JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

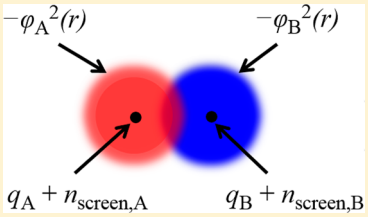
Screened Electrostatic Interactions in Molecular Mechanics

Bo Wang and Donald G. Truhlar*

Department of Chemistry, Chemical Theory Center, and Supercomputing Institute, University of Minnesota, 207 Pleasant Street SE, Minneapolis, Minnesota 55455-0431, United States

Supporting Information

ABSTRACT: In a typical application of molecular mechanics (MM), the electrostatic interactions are calculated from parametrized partial atomic charges treated as point charges interacting by radial Coulomb potentials. This does not usually yield accurate electrostatic interactions at van der Waals distances, but this is compensated by additional parametrized terms, for example Lennard-Jones potentials. In the present work, we present a scheme involving radial screened Coulomb potentials that reproduces the accurate electrostatics much more accurately. The screening accounts for charge penetration of one subsystem's charge cloud into that of another subsystem, and it is incorporated into the interaction potential in a way similar to what we proposed in a previous article (*J. Chem. Theory Comput.* **2010**, *6*, 3330) for combined quantum mechanical and molecular mechanical (QM/MM) simulations, but the screening parameters are reoptimized for MM. The optimization is carried out with electrostatic-potential-fitted partial atomic charges, but the optimized parameters should be useful with any realistic charge model. In the model we employ, the charge density of an atom is approximated as the sum of a point charge representing the nucleus and inner electrons and a smeared charge representing the outermost electrons; in particular, for all atoms except hydrogens, the smeared charge represents the two outermost electrons in the present model. We find that the charge penetration effect can cause very significant deviations from the popular point-charge model, and by comparison to electrostatic interactions calculated by symmetry-adapted perturbation theory, we find that the present results are considerably more accurate than point-charge electrostatic interactions. The mean unsigned error in electrostatics for a large and diverse data set (192 interaction energies) decreases from 9.2 to 3.3 kcal/mol, and the error in the electrostatics for 10 water dimers decreases from 1.7 to 0.5 kcal/mol. We could have decreased the average errors further, but at the cost of sometimes significantly overestimating the screening; instead we chose a more conservative (safer) parametrization that systematically underestimates the screening (which by definition means it improves over point charges) and only occasionally overestimates it. Despite this conservative choice, we find that the screened MM method is even more accurate for the electrostatics than unscreened QM/MM calculations. This new method is easy to implement in any MM program, and it can be used to develop more physical force fields for molecular simulations.



1. INTRODUCTION

Molecular mechanics (MM) is useful for many calculations on large systems, especially when a long simulation time is required or where electronic structure calculations of the potential energy functions are not practically affordable for the whole system. Molecular mechanics in its conventional form does not include bond breaking or intermediate stages of bond rearrangement, and it may be used for the nonreactive part of a system or, when chemical reactions are not involved, for the whole system. The force fields usually contain several terms, nominally corresponding to identifiable physical effects such as electrostatics, exchange repulsion, and dispersion. Sometimes induction is included as one or more separate terms, and sometimes it is absorbed in an average way into other terms. Intermolecular charge transfer is not usually included and must be ignored or compensated by empirical optimization of other terms.

The parametrization of MM force fields can be based on experimental data, electronic structure calculations, or both.[1−15] In practice, the individual terms are often not meaningful on their own; for example, there may be systematic errors in the electrostatics that are compensated to some extent by the fitting of van der Waals parameters, and the parameters may make up in an average way for neglect of charge transfer. The electrostatic interactions are sometimes intentionally overestimated to make up for neglected induction or for an increase in partial atomic charges due to solvent polarization. Fitting MM parameters such that individual terms are physically meaningful can improve the transferability of the force field.[16−20] Some widely used energy decomposition methods include Morokuma−Kitaura decomposition,[21,22] constrained space orbital variation analysis (CSOV),[23,24] reduced variational space (RVS),[25] natural energy decomposition analysis,[26] the extended transition state method with the natural orbitals for chemical valence (ETS-NOCV),[27] and symmetry-adapted perturbation theory (SAPT).[28,29] Although these decompositions are not completely unambiguous, the static electrostatic term is straightforward and is given unambiguously by, for example, the SAPT or ETS-NOCV schemes. The present article presents a method for improving the physical meaningfulness of the electrostatic terms in MM force fields so that they correspond better to the electrostatics calculated by SAPT.

Received: June 13, 2014

   dx.doi.org/10.1021/ct5005142 | *J. Chem. Theory Comput.* XXXX, XXX, XXX−XXX

In force field development, an accurate description of electrostatic interactions is crucial since electrostatic interactions often make the largest contribution to the total interaction energy. For simplicity, most force fields use the monopole approximation, which may also be called the point-charge model, and the point charges are usually (but not always) placed at the nuclei. Placing point charges also at off-nuclear positions can lead to improved accuracy, but it greatly increases the complexity for flexible molecules. Including higher-order multipole contributions can also improve the treatment of electrostatic interactions, but it too increases the complexity considerably; furthermore, models including higher multipoles are sometimes sensitive to the way the multipole terms are damped at short and medium ranges where the multipole series is no longer applicable and the higher-order asymptotic multipole terms diverge strongly in both single-center expansions and distributed multipole expansions.[30] One important contribution that is missing in the multipole expansion is the charge penetration effect. Charge penetration occurs when the electron densities of two molecules overlap; then the valence electrons of one center do not screen the nucleus of that center from interacting partners as fully as they do in the nonoverlap regions.

The charge penetration effect between molecules can be significant even at the equilibrium geometries of van der Waals molecules (the region of geometry near a van der Waals minimum is called the medium range); the importance of overlap at van der Waals geometries can be appreciated by the following consideration. An equilibrium van der Waals geometry is a geometry where attractive forces precisely balance the repulsive force, and the exchange repulsion that dominates the repulsive force is a result of overlap. Therefore, we may say that the charge penetration effects are of both short and medium range. Including the charge penetration effect can significantly improve the description of the electrostatics in van der Waals interactions, and various procedures have been proposed to include this effect in molecular modeling.[17,31−42] For example, the charge penetration effects may be described by Gaussian-type functions or damping functions, and accurate electron densities, electrostatic potentials, and electrostatic interaction energies may be used for fitting parameters in these functions. The charge penetration effect has already been included in the development of some new force fields.[18,43,44]

In a previous study,[39] we have developed a screened charge model for use in the QM−MM interaction terms of QM/MM calculations. In this method, the electron densities of atoms in the MM subsystem are approximated as screened charges with spherical charge distributions. The electrostatic energy from SAPT was used for fitting the parameters of the screened charges. In the present study, we derive the formulas to extend this screened charge model to pure molecular mechanics. In section 2, we present the screened charge model and give the formulas for its use in MM. Section 3 has the computational details. Section 4 describes the optimization of screening parameters and the validation. Section 5 shows the performance of the screened charge model for the electrostatic interactions. Section 6 summarizes the conclusions.

2. THEORY

In previous work,[39] we developed a screened charge model called outer density screening (ODS). In ODS, the electron density of an atom is partitioned into a point charge representing the nucleus and the electrons closest to the nucleus and a smeared charge representing n_{screen} electrons in the outer region. The spherical Slater-type orbital (STO) whose absolute square represents the smeared charge is given by

$$\phi = Br^{n-1}\exp(-\zeta r) \tag{1}$$

where B normalizes φ^2 to n_{screen}, r is the distance of the electron from the nucleus, n is the highest principal quantum number of the element, and ζ is the exponential parameter of the STO. The point charge scheme is the unscreened scheme, and it corresponds to the limit as $\zeta \to \infty$. The number n_{screen} of outer electrons was set to 1 for nonmetals except hydrogen, to 0 for metals, and to $1 - q$ for hydrogen, where q is the partial atomic charge on hydrogen. We optimized the ζ parameters for 10 elements (H, C, N, O, F, Si, P, S, Cl, and Br),[39] and we recommended that ζ parameters for other nonmetals with $Z \le 36$ be set equal to half of the Strand and Bonham exponential parameters for the outermost component of the electron density.[45] (Since ζ parameters are not needed for metals in the ODS scheme, this means that QM/MM ODS parameters are available for all elements except Sb, Te, I, Po, At, and noble gases.)

In the present paper, we optimize another set of parameters—this time for use in screened MM calculations. In this proof-of-concept study, we implement and test the scheme only for molecules containing the elements H, C, N, O, F, Si, P, S, and Cl. For hydrogen, we keep n_{screen} equal to $1 - q$, where q is partial atomic charge on the H, and we take its optimized ζ value from the QM/MM study, which optimized it as 1.32 a_0^{-1}, where a_0 is the Bohr radius (1 a_0 = 0.529177 Å). For C, N, O, F, Si, P, S, and Cl, we set n_{screen} to 2, rather than the previously used 1, and we optimize the ζ parameters for these elements using 202 interaction energies (192 in stage 1 and 10 more added in stage 2 as specified below). The choice of n_{screen} = 2 is quite reasonable since there are two electrons in the HOMO, which is the orbital most readily penetrated, in a closed shell-molecule, and the present treatment is designed to be very conservative, so we only screen electrons in the outermost orbital. The reason why the smaller value of 1 might be a better choice in QM/MM calculations is that when more electrons are included in the screening region of a QM/MM calculation, the screened charges, which are incorporated in one-electron integrals in the QM Hamiltonian, can overpolarize the QM region and cause unphysical charge flow in the absence of exchange repulsion potentials. Therefore, our parametrization involves a compromise between making the most complete possible treatment and making a simple model that gives improved electrostatics without making the formulation very complicated or expensive. In the MM calculations, there is no such concern, and we can place more electrons in the screening region, which is more physical, and which we found to yield better performance in MM calculations. (A comparison is shown in the Supporting Information.) To differentiate the parameters used in the original QM/MM paper from those optimized here, we label the new set of parameters MMODS, while the old set of parameters is still called ODS.

In order to use the screened charges in molecular mechanics, one needs to evaluate the electrostatic energy between two screened atoms A and B in the following equation.

$$E = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \frac{\rho_A(\mathbf{r}_1)\,\rho_B(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{2}$$

where ρ_A and ρ_B are charge densities for atoms A and B, and $\mathbf{r}_i$ denotes the position of electron i. A schematic figure of

dx.doi.org/10.1021/ct5005142 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

interactions between screened charges A and B is shown in Figure 1. The detailed evaluation of this expression is given in the next section.

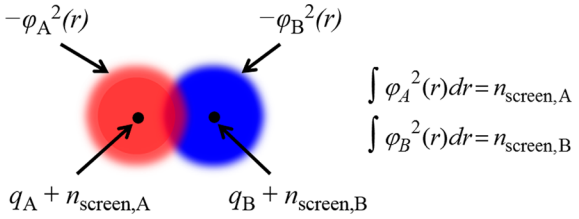


Figure 1. Schematic diagram of the interaction between two screened charge distributions, where the screening is done by the molecular mechanics outer density screening (MMODS) algorithm.

3. COMPUTATIONAL DETAILS

Here, we provide the detailed formulas for the evaluation of MM energy between two screened charges. The charge density of a screened charge A is written as

$$\rho_A = (q_A + n_{\text{screen,A}})\,\delta(\mathbf{r}_1 - \mathbf{r}_A) - n_{\text{screen,A}} \times [C_1 |\mathbf{r}_1 - \mathbf{r}_A|^{n_A - 1} e^{-\zeta_A |\mathbf{r}_1 - \mathbf{r}_A|}]^2 \quad (3)$$

and the Coulomb interactions between two screened charges A and B are written as

$$E = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \{(q_A + n_{\text{screen,A}})\,\delta(\mathbf{r}_1 - \mathbf{r}_A) - n_{\text{screen,A}} [C_1 |\mathbf{r}_1 - \mathbf{r}_A|^{n_A - 1} e^{-\zeta_A |\mathbf{r}_1 - \mathbf{r}_A|}]^2\} \times \{(q_B + n_{\text{screen,B}})\,\delta(\mathbf{r}_2 - \mathbf{r}_B) - n_{\text{screen,B}} \times [C_2 |\mathbf{r}_2 - \mathbf{r}_B|^{n_B - 1} e^{-\zeta_B |\mathbf{r}_2 - \mathbf{r}_B|}]^2\} \quad (4)$$

where C_1 and C_2 are the normalization constants, $\mathbf{r}_A$ and $\mathbf{r}_B$ are the postions of the nuclei of atoms A and B, n_A and n_B are the highest principal quantum numbers of atoms A and B, and ζ_A and ζ_B are the exponential parameter of the STOs for atoms A and B.

This expression is further separated into four terms:

$$E = E_1(\text{core}_A - \text{core}_B) + E_2(\text{core}_A - n_B) + E_3(n_A - \text{core}_B) + E_4(n_A, n_B) \quad (5)$$

where core_A and core_B represent the point charge contributions on atoms A and B, and n_A and n_B represent the smeared charge distributions of screened charges A and B with principal quantum numbers n_A and n_B, respectively. The first term is the point charge−point charge interaction. The second and third terms are the interactions between a point charge and a smeared charge of a given principal quantum number, which can be evaluated by integrating the smeared charge, precisely the same as in the already published QM/MM scheme.[39] The only new issue is the fourth term, which is written as follows:

$$E_4(n_A, n_B) = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} n_{\text{screen,A}} n_{\text{screen,B}} \times [C_1 |\mathbf{r}_1 - \mathbf{r}_A|^{n_A - 1} e^{-\zeta_A |\mathbf{r}_1 - \mathbf{r}_A|}]^2 [C_2 |\mathbf{r}_2 - \mathbf{r}_B|^{n_B - 1} e^{-\zeta_B |\mathbf{r}_2 - \mathbf{r}_B|}]^2 \quad (6)$$

where n_A and n_B are the highest principal quantum numbers of the atoms A and B.

Many methods have been proposed to evaluate the Coulomb integrals between two Slater-type orbitals,[46−49] and from that work we can obtain formulas for the first, second, and third row elements as follows. We define

$$s_A = \zeta_A R$$
$$s_B = \zeta_B R$$
$$\kappa = (\zeta_A{}^2 + \zeta_B{}^2)/(\zeta_A{}^2 - \zeta_B{}^2) \quad (7)$$

where R is the distance between atoms A and B. First, consider the case where s_A equals s_B, and let $s = s_A = s_B$.

$$E_4(1, 1) = n_{\text{screen,A}} n_{\text{screen,B}} (1/R)[1 - (1 + 11/8s + 3/4s^2 + 1/6s^3) \exp(-2s)]$$

$$E_4(2, 2) = n_{\text{screen,A}} n_{\text{screen,B}} (1/R)[1 - (1 + 419/256s + 163/128s^2 + 119/192s^3 + 5/24s^4 + 1/20s^5 + 1/120s^6 + 1/1260s^7) \times \exp(-2s)]$$

$$E_4(2, 3) = n_{\text{screen,A}} n_{\text{screen,B}} (1/R)[1 - (1 + 2621/1536s + 1085/768s^2 + 1733/2304s^3 + 331/1152s^4 + 239/2880s^5 + 1/54s^6 + 1/315s^7 + 1/2520s^8 + 1/34020s^9) \exp(-2s)]$$

$$E_4(3, 3) = n_{\text{screen,A}} n_{\text{screen,B}} (1/R)[1 - (1 + 5351/3072s + 765/512s^2 + 631/768s^3 + 191/576s^4 + 179/1728s^5 + 7/270s^6 + 1/189s^7 + 1/1134s^8 + 1/8505s^9 + 1/85050s^{10} + 1/1403325s^{11}) \exp(-2s)] \quad (8)$$

dx.doi.org/10.1021/ct5005142 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Next, consider the case where $s_A \neq s_B$; then

$$E_4(1,1) = n_{\mathrm{screen,A}} n_{\mathrm{screen,B}}(1/R)\{1 - (1-\kappa)^2[1/4(2+\kappa) + 1/4s_A]\exp(-2s_A) - (1+\kappa)^2[1/4(2-\kappa) + 1/4s_B]\exp(-2s_B)\}$$

$$\begin{aligned} E_4(1,2) = n_{\mathrm{screen,A}} n_{\mathrm{screen,B}}(1/R)\{&1 - (1-\kappa)^3[1/16(1 - 5\kappa - 4\kappa^2) - 1/8\kappa s_A]\exp(-2s_A) \\ &- (1+\kappa)^2[1/16(15 - 22\kappa + 15\kappa^2 - 4\kappa^3) + 3/8(3 - 3\kappa + \kappa^2)s_B + 1/4(2-\kappa){s_B}^2 + (1/12)\kappa {s_B}^3]\exp(-2s_B)\} \end{aligned}$$

$$\begin{aligned} E_4(2,2) = n_{\mathrm{screen,A}} n_{\mathrm{screen,B}}(1/R)\{&1 - (1-\kappa)^3[1/16(8 - \kappa - 27\kappa^2 - 30\kappa^3 - 10\kappa^4) + 1/32(11 - 19\kappa - 44\kappa^2 - 20\kappa^3)s_A \\ &+ 1/16(1 - 5\kappa - 4\kappa^2){s_A}^2 - (1/24)\kappa {s_A}^3)]\exp(-2s_A) - (1+\kappa)^3[1/16(8 + \kappa - 27\kappa^2 + 30\kappa^3 - 10\kappa^4) \\ &+ 1/32(11 + 19\kappa - 44\kappa^2 + 20\kappa^3)s_B + 1/16(1 + 5\kappa - 4\kappa^2){s_B}^2 + (1/24)\kappa {s_B}^3]\exp(-2s_B)\} \end{aligned}$$

$$\begin{aligned} E_4(1,3) = n_{\mathrm{screen,A}} n_{\mathrm{screen,B}}(1/R)\{&1 - (1-\kappa)^4[1/16(-1 - 2\kappa + 4\kappa^2 + 4\kappa^3) + 1/48(-1 + 4\kappa^2)s_A]\exp(-2s_A) \\ &- (1+\kappa)^2[1/16(17 - 36\kappa + 49\kappa^2 - 42\kappa^3 + 20\kappa^4 - 4\kappa^5) + 1/48(83 - 154\kappa + 159\kappa^2 - 88\kappa^3 + 20\kappa^4)s_B \\ &+ 1/12(15 - 22\kappa + 15\kappa^2 - 4\kappa^3){s_B}^2 + 1/6(3 - 3\kappa + \kappa^2){s_B}^3 + 1/18(2-\kappa){s_B}^4 + 1/90{s_B}^5]\exp(-2s_B)\} \end{aligned}$$

$$\begin{aligned} E_4(2,3) = n_{\mathrm{screen,A}} n_{\mathrm{screen,B}}(1/R)\{&1 - (1-\kappa)^4[1/96(-7 - 120\kappa - 119\kappa^2 + 170\kappa^3 + 294\kappa^4 + 112\kappa^5) \\ &+ 1/32(-5 - 26\kappa + 6\kappa^2 + 52\kappa^3 + 28\kappa^4)s_A + 1/16(-1 - 2\kappa + 4\kappa^2 + 4\kappa^3){s_A}^2 + 1/144(-1 + 4\kappa^2){s_A}^3)]\exp(-2s_A) \\ &- (1+\kappa)^3[1/96(103 - 217\kappa + 23\kappa^2 + 525\kappa^3 - 800\kappa^4 + 490\kappa^5 - 112\kappa^6) \\ &+ 5/48(13 - 14\kappa - 26\kappa^2 + 67\kappa^3 - 52\kappa^4 + 14\kappa^5)s_B + 1/12(8 + \kappa - 27\kappa^2 + 30\kappa^3 - 10\kappa^4){s_B}^2 \\ &+ 1/72(11 + 19\kappa - 44\kappa^2 + 20\kappa^3){s_B}^3 + 1/72(1 + 5\kappa - 4\kappa^2){s_B}^4 + 1/180\kappa {s_B}^5)]\exp(-2s_B)\} \end{aligned}$$

$$\begin{aligned} E_4(3,3) = n_{\mathrm{screen,A}} n_{\mathrm{screen,B}}(1/R)\{&1 - (1-\kappa)^4[1/96(48 - 137\kappa - 836\kappa^2 - 993\kappa^3 + 448\kappa^4 + 1792\kappa^5 + 1344\kappa^6 + 336\kappa^7) \\ &+ 1/96(19 - 290\kappa - 727\kappa^2 - 120\kappa^3 + 1092\kappa^4 + 1120\kappa^5 + 336\kappa^6)s_A \\ &+ 1/72(-7 - 120\kappa - 119\kappa^2 + 170\kappa^3 + 294\kappa^4 + 112\kappa^5){s_A}^2 + 1/72(-5 - 26\kappa + 6\kappa^2 + 52\kappa^3 + 28\kappa^4){s_A}^3 \\ &+ 1/72(-1 - 2\kappa + 4\kappa^2 + 4\kappa^3){s_A}^4 + 1/1080(-1 + 4\kappa^2){s_A}^5)]\exp(-2s_A) \\ &- (1+\kappa)^4[1/96(48 + 137\kappa - 836\kappa^2 + 993\kappa^3 + 448\kappa^4 - 1792\kappa^5 + 1344\kappa^6 - 336\kappa^7) \\ &+ 1/96(19 + 290\kappa - 727\kappa^2 + 120\kappa^3 + 1092\kappa^4 - 1120\kappa^5 + 336\kappa^6)s_B \\ &+ 1/72(-7 + 120\kappa - 119\kappa^2 - 170\kappa^3 + 294\kappa^4 - 112\kappa^5){s_B}^2 + 1/72(-5 + 26\kappa + 6\kappa^2 - 52\kappa^3 + 28\kappa^4){s_B}^3 \\ &+ 1/72(-1 + 2\kappa + 4\kappa^2 - 4\kappa^3){s_B}^4 + 1/1080(-1 + 4\kappa^2){s_B}^5)]\exp(-2s_B)\} \end{aligned} \tag{9}$$

Equations 8 and 9 can be derived from eqs 34 and 34a in ref 46, eq 30 in ref 48, and pages 89 and 102 in ref 49. *Mathematica*[50] was used to simplify the equations. We do not show $E_4(1,2)$ and $E_4(1,3)$ for $s = s_A = s_B$ because these terms are not needed.

4. PARAMETRIZATION

The parametrization proceeds in two stages. In stage one, we use the same test set as used in ref 39, but only include the molecules that contain H, C, N, O, F, Si, P, S, and Cl. This test set contains 37 dimers, each calculated with one or two basis sets (aug-cc-pVTZ[51,52] and def2-TZVP[53]) and three kinds of geometries, as we used before. The 37 molecules in the test suite are listed in Table 1. The geometry that was originally called equilibrium geometry is now called standard geometry (because actually three of those geometries are not optimized geometries, as explained previously[39]). We do calculations with the aug-cc-pVTZ basis set for only 27 out of 37 molecules in the test suite, and only with def2-TZVP for the other 10. There are $(27 + 37) \times 3 = 192$ test cases in stage 1.

The electrostatic energy calculated by point MM charges or screened MM charges is called $E_{\mathrm{elst}}^{\mathrm{MM}}$, and the static electrostatic energy from SAPT is called $E_{\mathrm{elst}}^{(10)}$. The mean unsigned error (MUE) over the test suite is calculated by

$$\mathrm{MUE} = \sum_{B=1}^{1\text{ or }2}\sum_{M=1}^{\text{molecules}}\sum_{G=1}^{3}(|E_{\mathrm{elst}}^{\mathrm{MM}}(\mathrm{MM};B,M,G) - E_{\mathrm{elst}}^{(10)}(\mathrm{SAPT};B,M,G)|) \tag{10}$$

dx.doi.org/10.1021/ct5005142 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 1. 37 Dimers in the Test Suite[a]

$NH_3\cdots NH_3$	$H_2O\cdots H_2O$	HF···HF	$H_2S\cdots H_2S$	HCl···HCl
$CH_3CN\cdots CH_3CN$	$CH_3OH\cdots CH_3OH$	$(CH_3)_2CO\cdots(CH_3)_2CO$	$(CH_3)_2SO\cdots(CH_3)_2SO$ I	$(CH_3)_2SO\cdots(CH_3)_2SO$ II
$C_2H_4\cdots F_2$	$C_2H_2\cdots ClF$	$NH_3\cdots H_2O$	$NH_3\cdots HF$ I	$NH_3\cdots H_2S$
$NH_3\cdots HCl$ I	$NH_3\cdots ClF$ I	$NH_3\cdots HCl$ II	$NH_3\cdots ClF$ II	$NH_3\cdots HF$ II
$H_2O\cdots HF$	$H_2O\cdots PH_3$	$H_2O\cdots H_2S$	$H_2O\cdots HCl$	$HCl\cdots H_2S$
$HCONH_2\cdots H_2O$	$(CH_3)_2CO\cdots H_2O$	$H_3SiOSiH_3\cdots H_2O$	$PH_2CHCH_2\cdots H_2O$	$CH_3SH_2\cdots H_2O$
$H_2O\cdots OH^-$	$HCOO^-\cdots H_2O$	$H_2O\cdots Cl^-$	$H_3CNH_3^+\cdots H_2O$	$(CH_3)_2COH^+\cdots H_2O$
$(HSO_4^-\cdots NH_4^+)\cdots H_2O$	$HSO_4^-\cdots(NH_4^+\cdots H_2O)$			

[a]In some cases where two configurations of a dimer were tested, the two configurations are labeled as I and II. The detailed description is in ref 39.

where B labels the basis sets, M labels the molecules in the database, and G labels the geometries for each molecule in the database.

The MM charges used are Hartree−Fock/CHELPG[54] monomer charges obtained with the monomer basis set being used for SAPT calculations.

The SAPT calculations are carried out using SAPT2008.[28,55] All of the MM calculations are carried out in a modified TINKER program, called TINKERPLUS, which includes the evaluations of electrostatics between screened charges.[56]

During stage 1 of the parameter optimization, the ζ parameters of C, N, O, F, Si, P, S, and Cl are optimized to minimize the MUE shown in the above equation. We found that when both of the two screened charges are second row or third row elements, there is sometimes a numerical instability (due to round-off error in dealing with large numbers) of the screened electrostatics if the difference of their ζ values is not exactly zero but smaller than or equal to 0.02 a_0^{-1}. Therefore, in our optimization, we avoid such situations by optimizing the ζ value on a grid with a spacing of 0.03 a_0^{-1}. Therefore, because we avoided having small differences in ζ values, this will not be a problem in applications. Our generally preferred procedure is to include all of the considered data in a large training set (to obtain the most robust possible parametrization) rather than dividing the data into training data and test data; however, in the present case we used a slightly more complex scheme involving two stages with most of the data in stage 1 but additional data in stage 2.

In the second stage of parametrization, we added 10 water dimer geometries to the test set to increase our diversity for this important interaction (these dimer data are the same as used in ref 39), and we impose the constraint that the screened electrostatics should usually bring the results closer to the SAPT electrostatics but should not significantly overshoot the difference between the full electrostatics and the point charge approximation to the electrostatics. Therefore, we adjusted O, F, and Cl parameters for 10 H_2O dimers, for HF dimers with three geometries and two basis sets, and for HCl dimers with three geometries and two basis sets, respectively. For each of these three elements, we increased the ζ value for that element (with ζ values for other elements fixed at their stage-1 values) until only one of the electrostatic interaction energies in 10 H_2O dimers and no electrostatic interaction energies in HF and HCl dimers overshoot their SAPT goals by more than 10%. This stage leads to increases in the ζ values for oxygen, fluorine, and chlorine.

Table 2 shows the optimized parameters used for the screened charge model in MM calculations; this parameter set is labeled MMODS. Note that elements not in Table 2 are not studied in this paper. Extending the parametrization to a larger portion of the periodic table would be an interesting project for future work.

Table 2. ζ Values (in a_0^{-1})

atom	H	C	N	O	F
MMODS	1.32	1.45	1.15	1.36	1.33
atom	Si	P	S	Cl	
MMODS	1.21	0.91	1.27	1.21	

5. RESULTS AND DISCUSSIONS

The overall performance of the screened charge model with MMODS parameters is shown in Table 3. In the Supporting Information, we also compare several different sets of screening parameters we have used before.[39,41] To make a more interesting comparison, we also include the QM/MM results with the point charge model and the screened charge model with ODS parameters from ref 39 in Table 3. The screened charge model outperforms the point charge model not only in the QM/MM calculations, as shown before (as well as in Table 3), but also in MM calculations, as shown in Table 3. The mean unsigned errors over the 192 values in the test suite are 9.2 kcal/mol for unscreened MM (point charge model) and 3.3 kcal/mol for screened MM calculations with MMODS parameters. The improvement of screened MM over conventional MM is a factor of 2.8, which is quite significant.

It is surprising that the screened MM performs as well as screened QM/MM calculations because in the QM/MM calculations one of the monomers is always described by the accurate charge density while in the MM calculations the charge distributions of both of the monomers are approximated. A related finding is that the point charge model gives similar MUEs for the QM/MM (8.3 kcal/mol) and the MM calculations (9.2 kcal/mol). This indicates that when calculating the electrostatic interactions between two monomers, using the accurate charge density of only one monomer, as is done in unscreened QM/MM calculations, is not adequate to quantitatively model the charge penetration effect. This is consistent with the finding that the screened MM model presented here can actually be more accurate than the conventional (point-charge) QM/MM method. We recommend using the previous parameters (ODS) for interaction of an MM system with a polarizable QM system, and we recommend using the present parameters (MMODS) for MM−MM interactions.

To better understand the trends of screening effects for various geometries and basis sets, Table 4 shows the example of the mixed dimer of formamide and water. We can see that using the screened charge model, the performance is better for all three structures and both basis sets.

When studying the optimized ζ parameters from Table 2, we found that the ζ values for the third-row elements (those with highest principal quantum number $n = 3$) are smaller than those for

dx.doi.org/10.1021/ct5005142 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 3. MSE and MUE of Electrostatic Energies (kcal/mol) Using the QM/MM and MM Methods with the Point Charge Model and Screened Charge Models (Exact Values are SAPT Electrostatic Energies)[a]

		st/aT		st/dT		com/aT		com/dT		ext/aT		ext/dT		all	
method	charge model	MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE
QM/MM[b]	point charge	6.7	6.7	6.5	6.5	15.4	15.4	15.2	15.2	3.1	3.1	2.9	2.9	8.3	8.3
QM/MM[b]	ODS	1.6	2.0	1.2	2.1	3.7	5.3	2.4	5.1	1.0	1.1	1.0	1.1	1.8	2.8
MM	point charge	7.7	7.7	7.4	7.4	16.5	16.5	16.3	16.3	3.8	3.8	3.7	3.7	9.2	9.2
MM	MMODS	2.6	2.7	2.3	2.4	6.2	6.4	5.6	5.8	1.4	1.5	1.3	1.4	3.2	3.3

[a]The test suite includes 37 dimers. Abbreviations: st, standard; com, compressed; ext, extended; aT, aug-cc-pVTZ; dT, def2-TZVP. [b]QM/MM results were recalculated from ref 39 for 37 dimers.

Table 4. Electrostatic Energies (kcal/mol) of $HCONH_2\cdots H_2O$ Dimer in QM/MM and MM Calculations Compared with SAPT Results and MUE (kcal/mol) of MM and QM/MM Calculations over Three Geometries and Two Basis Sets[a]

		standard/aT	standard/dT	compressed/aT	compressed/dT	extended/aT	extended/dT	MUE
SAPT[b]		−11.3	−11.6	−21.9	−22.3	−6.6	−6.9	
QM/MM[b]	point charge	−7.4/−7.3[c]	−8.0/−7.5	−10.5/−11.3	−11.4/−11.5	−5.1/−5.0	−5.5/−5.1	5.4
QM/MM[b]	ODS	−10.5/−9.8	−10.8/−10.0	−22.0/−18.6	−22.5/−18.8	−5.9/−5.7	−6.2/−5.8	1.3
MM	point charge	−6.5	−7.0	−9.9	−10.6	−4.4	−4.8	6.2
MM	MMODS	−9.7	−10.0	−18.7	−19.1	−5.5	−5.8	2.0

[a]Abbreviations: aT, aug-cc-pVTZ; dT, def2-TZVP. [b]SAPT and QM/MM results are from ref 39. [c]x/y denotes that the electrostatic energy is x when $HCONH_2$ is the QM region and is y when H_2O is the QM region.

the second-row elements (those with highest principal quantum number $n = 2$). This is consistent with the well-known feature that the electron distributions of third-row elements are more diffuse.

We should keep in mind that if we optimized ζ values for different types of functionalities of the same element, we would get different values, and so the optimized values presented here are compromises. For example, in the H_2S dimer, the optimized ζ parameter for S would be ∼1.1, while the optimized parameter for S in $(CH_3)_2SO$ dimer would be ∼1.3. Although our goal in the present paper was to present a parameter set in which the parameters depend only on atomic numbers and not on atom types, the results for sulfur indicate that using different ζ values for atoms in different bonding environments can further improve the performance of the screened charge model.

Table 5 shows the results for 10 H_2O dimers (differing in their geometries). The largest error of these 10 water dimers is

Table 5. Electrostatic Interactions (kcal/mol) between H_2O Dimers with 10 Different Geometries Using SAPT Calculations and MM Calculations[a]

index of dimer	SAPT	PC	MMODS
1	−8.2	−5.0	−6.9
2	−7.0	−4.6	−6.3
3	−6.9	−4.8	−6.5
4	−6.8	−4.4	−6.3
5	−6.0	−4.0	−6.0
6	−5.7	−4.2	−6.2
7	−4.9	−3.4	−4.5
8	−1.5	−0.9	−1.1
9	−5.0	−3.9	−4.9
10	−2.9	−3.1	−3.6
MUE		1.7	0.5

[a]The MUE is calculated by comparison with the SAPT results. The geometries are from ref 57.

for the most stable water dimer for both the point charge model and the screened charge model. With the point-charge model, the mean unsigned error in the electrostatics is 1.7 kcal/mol, and with the stage-1 parameters, the screened electrostatics reduce this error to 1.1 kcal/mol. The stage-2 parameters further reduce this error to 0.5 kcal/mol at the price of raising the error in the stage-1 test data. Fitting parameters is always a matter of judgment. The parametrization scheme and parameters that we settled on are the ones that we recommend as most robust for general application.

We also want to mention that the only case in our test suite that screened MM overshoots the SAPT electrostatic energy by more than 10% is $NH_3\cdots ClF$ II with the extended geometry. $NH_3\cdots ClF$ II has the geometry of optimized $NH_3\cdots ClF$, with Cl and F switching their positions (this is one of the three standard geometries mentioned above that are not equilibrium geometries). The results are shown in Table 6. The point charge model predicts the wrong sign of the electrostatics, while our screened charge model corrects that by including the charge penetration effects. Therefore, the screened charge model works quite well for this difficult case.

We emphasize that in the present study, we have not included any multipole expansion terms in the evaluation of Coulomb interactions in the screened charge model, which has been done in work by others.[43,44] This keeps our model as simple as possible. Including multipole terms may further improve the performance of the charge model, but at the cost of complexity, as already discussed. The present model would be very easy to incorporate into any molecular mechanics program; it just requires replacing $1/R$ by a different function of R. It is true that the computational cost in the screened charge model increases compared to the point charge model because there are more exponentials and polynomials to evaluate. However, the charge penetration effect is short or medium ranged, and in a well optimized program one would just include the charge penetration effect for the nearby atom pairs. In a preliminary test, we have found that the screened charges can be replaced by point charges for atom pairs that are separated by several angstroms, and one can use a cutoff distance for screening. If one employs such a cutoff, the total computational cost of the screened charge model is very similar to that of the conventional point charge model.

dx.doi.org/10.1021/ct5005142 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 6. Electrostatic Interactions (kcal/mol) between NH_3···ClF II Using SAPT Calculations and MM Calculations[a]

		standard/aT	standard/dT	compressed/aT	compressed/dT	extended/aT	extended/dT
SAPT		−6.0	−5.7	−24.2	−23.8	−0.9	−0.6
MM	point charge	2.6	3.0	3.9	4.4	1.9	2.1
MM	MMODS	−6.5	−6.2	−19.2	−18.7	−1.5	−1.2

[a]Abbreviations: aT, aug-cc-pVTZ; dT, def2-TZVP.

In the current calculations, all of the SAPT calculations and charge analysis are at the Hartree−Fock level. In particular, the electrostatic interaction is calculated at the HF level, which does not include electron correlation. The partial atomic charges used for MM electrostatic energy calculations were also derived at the HF level, and so this is a consistent test of screening. If both the partial atomic charges and the electrostatic components of the interaction energies were computed at the same correlated level, we would expect to obtain very similar screening parameters and conclusions as in the present work where they are both calculated at the Hartree−Fock level.

There are two ways to use the improved electrostatics presented in this paper. In the long term, one could envision using improved electrostatics in force fields in which one attempts to make all terms as physically accurate as possible for individual components of the interaction energies. In the shorter term, one recognizes that molecular mechanics force fields will still contain empirical parameters such as van der Waals parameters that in part make up for deficiencies in electrostatics, polarization, charge transfer, damping of dispersion-like correlation terms, and missing higher-order correlation effects. However, the electrostatic terms are large and sometimes dominant, and we believe that making these terms more physical will contribute to making it easier to obtain transferable parametrization of the rest of the force field.

6. CONCLUSIONS

In the present study, we formulated the screened charge model for electrostatic interactions for use in molecular mechanics. The charge penetration effect is captured by the use of a smeared charge distribution. We applied the new charge model to a large test suite and compared the electrostatics to that from SAPT calculations. With conservative screening parameters, the performance is better than that for the point charge scheme by a factor of 2.8—in particular, the mean unsigned error in the electrostatic interactions drops from 9.2 to 3.3 kcal/mol. The screened MM model even yields more accurate electrostatics than QM/MM calculations with point charges. Improving the description of the electrostatics can certainly lead to more physical force fields, and that should eventually lead to more accurate force fields.

■ ASSOCIATED CONTENT

Supporting Information

Performance of screened charge models with various sets of screening parameters. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: truhlar@umn.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Kaining Duanmu, Laura Fernandez, and Pragya Verma for helpful discussions. This work was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences under award DE-FG02-12ER16362.

■ REFERENCES

(1) Hagler, A. T.; Lifson, S.; Dauber, P. *J. Am. Chem. Soc.* **1979**, *101*, 5122−5130.
(2) Jorgensen, W. L.; Swenson, C. J. *J. Am. Chem. Soc.* **1985**, *107*, 569−578.
(3) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024−10035.
(4) Allinger, N. L.; Zhou, X.; Bergsma, J. *J. Mol. Struct.: THEOCHEM* **1994**, *312*, 69−83.
(5) Hagler, A. T.; Ewig, C. S. *Comput. Phys. Commun.* **1994**, *84*, 131−155.
(6) 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.
(7) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490−519.
(8) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(9) 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.; 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.
(10) Chen, B.; Siepmann, J. I. *J. Phys. Chem. B* **1999**, *103*, 5370−5379.
(11) Ponder, J. W.; Case, D. A. *Adv. Protein Chem.* **2003**, *66*, 27−85.
(12) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(13) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(14) MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., III. *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(15) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656−1676.
(16) Torheyden, M.; Jansen, G. *Mol. Phys.* **2006**, *104*, 2101−2138.
(17) Tafipolsky, M.; Engels, B. *J. Chem. Theory Comput.* **2011**, *7*, 1791−1803.
(18) McDaniel, J. G.; Schmidt, J. R. *J. Phys. Chem. A* **2013**, *117*, 2053−2066.
(19) Ansorg, K.; Tafipolsky, M.; Engels, B. *J. Phys. Chem. B* **2013**, *117*, 10093−10102.
(20) McDaniel, J. G.; Schmidt, J. R. *J. Phys. Chem. B* **2014**, *118*, 8042−8053.
(21) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325−340.
(22) Morokuma, K.; Kitaura, K. In *Chemical Applications of Atomic and Molecular Electrostatic Potentials*; Politzer, P., Truhlar, D. G., Eds.; Plenum Press: New York, 1981; pp 215−242.
(23) Bagus, P. S.; Hermann, K.; Bauschlicher, C. W., Jr. *J. Chem. Phys.* **1984**, *80*, 4378−4386.

(24) Bagus, P. S.; Hermann, K.; Bauschlicher, C. W., Jr. *J. Chem. Phys.* **1984**, *81*, 1966−1974.
(25) Stevens, W. J.; Fink, W. H. *Chem. Phys. Lett.* **1987**, *139*, 15−22.
(26) Glendening, E. D.; Streitwieser, A. *J. Chem. Phys.* **1994**, *100*, 2900−2909.
(27) Mitoraj, M. P.; Michalak, A.; Ziegler, T. A. *J. Chem. Theory Comput.* **2009**, *5*, 962−975.
(28) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887−1930.
(29) Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. *J. Chem. Phys.* **2005**, *123*, 214103.
(30) (a) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233−239. (b) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341−350. (c) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933−5947.
(31) Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. *J. Chem. Phys.* **1996**, *105*, 1968−1986.
(32) Freitag, M. A.; Gordon, M. S.; Jensen, J. H.; Stevens, W. J. *J. Chem. Phys.* **2000**, *112*, 7300−7306.
(33) Piquemal, J. P.; Gresh, N.; Giessner-Prettre, C. *J. Phys. Chem. A* **2003**, *107*, 10353−10359.
(34) Cisneros, G. A.; Piquemal, J. P.; Darden, T. A. *J. Phys. Chem. B* **2006**, *110*, 13682−13684.
(35) Piquemal, J. P.; Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darden, T. A. *J. Chem. Phys.* **2006**, *124*, 104101.
(36) Werneck, A. S.; Filho, T. M. R.; Dardenne, L. E. *J. Phys. Chem. A* **2007**, *112*, 268−280.
(37) Cisneros, G. A.; Tholander, S. N.-I.; Parisel, O.; Darden, T. A.; Elking, D.; Perera, L.; Piquemal, J. P. *Int. J. Quantum Chem.* **2008**, *108*, 1905−1912.
(38) Elking, D. M.; Cisneros, G. A.; Piquemal, J. P.; Darden, T. A.; Pedersen, L. G. *J. Chem. Theory Comput.* **2010**, *6*, 190−202.
(39) Wang, B.; Truhlar, D. G. *J. Chem. Theory Comput.* **2010**, *6*, 3330−3342.
(40) (a) Stone, A. J. *J. Phys. Chem. A* **2011**, *115*, 7017−7027. (b) Lu, Z.; Zhou, N.; Wu, Q.; Zhang, Y. *J. Chem. Theory Comput.* **2011**, *7*, 4038−4049.
(41) Wang, B.; Truhlar, D. G. *J. Chem. Theory Comput.* **2012**, *8*, 1989−1998.
(42) Cisneros, G. A.; Karttunen, M.; Ren, P.; Sagui, C. *Chem. Rev.* **2014**, *114*, 779−814.
(43) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J. P. *J. Chem. Theory Comput.* **2007**, *3*, 1960−1986.
(44) Gordon, M. S.; Smith, Q. A.; Xu, P.; Slipchenko, L. V. *Annu. Rev. Phys. Chem.* **2014**, *64*, 553−578.
(45) Strand, T. G.; Bonham, R. A. *J. Chem. Phys.* **1964**, *40*, 1686−1691.
(46) Roothaan, C. C. J. *J. Chem. Phys.* **1951**, *19*, 1445−1458.
(47) Silver, D. M.; Ruedenberg, K. *J. Chem. Phys.* **1968**, *49*, 4301−4305.
(48) Silver, D. M.; Ruedenberg, K. *J. Chem. Phys.* **1968**, *49*, 4306−4311.
(49) Kotani, M.; Amemiya, A. In *Table of Molecular Integrals*; Maruzen Company: Tokyo, 1955.
(50) *Mathematica*, version 9.0; Wolfram Research, Inc.: Champaign, IL, 2012.
(51) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(52) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796−6806.
(53) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297−3305.
(54) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361−373.
(55) Bukowski, R.; Cencek, W.; Jankowski, P.; Jeziorski, B.; Jeziorska, M.; Kucharski, S. A.; Lotrich, V. F.; Misquitta, A. J.; Moszynski, R.; Patkowski, K.; Podeszwa, R.; Rybak, S.; Szalewicz, K.; Williams, H. L.; Wheatley, R. J.; Wormer, P. E. S.; Żuchowski, P. S. SAPT2008: "An Ab Initio Program for Many-Body Symmetry-Adapted Perturbation Theory Calculations of Intermolecular Interaction Energies." http://www.physics.udel.edu/~szalewic/SAPT/index.html.
(56) Wang, B.; Truhlar, D. G.; *TINKERPLUS*, version 2014; University of Minnesota: Minneapolis, MN, 2014, based on TINKER 6.3: Ponder, J. W. *TINKER*, version 6.3; Washington University: St. Louis, MO, 2014. https://t1.chem.umn.edu/truhlar/index.htm#software.
(57) Reinhardt, P.; Piquemal, J. P. *Int. J. Quantum Chem.* **2009**, *109*, 3259−3267.