and Dispersion/Polarizability Substituent Constants. *J. Am. Chem. Soc.* **2011**, *133* (11), 3854−3862.

(13) Hunter, C. A.; Sanders, J. K. M. The nature of .pi.-.pi. interactions. *J. Am. Chem. Soc.* **1990**, *112* (14), 5525−5534.

(14) Hohenstein, E. G.; Duan, J. N.; Sherrill, C. D. Origin of the Surprising Enhancement of Electrostatic Energies by Electron-Donating Substituents in Substituted Sandwich Benzene Dimers. *J. Am. Chem. Soc.* **2011**, *133* (34), 13244−13247.

(15) Tafipolsky, M.; Engels, B. Accurate Intermolecular Potentials with Physically Grounded Electrostatics. *J. Chem. Theory Comput.* **2011**, *7* (6), 1791−1803.

(16) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *94* (7), 1887−1930.

(17) Mu, X.; Wang, Q.; Wang, L.-P.; Fried, S. D.; Piquemal, J.-P.; Dalby, K. N.; Ren, P. Modeling Organochlorine Compounds and the σ-Hole Effect Using a Polarizable Multipole Force Field. *J. Phys. Chem. B* **2014**, *118* (24), 6456−6465.

(18) Chipman, D. M. Charge penetration in dielectric models of solvation. *J. Chem. Phys.* **1997**, *106* (24), 10194−10206.

(19) Chipman, D. M. Reaction field treatment of charge penetration. *J. Chem. Phys.* **2000**, *112* (13), 5558−5565.

(20) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Polarizable dielectric model of solvation with inclusion of charge penetration effects. *J. Chem. Phys.* **2001**, *114* (13), 5691−5701.

(21) Freitag, M. A.; Gordon, M. S.; Jensen, J. H.; Stevens, W. J. Evaluation of charge penetration between distributed multipolar expansions. *J. Chem. Phys.* **2000**, *112* (17), 7300−7306.

(22) Piquemal, J. P.; Gresh, N.; Giessner-Prettre, C. Improved formulas for the calculation of the electrostatic contribution to the intermolecular interaction energy from multipolar expansion of the electronic distribution. *J. Phys. Chem. A* **2003**, *107* (48), 10353−10359.

DOI: 10.1021/acs.jctc.5b00267
J. Chem. Theory Comput. 2015, 11, 2609−2618

(23) Gresh, N.; Piquemal, J. P.; Krauss, M. Representation of Zn(II) complexes in polarizable molecular mechanics. Further refinements of the electrostatic and short-range contributions. Comparisons with parallel ab initio computations. *J. Comput. Chem.* **2005**, *26* (11), 1113−1130.
(24) Slipchenko, L. V.; Gordon, M. S. Electrostatic energy in the effective fragment potential method: Theory and application to benzene dimer. *J. Comput. Chem.* **2007**, *28* (1), 276−291.
(25) Werneck, A. S.; Filho, T. M. R.; Dardenne, L. E. General methodology to optimize damping functions to account for charge penetration effects in electrostatic calculations using multicentered multipolar expansions. *J. Phys. Chem. A* **2008**, *112* (2), 268−280.
(26) Kumar, R.; Wang, F. F.; Jenness, G. R.; Jordan, K. D. A second generation distributed point polarizable water model. *J. Chem. Phys.* **2010**, *132* (1), No. 014309.
(27) Stone, A. J. Electrostatic Damping Functions and the Penetration Energy. *J. Phys. Chem. A* **2011**, *115* (25), 7017−7027.
(28) Spackman, M. A. The use of the promolecular charge density to approximate the penetration contribution to intermolecular electrostatic energies. *Chem. Phys. Lett.* **2006**, *418* (1−3), 158−162.
(29) Cisneros, G. A. Application of Gaussian Electrostatic Model (GEM) Distributed Multipoles in the AMOEBA Force Field. *J. Chem. Theory Comput.* **2012**, *8* (12), 5072−5080.
(30) Wang, B.; Truhlar, D. G. Screened Electrostatic Interactions in Molecular Mechanics. *J. Chem. Theory Comput.* **2014**, *10* (10), 4480−4487.
(31) Cisneros, G. A.; Tholander, S. N. I.; Parisel, O.; Darden, T. A.; Elking, D.; Perera, L.; Piquemal, J. P. Simple formulas for improved point-charge electrostatics in classical force fields and hybrid quantum mechanical/molecular mechanical embedding. *Int. J. Quantum Chem.* **2008**, *108* (11), 1905−1912.
(32) Wang, B.; Truhlar, D. G. Including Charge Penetration Effects in Molecular Modeling. *J. Chem. Theory Comput.* **2010**, *6* (11), 3330−3342.
(33) Wang, B.; Truhlar, D. G. Partial Atomic Charges and Screened Charge Models of the Electrostatic Potential. *J. Chem. Theory Comput.* **2012**, *8* (6), 1989−1998.
(34) Wang, Q. T.; Bryce, R. A. Improved Hydrogen Bonding at the NDDO-Type Semiempirical Quantum Mechanical/Molecular Mechanical Interface. *J. Chem. Theory Comput.* **2009**, *5* (9), 2206−2211.
(35) Hohenstein, E. G.; Sherrill, C. D. Density fitting of intramonomer correlation effects in symmetry-adapted perturbation theory. *J. Chem. Phys.* **2010**, *133* (1), No. 014101.
(36) Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. Levels of symmetry adapted perturbation theory (SAPT). I. Efficiency and performance for interaction energies. *J. Chem. Phys.* **2014**, *140* (9), No. 094106.
(37) Rezac, J.; Riley, K. E.; Hobza, P. S66: A Well-Balanced Database of Benchmark Interaction Energies Relevant to Biomolecular Structures. *J. Chem. Theory Comput.* **2011**, *7* (8), 2427−2438.
(38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian: Wallingford, CT, USA, 2009.
(39) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. Psi4: An open-source ab initio electronic structure program. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2* (4), 556−565.
(40) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90* (2), 1007−1023.
(41) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96* (9), 6796−6806.
(42) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-set convergence in correlated calculations on Ne, N2, and H2O. *Chem. Phys. Lett.* **1998**, *286* (3−4), 243−252.
(43) Rezac, J.; Hobza, P. Extrapolation and Scaling of the DFT-SAPT Interaction Energies toward the Basis Set Limit. *J. Chem. Theory Comput.* **2011**, *7* (3), 685−689.
(44) Wu, J.; Chattree, G.; Ren, P. Automation of AMOEBA polarizable force field parameterization for small molecules. *Theor. Chem. Acc.* **2012**, *131* (3), 1−11.
(45) Stone, A. J.; Alderton, M. Distributed multipole analysis. *Mol. Phys.* **1985**, *56* (5), 1047−1064.
(46) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J. P. Anisotropic, polarizable molecular mechanics studies of inter- and intramoecular interactions and ligand-macromolecule complexes. A bottom-up strategy. *J. Chem. Theory Comput.* **2007**, *3* (6), 1960−1986.

DOI: 10.1021/acs.jctc.5b00267
J. Chem. Theory Comput. 2015, 11, 2609−2618