JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Multipole-Based Force Fields from ab Initio Interaction Energies and the Need for Jointly Refitting All Intermolecular Parameters

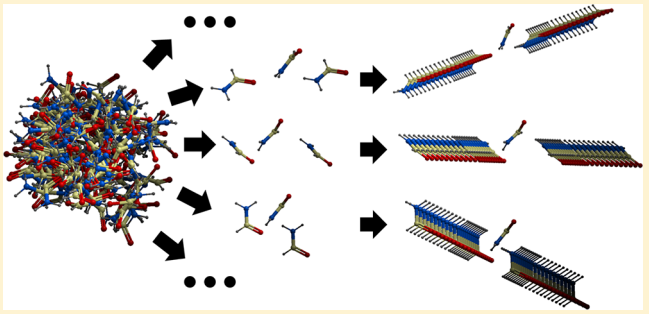
Christian Kramer,*[,†,‡] Peter Gedeck,*[,§] and Markus Meuwly*[,‡]

[†]Novartis Institutes for BioMedical Research, Basel, Switzerland

[‡]Department of Chemistry, University of Basel, Klingelbergstrasse 80, 4056 Basel, Switzerland

[§]Novartis Institutes for Tropical Diseases, Singapore

ABSTRACT: Distributed atomic multipole (MTP) moments promise significant improvements over point charges (PCs) in molecular force fields, as they (a) more realistically reproduce the *ab initio* electrostatic potential (ESP) and (b) allow to capture anisotropic atomic properties such as lone pairs, conjugated systems, and σ holes. The present work focuses on the question of whether multipolar electrostatics instead of PCs in standard force fields leads to quantitative improvements over point charges in reproducing intermolecular interactions. To this end, the interaction energies of two model systems, benzonitrile (BZN) and formamide (FAM) homodimers, are characterized over a wide range of dimer conformations. It is found that although with MTPs the monomer *ab initio* ESP can be captured better by about an order of magnitude compared to point charges (PCs), this does not directly translate into better describing *ab initio* interaction energies compared to PCs. Neither ESP-fitted MTPs nor refitted Lennard-Jones (LJ) parameters alone demonstrate a clear superiority of atomic MTPs. We show that only if both electrostatic and LJ parameters are jointly optimized in standard, nonpolarizable force fields, atomic are MTPs clearly beneficial for reproducing *ab initio* dimerization energies. After an exhaustive exponent scan, we find that for both BZN and FAM, atomic MTPs and a 9−6 LJ potential can reproduce *ab initio* interaction energies with ∼30% (RMSD 0.13 vs 0.18 kcal/mol) less error than point charges (PCs) and a 12−6 LJ potential. We also find that the improvement due to using MTPs with a 9−6 LJ potential is considerably more pronounced than with a 12−6 LJ potential (≈ 10%; RMSD 0.19 versus 0.21 kcal/mol).



1. INTRODUCTION

Force field simulations have become one of the standard and widely used tools in molecular sciences, providing insight and allowing interpretations of phenomena at the molecular level.[1−3] Contrary to a quantum mechanical approach, the energy in empirical force fields is calculated as a sum over phenomenological terms. In all major force fields for biochemical simulations, including AMBER,[1] CHARMM,[3] GROMOS,[4] and OPLS,[5] intermolecular forces are approximated by point charges (PCs) and a Lennard-Jones[6] (LJ) term. PCs describe the electrostatic interactions, whereas the LJ term accounts for all remaining intermolecular attractive and repulsive interactions.

Force-field parameters can be derived by different strategies: Historically, most force-field parameters were obtained from fitting the results of simulations to spectroscopic (AMBER, CHARMM) or thermodynamic data (OPLS, GROMOS). This is consistent with the notion that such data are also used to subsequently validate force fields, and the expectation is that correct macroscopic and spectroscopic observables (e.g., solvation free energies or NMR scalar coupling constants) can only be calculated if the microscopic behavior of the simulations—the movements of the atoms—is correct.

Unfortunately, this approach has a number of significant shortcomings: First, there is not always a sufficient number of consistent, high-quality spectroscopic and thermodynamic data available to develop parameters for a large collection of molecules, especially molecules of pharmaceutical interest. Second, even if "only" normal proteins should be simulated, the parameter space for force fields that use nonstandard functional forms like MTPs or polarizability grows rapidly, and fitting those parameters requires a large amount of experimental data. Third, fitting force fields to experimental data is very time-consuming, since the iteration cycles for the evaluation of a new parameter guess are very long if each guess is validated by a set of molecular mechanics simulations.

Therefore, methods for improving force fields by fitting to the results of high-level *ab initio* calculations are being developed by others and us.[7−15] Compared to most experiments, *ab initio* calculations are relatively rapid and cheap and yield access to an endless number of properties which are not always experimentally accessible (such as interaction energies at clearly defined geometries). They can be used to calculate

Received: October 12, 2012
Published: February 15, 2013

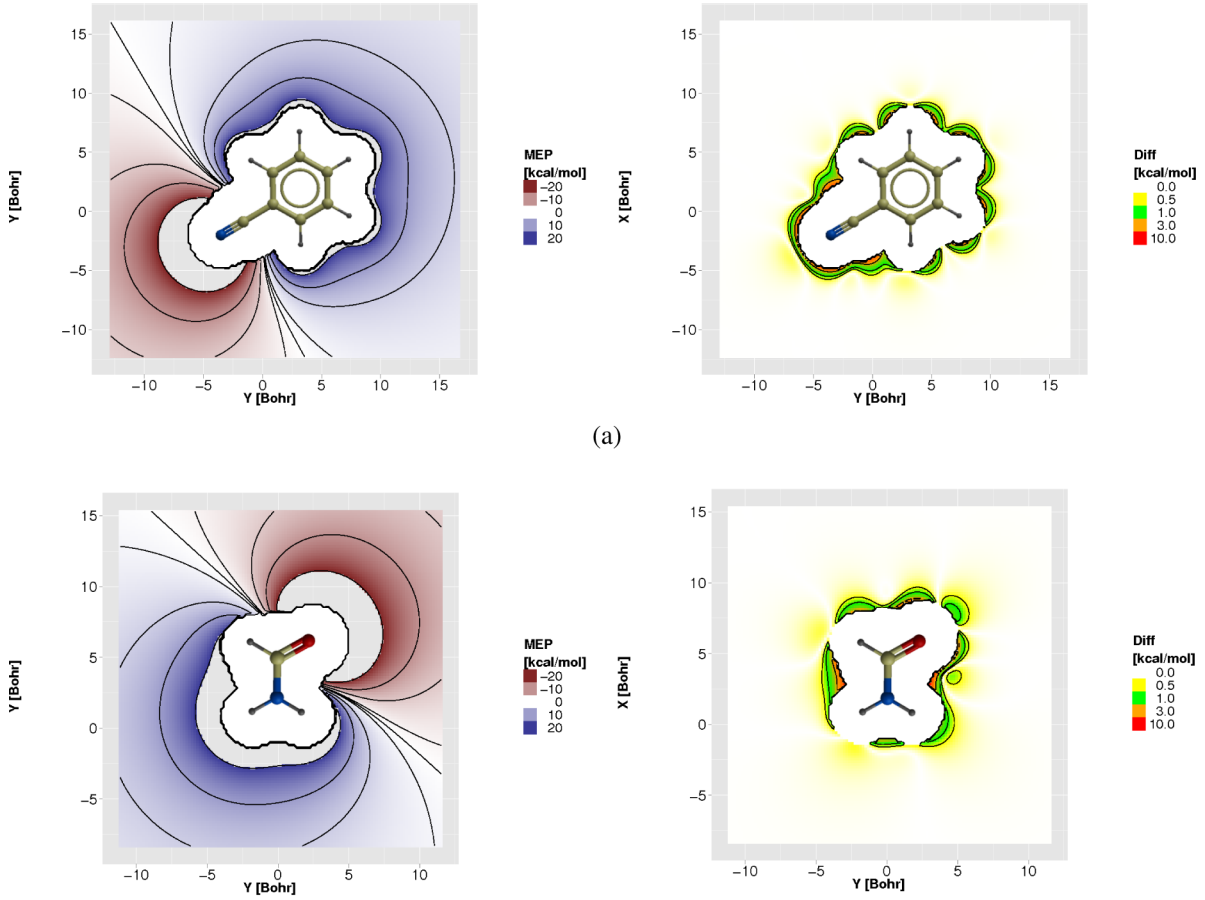


Figure 1. ESP from fitted MTPs (isocontours at −20, −10, −5, −1, 0, 1, 5, 10, and 20 kcal/mol) and difference to the *ab initio* ESP (isocontours at 0.5, 1, 3, and 10 kcal/mol) for (a) BZN and (b) FAM.

complete potential energy surfaces (PESs), which in turn are target functions for fitting force field (FF) parameters much more efficiently than to thermodynamic data. Modern high-level *ab initio* methods can provide results with errors within experimental uncertainty. Therefore, the *ab initio* PES is the relevant quantity to fit to, and closely reproducing it should yield meaningful macroscopic observables.

Fitting to *ab initio* interaction energies allows using energy decomposition techniques and fitting individual energy terms to electrostatics, dispersion, repulsion, and polarization. However, most of the current standard force field parametrizations—which are hard-coded in the more widely used molecular mechanics programs—do not explicitly include polarizability terms, and the LJ expression for the repulsion ($\propto r^{-12}$) is known to be suboptimal. Therefore, the physical energy terms (see above) cannot be easily described individually, and certain effects are captured implicitly. For example, ESP charges are usually scaled to account for polarization effects in condensed-phase simulations. The importance of mutual polarization has recently been demonstrated,[16] and proper inclusion of polarizability into force fields is an exciting avenue in modern force field development. One should also remember that parametrizing such terms in a consistent manner is a formidable endeavor, requires profound changes in the computer codes, and will usually be computationally much more demanding than the standard expressions for nonbonded interactions.

While force fields based on energy decomposition techniques are still under development and a number of critical questions remain to be fully addressed (e.g., transferability of gas-phase-derived parameters to be used in condensed-phase simulations), there is a rapidly growing body of successful studies where purely *ab initio* derived parameters have been used to compute macroscopic properties.[17−20] Nevertheless, it needs to be emphasized that the ultimate test of a parametrized force field is still comparison with experimental data, which may even require conformational sampling, depending on the observables of interest.[21]

One of the essential ingredients in a force field derived from *ab initio* calculations are PCs, although it was demonstrated that it may also be possible to derive suitable PCs from very high resolution X-ray structures (0.5 Å and better).[22] Mulliken[23] and Coulson[24] charges are calculated directly from the wave functions, whereas RESP[25] and CHELPG[26] charges are fitted to reproduce the *ab initio* electrostatic potential (ESP) around the molecule. The ESP is a particularly interesting quantity derived from the solution of the electronic Schrödinger equation because it is an easy-to-calculate measure for the electrostatic contribution to interaction energies. While PCs have been used for a long time, it has always been clear that they are not able to exactly reproduce the ESP outside of a molecule.[27] Since PCs are isotropic by definition, they lack spatial resolution and cannot realistically reproduce the ESP around conjugated systems, aromatic systems, or lone pairs.[28]

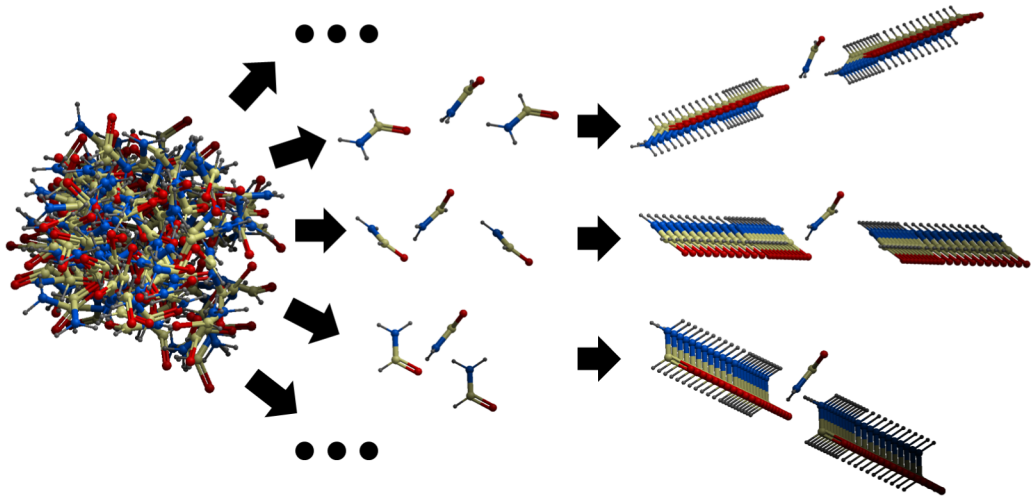


Figure 2. Sampling scheme used to generate the ensemble of homodimers. First, 100 relative conformations are randomly generated while keeping one monomer fixed. Then, for each relative orientation, 22 snapshots are created along the vector connecting the center of mass of both monomers in both directions.

Atomic multipoles (MTPs) offer an attractive solution to this problem since they are able to reduce the error in reproducing the *ab initio* ESP by up to 90%.[15] MTP electrostatics for flexible molecules have only recently been implemented,[8] but there are still a number of questions concerning the propagation of torques in MD simulations.

One important, yet unanswered question for practical applications of MTPs is whether—and if so, by how much—they improve the accuracy of intermolecular interaction energies, i.e., whether a more accurate representation of the ESP (as provided by MTPs) directly translates into quantitatively better interaction energies compared to *ab initio* calculations. We address this question by considering interaction energies between two homodimer model systems (benzonitrile (BZN) and formamide (FAM)) where the ESP outside the molecule is better reproduced with MTPs than with PCs. To this end, we use a standard force field expression with a LJ term for the dispersion/repulsion energy and omit polarizability terms in order to stay as close as possible to the current standard. Using this approach, we implicitly assume that the error introduced from omitting polarizability is similar for both PCs and MTPs.

For the monomers, the ESP outside the molecule can be modeled very accurately by atomic MTPs up to quadrupoles, whereas optimized atomic monopoles give a roughly 10-fold less accurate ESP. Several thousand different relative orientations of the homodimers are generated to sample the PES, and interaction energies from *ab initio* calculations are compared to those from force fields with (a) PCs and (b) MTPs up to quadrupoles. We show that only going from PCs to MTPs does not improve the reproduction of the *ab initio* energies, and both charge schemes give unsatisfactory overall interaction energies if standard LJ parameters are used. Also, refitting LJ parameters to the difference between *ab initio* energies and electrostatic energies does not clearly show the superiority of MTPs. Only if *both* electrostatic and LJ parameters are fitted at the same time, MTPs are clearly superior. The overall errors decrease and the agreement further improves when the 12−6 LJ term is replaced by a 9−6 LJ term. This leads us to the conclusion that if MTPs should be introduced into force field simulations to replace standard PCs, all intermolecular parameters have to be adjusted simultaneously.

2. METHODS

2.1. Multipole Parameters. BZN and FAM were selected as model systems because they are small and rigid and they contain various strongly anisotropic charge distribution features, e.g, an unequal distribution of electrons around the individual atoms. These features are the aromatic π system and the cyano group in BZN and the carbonyl and the highly polar H-bond donating and accepting features in FAM, rendering all atoms ideal candidates for MTPs.

Both the BZN and the FAM monomers have been geometry optimized at the MP2/6-311+G* level using Gaussian 03.[29] Atomic MTP moments have been calculated from the *ab initio* electron density using the GDMA2 program[30] (henceforth called GDMA). Starting from GDMA MTP moments as an initial guess, atomic MTPs up to quadrupoles have been fit to the *ab initio* electron density in the first interaction belt (the area between 1.66 and 2.2 times outside the van der Waals radii) using a simplex optimization as previously described[15] and yields an error of 0.033 and 0.036 kcal/mol on average for BZN and FAM, respectively. PCs have been fit to the individual atoms in the same way as the MTPs. With PCs, the ESP in the first interaction belt is reproduced with an average error of 0.18 kcal/mol (BZN) and 0.77 kcal/mol (FAM). For details of the fitting procedure, see Kramer et al.[15] The fit is constrained to yield the same charges for chemically equivalent atoms. The ESP resulting from the MTPs and the difference to the ESP arising from the *ab initio* calculations are shown in Figure 1.

2.2. Sampling. For the BZN homodimer, 100 relative conformations have been generated by randomly rotating one monomer. For each of the orientations, a series of conformations along a random vector was generated by sliding one monomer along that vector in both directions. Snapshots were taken at the minimum distance between any two atoms of the two monomers of 1.5 to 4.0 Å in steps of 0.25 Å and of 4.0 to 10.0 Å in steps of 0.5 Å. For the FAM dimer, the same scheme has been used to generate the ensemble, except that the upper threshold for the minimum distance between any two atoms of the monomers is 8.0 Å. At this distance, the

interaction energies are almost zero. The procedure is schematically shown in Figure 2.

2.3. Energy Calculations. Standard force field parameters for both monomers have been obtained from SwissPARAM.[31] For internal degrees of freedom SwissPARAM uses MMFF parameters,[32] whereas the electrostatics and LJ parameters are obtained from the closest atom types in CHARMM22.[33] The LJ parameters and the PCs obtained from SwissPARAM are listed in Table 1, and the atom type assignment is shown in Figure 3.

Table 1. LJ $r_{min,i}$ [Å] and ε_i [kcal/mol] Parameters Obtained from SwissPARAM[31]

monomer	atom type	ε_i	$2^*r_{min,i}$
BZN	CB	−0.070	3.98
	CBB	−0.070	3.98
	CSP	−0.068	4.16
	HCMM	−0.022	2.64
	NSP	−0.20	3.70
FAM	NC═O	−0.20	3.70
	C═O	−0.11	4.00
	O═C	−0.12	3.40
	HNCO	−0.046	0.44
	HC═O	−0.022	2.64

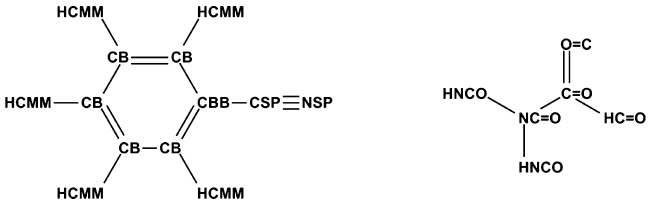


Figure 3. Atom type assignment for BZN (left) and FAM (right).

All force-field interaction energies have been calculated using our own Python implementation of MTP electrostatics based on atomic MTPs within their local reference axis system.[15] A number of MTP parameters, especially most of the Q_{xy}, Q_{xz}, and Q_{yz} parameters, can be set to zero due to symmetry, see Kramer et al.[15] Single point *ab initio* interaction energies for the dimers were calculated for each snapshot in Gaussian 03[29] at the MP2/6-311+G* level and corrected for basis set superposition error (BSSE) by using the counterpoise correction.[34] The conventional counterpoise method fails in some cases.[35] However, for a range of H-bonded homodimers, including the FAM dimer investigated here, it was found that including BSSE corrections is mandatory and the issue of using BSSE-corrected or uncorrected energies remains controversial.[36] Thus, for fitting a higher-accuracy force field it might be worth to consider alternative *ab initio* methods such as local MP2[37] and eventually fit to experimental data, including spectroscopic and thermodynamic observables. While the level of theory employed in the present study may not be the highest possible, it is computationally affordable and sufficient to generate realistic interaction energies.

The FAM dimer and multimer have been used as a model system for many different previous studies, including proton transfer,[41,42] the connection between low-mode vibrations and hydrogen bond strength,[43] and tautomerization.[44] Comparison with dispersion-corrected functionals (M06-2X[38] and wB97DX[39]) and a larger basis set (aug-cc-PVTZ) suggest that the *ab initio* level used here underestimates hydrogen bonding energies in FAM. The dissociation energy for the global minimum conformation of the FAM dimer (cyclic with two hydrogen bonds) is 12.20 kcal/mol at the MP2/6-311+G* level, while it is 15.16 kcal/mol on the M06-2X/aug-cc-PVTZ level and 15.39 kcal/mol on the wB97xD/aug-cc-PVTZ level (all optimizations done on the MP2/6-311+G* level). This compares with ∼15.8 kcal/mol from MP2 calculations with explicitly correlated wave functions.[40] However, the minimum energy conformation is not part of the sampled structures since we use a fixed monomer geometry and randomly generated relative orientations. For one of the minimum energy conformations generated in our sampling, the MP2/6-311+G* dissociation energy is 6.49 kcal/mol, whereas the higher level dissociation energies are 7.55 kcal/mol for wB97xD/aug-cc-PVTZ and 7.64 kcal/mol for M06-2X/aug-cc-PVTZ.

Since the purpose of the present work is not the development of a general-purpose force field but rather to provide insight into the performance of monomer, ESP-fitted PCs and MTPs to compute interaction energies, BSSE-corrected MP2/6-311+G* energies are deemed sufficient. We also note that recent CHARMM-parametrizations have been carried out with geometries optimized at the MP2/6-31+G(d) level and single point calculations at the MP2/cc-pVTZ level.[45]

2.4. LJ Parameter Optimization. The force field interaction energy for each snapshot is calculated according to

$$E_{inter,FF} = E_{Dimer,FF} - E_{Monomer1,FF} - E_{Monomer2,FF} = E_{Estat} + E_{LJ} \quad (1)$$

which is decomposed into the electrostatic energy terms E_{Estat} and the LJ energy terms E_{LJ}. The electrostatic term can be written as a multipole expansion of the charge distribution, leading to an ESP $\Phi(\vec{R})$:

$$\Phi(\vec{R}) = \frac{1}{4\pi\varepsilon_0}\int_V \frac{\rho(\vec{r})\,\mathrm{d}r}{|\vec{R}-\vec{r}|} = \frac{1}{4\pi\varepsilon_0}\left(\frac{q}{|\vec{R}|} + \mu_i\frac{R_i}{\vec{R}^3} + Q_{ij}\frac{3R_iR_j - \vec{R}^2\delta_{ij}}{2\vec{R}^5} + \ldots\right) \quad (2)$$

In conventional force fields, only the first term, i.e., the PCs, is used. Most major force fields, including CHARMM, use the 12−6 LJ potential to represent the nonelectrostatic intermolecular interaction terms.

$$E_{LJ,ij} = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right] \quad (3)$$

Here, i and j are atoms of monomers one and two, respectively. Since all internal degrees of freedom are kept rigid for both homodimers examined, intramolecular LJ terms can be ignored. The variables ε_{ij} and σ_{ij} can be interpreted as the depth of the potential well and the distance between the two atoms i and j where the interaction energy equals zero, respectively. CHARMM uses a slightly rewritten version of the LJ potential with $r_{min,ij}$, the distance where the LJ term has its minimum, i.e., where it is the most attractive. This is formulated as

$$E_{\mathrm{LJ},ij} = \varepsilon_{ij}\left[\left(\frac{r_{\min,ij}}{r}\right)^{12} - 2\left(\frac{r_{\min,ij}}{r}\right)^{6}\right] \quad (4)$$

It has been noted[46] that the r^{-12} term is too steep and yields too repulsive energies. Further, the −12 exponent is chosen for computational convenience and is not physically justified. An alternative to the 12−6 LJ potential is the 9−6 potential,[46] which can be written as

$$E_{\mathrm{LJ},ij} = \varepsilon_{ij}\left[2\left(\frac{r_{\min,ij}}{r}\right)^{9} - 3\left(\frac{r_{\min,ij}}{r}\right)^{6}\right] \quad (5)$$

The $r_{\min,ij}$ and ε_{ij} parameters are usually derived from individual parameters for atoms i and j that are combined to give the diatomic ε_{ij} and $r_{\min,ij}$ parameters. CHARMM, for example, uses the Lorenz−Berthelot mixing rules $\varepsilon_{ij} = (\varepsilon_i \cdot \varepsilon_j)^{1/2}$ and $r_{\min,ij} = 0.5(r_{\min,i} + r_{\min,j})$.[47] Baker et al. have recently investigated alternative mixing rules and found that there is scope for improvement.[48] The atomic parameters have been derived either from virial coefficients of the ideal gas equation or from crystal structures.[49]

The electrostatic and LJ parameters were optimized using a standard Nelder-Mead simplex implemented in SciPy.[50] We examined several other optimization algorithms that use numerical gradients, but it turned out that the standard simplex gave the best results. For each optimization, a maximum of 100 000 steps were allowed, whereas the optimization usually converged far before the maximum number of steps was reached. The convergence criterion was set to a 10^{-6} difference in the sum of the weighted squared difference between *ab initio* and force-field energy. Relative orientations with highly repulsive *ab initio* interaction energies (E_{abinitio} > 10 kcal/mol) were excluded from all further calculations, since they are less likely to be populated but dominate the fit if included. Boltzmann weights have been assigned to all repulsive conformations where the minimum of *ab initio* and force-field energy is larger than zero in order to give the highest weights to those conformations that are most likely sampled. It is necessary to use the minimum of both *ab initio* and force-field energy to ensure that false positive conformations do not occur. If only the *ab initio* energies were used for weighting, relative conformations that are repulsive according to *ab initio* energies but attractive according to force field energies might occur because they are not penalized. The minimization can then be written as

$$\text{minimize} \sum_{k}^{N_s} w_k (E_{\mathrm{abinitio},k} - E_{\mathrm{forcefield},k})^2 \quad (6)$$

where N_s is the number of snapshots and weights

$$w_k = \begin{cases} 1 & \text{if } \Delta E < 0 \\ \exp^{-\Delta E/RT} & \text{else} \end{cases} \quad (7)$$

where $\Delta E = \min(E_{\mathrm{abinitio},k}, E_{\mathrm{forcefield},k})$, RT = 0.59179 kcal/mol at T = 298 K, $E_{\mathrm{abinitio},k}$ is the *ab initio* interaction energy, and $E_{\mathrm{forcefield},k}$ is the force field energy calculated for snapshot k. A flowchart of the overall fit process is shown in Figure 4.

2.5. Full Exponent Search using Mie Potentials. In replacing the LJ potential by the more general Mie potential,[51] $E_{\mathrm{LJ},ij} = (A_{ij,12}/r_{ij}^{12}) - (B_{ij,6}/r_{ij}^{6})$ an atom-pairwise expression for the parameters $A_{ij,12}$ and $B_{ij,6}$ is obtained. This permits

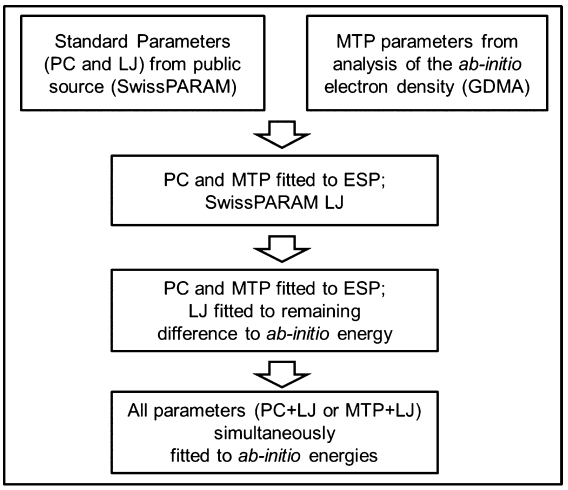


Figure 4. Flowchart to summarize the overall fit process.

determining $A_{ij,12}$ and $B_{ij,6}$ directly from a linear least-squares fit, which is computationally much more efficient and stable:

$$\begin{aligned} &\text{minimize} \sum_{k}^{N_s} (E_{\mathrm{abinitio},k} - E_{\mathrm{Estat},k} - \mathrm{E}_{\mathrm{LJ},k})^2 \\ &= \text{minimize} \sum_{k}^{N_s} \left((E_{\mathrm{abinitio},k} - E_{\mathrm{Estat},k}) \right. \\ &\quad \left. - \left(\sum_{i=1}^{\mu} \sum_{j=1}^{\nu} \frac{A_{ij,12}}{r_{ij,k}^{12}} - \frac{B_{ij,6}}{r_{ij,k}^{6}} \right) \right)^2 \end{aligned} \quad (8)$$

Here, μ and ν denote the numbers of atoms in both monomers which do not need to be identical. Since the exponents (m = 12, n = 6) are not necessarily the optimal coefficients in a least-squares sense (although we note that there is a theoretical justification for r^{-6}), we decided to treat them as adjustable parameters m and n. As a linear regression, the fit is very efficient and guaranteed to find the global minimum. Fitting a Mie potential avoids mixing rules, since only pairwise parameters are fitted. If both $A_{ij,m}$ and $B_{ij,n}$ are positive, the distance $r_{\min,ij}$ at which $V_{\mathrm{LJ},ij}(r_{ij})$ has its extremum can be obtained from $r_{\min,ij} = (mA_{ij,m}/nB_{ij,n})^{1/m-n}$ and ε_{ij} as the energy of the extremum of the potential follows from $\varepsilon_{ij} = A_{ij,m}(mA_{ij,m}/nB_{ij,n})^{m/n-m} - B_{ij,n}(mA_{ij,m}/nB_{ij,n})^{n/n-m}$ with $n < m$. Also, $A_{ij,m}$ and $B_{ij,n}$ are not restrained to be positive, so results might be obtained where no pair-specific ε_{ij} and $r_{\min,ij}$ can be calculated from the parameters found.

If Mie potentials should be used for fitting practical force fields, some restraints might have to be used because terms with similar exponents are highly correlated. For the work presented here, this does not play a role, since we are not interested in the actual Mie coefficients.

A deeper analysis of direct pairwise fits of the Mie potential is beyond the scope of this manuscript. For the remaining manuscript, we will use the Mie potential to compare various (n,m) exponent combinations. All fits were done in R[52] with the standard linear regression routine *lm*, setting the y axis intersection to zero. Weights were not used in order to keep the fit efficient, but all relative orientations with *ab initio* interaction energies >10 kcal/mol were excluded from the fits.

3. RESULTS

3.1. Comparison of Interaction Energies with Standard LJ Parameters. For both dimer ensembles (5400 conformations for BZN and 4400 for FAM), the *ab initio* energies with LJ parameters obtained from SwissParam and either SwissPARAM PCs, GDMA MTPs, or ESP-fitted PC and MTP parameters were calculated and compared with the MP2/6-311+G* energies ($E_{\mathrm{abinitio},k}$). The results are shown in Table 2. The absolute value of the root mean squared error (RMSE) depends on the number of relative orientations with a large distance between the monomers, because at long-range (large r) the interaction energy and the error are small. Therefore, the absolute values of the error can only be compared within the same ensemble and should not be compared between the two homodimer series. Parameters from SwissPARAM are not expected to give energies that compare well with *ab initio* energies because they were not developed to reproduce gas phase dimerization energies. They only serve as a largely arbitrary reference point from a typical force field parametrization for comparison. For BZN, the overall error with ESP-fitted PCs and SwissPARAM LJ parameters (0.67 kcal/mol) is slightly lower than with SwissPARAM PCs and SwissPARAM LJ parameters (0.69 kcal/mol) but equal to the error obtained with ESP-fitted MTPs and SwissPARAM LJ parameters (0.67 kcal/mol). The combination of GDMA MTPs and SwissPARAM LJ parameters yields an error of 0.65 kcal/mol, but all errors are similar. Refitting the LJ parameters for all charge models to the difference between *ab initio* and electrostatic energy reduces the error by more than 50% to 0.24−0.25 kcal/mol. Although they reproduce the electrostatic field much better, MTPs do not outperform PCs.

Table 2. Root Mean Squared Error [kcal/mol] for PC and MTP Electrostatics Representations with Standard 12-6 and Refitted 12-6 LJ Parameters[a]

monomer	charge model	charge parameters	SwissPARAM LJ	*ab initio* fit LJ
BZN	PC	SwissPARAM	0.69	−−[b]
		ESPfit	0.67	0.24
	MTP	GDMA	0.65	0.25
		ESPfit	0.67	0.25
FAM	PC	SwissPARAM	0.90	−−[b]
		ESPfit	1.14	0.56
	MTP	GDMA	1.13	0.55
		ESPfit	1.12	0.51

[a]Abbreviations. ESPfit: PCs and MTPs from fitting to the ESP. *Ab initio* fit: LJ parameters (ε and σ) from fitting to *ab initio* interaction energies. GDMA PCs have not been examined. SwissPARAM only gives PCs, so there is no entry for SwissPARAM MTPs. [b]Not examined.

For FAM, the error obtained with GDMA MTPs (1.13 kcal/mol) and ESP-fitted PCs (1.14 kcal/mol) and MTPs (1.12 kcal/mol) is significantly larger than the error obtained with SwissPARAM PCs (0.90 kcal/mol), if combined with SwissPARAM LJ parameters. Again, the error is reduced by more than 50%, if the LJ parameters are refitted to *ab initio* energies. For FAM, ESP-fitted MTPs are slightly superior to ESP-fitted PCs (0.51 vs 0.56 kcal/mol), if the LJ parameters are refitted to *ab initio* energies.

3.2. Refitting 12−6 LJ Parameters. A LJ ansatz attempts to capture several physically different phenomena: The attractive r^{-6} term is used to model the dispersion interactions between induced dipoles, whereas the r^{-12} term is repulsive and should describe all short-range repulsive interactions except electrostatic repulsion. If the electrostatic model is modified, the LJ model also has to be adjusted. Therefore, we fitted the 12−6 LJ parameters to the difference between the *ab initio* interaction energies and the electrostatic energies based on ESP-fitted PCs and MTPs, respectively. For BZN and PCs, the RMSE was 0.24 kcal/mol, and with MTPs it was 0.25 kcal/mol. This is a significant improvement compared to the result obtained with SwissPARAM LJ parameters (0.65 and 0.69 kcal/mol for PCs and MTPs, respectively), but the MTPs perform slightly worse than PCs. Also, ESPfit MTPs do not yield a lower error than unfitted GDMA MTPs. For FAM, the RMSE with fitted 12−6 LJ parameters is 0.56 kcal/mol for ESP-fitted PCs, 0.55 kcal/mol for GDMA MTPs, and 0.51 kcal/mol for ESP-fitted MTPs. This is also significantly better than the result obtained with standard LJ parameters (1.12−1.14 kcal/mol), but the ESP-fitted MTPs yield only slightly lower errors than ESP-fitted PCs. The results are summarized in Table 2.

3.3. Exponent Search. Using the pairwise decomposition method described in section 2.5, we scanned all combinations between 4 and 12 for the exponents m and n of the Mie potential (see Methods). The RMSE for all BZN dimer conformations with *ab initio* energies below 10 kcal/mol and for all combinations of m and n and both PCs and MTPs is summarized in Tables 3 and 4.

The exponent combinations with the lowest errors are 10−6, 9−6, 9−7, and 8−7. MTPs give slightly better results than PCs. Notice that the errors from the exponent search using Mie potentials (e.g ESPfit MTP + Mie(12,6): 0.37 kcal/mol, Table

Table 3. Root Mean Squared Error in kcal/mol after Fitting Mie Potentials Combined with MTP Electrostatics Representations for Different Exponents of the Mie Potential and the BZN Dimer[a]

		m							
		12	11	10	9	8	7	6	5
n	11	0.78							
	10	0.71	0.64						
	9	0.63	0.57	0.48					
	8	0.54	0.47	0.40	0.34				
	7	0.45	0.39	0.33	**0.29**	**0.27**			
	6	0.37	0.32	**0.29**	**0.28**	0.32	0.39		
	5	0.33	0.31	0.31	0.35	0.42	0.50	0.61	
	4	0.37	0.38	0.41	0.47	0.55	0.64	0.74	0.86

[a]Best four solutions highlighted in bold.

Table 4. Root Mean Squared Error in kcal/mol after Fitting Mie Potentials Combined with PC Electrostatics Representations for Different Exponents of the Mie Potential and the BZN Dimer[a]

		m							
		12	11	10	9	8	7	6	5
n	11	0.77							
	10	0.70	0.63						
	9	0.62	0.55	0.48					
	8	0.54	0.47	0.41	0.34				
	7	0.46	0.39	0.34	**0.30**	**0.29**			
	6	0.38	0.33	**0.30**	**0.30**	0.33	0.40		
	5	0.34	0.32	0.33	0.36	0.42	0.51	0.62	
	4	0.38	0.39	0.42	0.48	0.55	0.64	0.75	0.86

[a]Best four solutions highlighted in bold.

Table 5. Root Mean Squared Error in kcal/mol after Fitting Mie Potentials Combined with MTP Electrostatics Representations for Different Exponents of the Mie Potential and the FAM Dimer[a]

		m										
		12	11	10	9	8	7	6	5	4	3	2
n	11	0.60										
	10	0.59	0.58									
	9	0.58	0.56	0.54								
	8	0.56	0.54	0.52	0.50							
	7	0.55	0.53	0.50	0.48	0.46						
	6	0.53	0.50	0.48	0.46	0.44	0.43					
	5	0.50	0.48	0.45	0.43	0.42	0.43	0.45				
	4	0.48	0.45	0.43	**0.41**	0.42	0.44	0.49	0.56			
	3	0.45	0.43	**0.41**	**0.41**	0.43	0.48	0.55	0.63	0.73		
	2	0.44	**0.41**	**0.40**	0.42	0.46	0.53	0.63	0.73	0.84	0.96	
	1	0.45	0.43	0.42	0.45	0.51	0.60	0.72	0.84	0.97	1.09	1.21

[a]Best five solutions highlighted in bold.

Table 6. Root Mean Squared Error in kcal/mol after Fitting Mie Potentials Combined with PC Electrostatics Representations for Different Exponents of the Mie Potential and the FAM Dimer[a]

		m										
		12	11	10	9	8	7	6	5	4	3	2
n	11	0.67										
	10	0.66	0.65									
	9	0.65	0.64	0.63								
	8	0.64	0.63	0.61	0.60							
	7	0.63	0.61	0.60	0.58	0.56						
	6	0.62	0.59	0.57	0.56	0.54	0.53					
	5	0.60	0.57	0.55	0.53	0.52	0.53	0.54				
	4	0.57	0.54	0.52	0.51	0.51	0.53	0.56	0.61			
	3	0.54	0.52	**0.50**	**0.50**	0.51	0.55	0.60	0.68	0.76		
	2	0.52	**0.50**	**0.49**	**0.50**	0.53	0.59	0.67	0.76	0.86	0.97	
	1	0.52	**0.49**	**0.49**	0.52	0.57	0.65	0.75	0.86	0.98	1.09	1.20

[a]Best seven solutions highlighted in bold.

3) should not be compared with errors obtained from fitting the LJ potential (e.g ESPfit MTP + LJ(12,6): 0.25 kcal/mol, Table 2) because for the exponent search conformations with repulsive interaction energy between 0 and 10 kcal/mol they have not been down-weighted. Interaction energies for these conformations are usually reproduced the worst—therefore the RMSE of the exponent search fit is larger than the error from the direct fit of ε and σ with identical parameters. In addition, the two approaches cannot be compared because there are fewer degrees of freedom in fitting standard LJ potentials with ε and σ restrained to be positive.

The RMSEs for the FAM dimers with both ESP-fitted charge representations are shown in Tables 5 and 6. Since initially it seemed that there was a minimum at 9−4, the analysis for FAM was extended to all exponents from 1 to 12.

The best fit to *ab initio* energies for the FAM dimer has been obtained with the 10−2 (MTP and PC), 10−1 (PC), and 11−1 (PC) exponents for the Mie potential. Assuming that this is due to electrostatics not fully accounted for by the ESP-fitted charge model, an attempt was made to use a LJ potential augmented by a $1/r$ term. It was found that a 7−6−1 or a 8−5−1 combination and MTP electrostatics gave fits with the lowest

dx.doi.org/10.1021/ct300888f | J. Chem. Theory Comput. 2013, 9, 1499−1511

error (0.33 kcal/mol, 9−6−1: 0.34 kcal/mol). This shows that a contribution which decays with r^{-1}—i.e., electrostatics—is not fully accounted for by the charge scheme. Fixed PCs or MTPs are likely to be insufficient for correctly describing the electrostatics for the FAM dimer because in H-bonded systems polarization or charge-transfer become important. A strong indication for this can be seen from the error which for FAM is approximately twice as large as for BZN. Nevertheless, the overall interaction energies in the FAM dimer can be better reproduced with MTPs than with PCs. Also the r^{-12} term does not give the best fit—the best fits were obtained with a repulsive term proportional to r^{-7} or r^{-8}, closely followed by the r^{-9} term.

In summary, these considerations show that for the repulsive term, an exponent $m < 12$, in particular $m = 8$ or $m = 9$, gives the best agreement with *ab initio* energies.

3.4. Refitting 9−6 LJ Parameters. From the exponent search, it becomes clear that the r^{-12} term for the repulsive potential is too steep. We therefore examined the use of a 9−6 LJ potential instead of a 12−6 potential, and the fit was repeated.

The root mean squared errors for the SwissPARAM parameter set, the ESP-fitted electrostatics, and the ESP-fitted electrostatics in combination with refitted LJ parameters are shown in Table 7.

Table 7. Root Mean Squared Error [kcal/mol] for PC and MTP Electrostatics Representations with Standard 9-6 and Refitted 9-6 LJ Parameters[a]

monomer	charge model	charge parameters	SwissPARAM LJ	*ab initio* fit LJ
BZN	PC	SwissPARAM	0.55	−−[b]
		ESPfit	0.45	0.21
	MTP	GDMA	0.47	0.24
		ESPfit	0.46	0.22
FAM	PC	SwissPARAM	1.06	−−[b]
		ESPfit	1.49	0.50
	MTP	GDMA	1.57	0.54
		ESPfit	1.55	0.49

[a]ESPfit denotes charge parameters that have been obtained by refitting to the ESP. *Ab initio* fit means parameters have been obtained by refitting to *ab initio* interaction energies. GDMA PCs have not been examined. SwissPARAM only gives PCs, so there is no entry for SwissPARAM MTPs. [b]Not examined.

With a 9−6 LJ potential, the errors are typically lower than those from $(m,n) = (12,6)$. The exception is FAM with SwissPARAM LJ parameters, for which the error with 9−6 (1.06 kcal/mol) is somewhat larger than for 12−6 (0.90 kcal/mol). The best results for BZN are ESP-fitted PCs combined with the 9−6 LJ potential and *ab initio* fitted LJ parameters (RMSE = 0.21 kcal/mol, compared to 0.24 kcal/mol for the 12−6 potential). The best results for FAM are obtained with MTPs and the 9−6 LJ potential and *ab initio* fitted LJ parameters (RMSE = 0.49 kcal/mol, compared to 0.50 kcal/mol). With the 9−6 LJ potential, all results with ESP-fitted MTPs are slightly better than the results with GDMA MTPs.

3.5. Simultaneous Fit of Electrostatic and LJ Parameters. For deriving intermolecular interaction potentials, the ESP is used as a convenient measure to quantify the contribution of the electrostatic part to the total interaction energy. It does not necessarily give the best parameters possible for interaction energies, since for example mutual polarization is ignored. Also, the quality in reproducing the ESP close to the molecular surface is limited due to the PC/point MTP approximation and because different concepts and definitions of a molecular surface exist. For each of the dimer series and LJ exponents, we therefore *simultaneously* refitted all electrostatic and LJ parameters. The RMSEs obtained are summarized in Table 8.

If all intermolecular interaction parameters are fitted simultaneously, the result obtained with MTPs is clearly superior to that obtained with PCs. For the 9−6 LJ potential, this is even more pronounced (RMSE of 0.13 vs 0.19 kcal/mol for BZN and 0.19 vs 0.30 kcal/mol for FAM) than for the 12−6 LJ potential (RMSE of 0.18 vs 0.21 kcal/mol for BZN and 0.24 vs 0.30 kcal/mol for FAM). Overall, the results obtained with the 9−6 LJ potential are also better than those obtained with the 12−6 LJ potential. With refitted electrostatic parameters, the errors for FAM drop dramatically compared to the ESP-fitted electrostatic parameters.

All parameter fits have been run without restraints on the parameters with the exception of some atomic dipole and quadrupole parameters which could be set to zero due to symmetry.[15] Although restraints have not been used in the fit, the values of the parameters found are in the same overall range compared to the SwissPARAM parameters and physically reasonable. The final parameters obtained for the 9−6 LJ potential are reported in Table 9.

As a general trend, we find that the charges obtained within a MTP representation are considerably smaller than those obtained from SwissPARAM. The partial charges obtained from fitting MTPs to *ab initio* interaction energies are usually also smaller than PCs from fitting MTPs to the monomer ESP. This is in good agreement with the empirical practice of decreasing ESP-fitted charges in condensed-phase simulations by a certain factor in order to account for polarization damping effects. Interestingly, we find that the directional dipole and quadrupole parameters are proportionally less decreased than the charge parameters, when comparing ESP-fitted to *ab initio* fitted MTPs.

For BZN, the partial charges in a MTP representation are considerably smaller than both PCs from SwissPARAM and those optimized to *ab initio* interaction energies. This is evident for the carbon atoms on the aromatic ring (CB and CBB, see Figure 3), where the MTP charges are roughly one order of magnitude smaller than the PCs ($-0.018e$ vs $-0.15e$). If only PCs are used, partial charges need to be larger in order to capture the quadrupole of the aromatic ring. For MTP electrostatics, an appreciable magnitude for $Q_{20} = Q_{z^2}$ is found for the aromatic carbons CB and CBB, which represent the charge distribution of the aromatic system. Interestingly, both dipole and quadrupole parameters increase by a factor of two when going from monomer ESP-fitted MTPs to dimer *ab initio* interaction energy-fitted MTPs. Aromatic−aromatic interactions with their preferred face-to-face and edge-to-face orientation are highly anisotropic, and multipoles provide a natural framework for describing such interactions. They can most probably not be properly modeled with PCs alone. These benefits will have to be analyzed more carefully when developing general-purpose MTP force fields for practical applications. The charge on CSP completely vanishes, and all electrostatic interactions are mediated through the dipole and quadrupole moments. The μ_{10} coefficient for NSP describes the lone pair on the cyano-nitrogen. Interestingly, this dipole also

Table 8. Root Mean Squared Error [kcal/mol] for PC and MTP Electrostatics Representations with Standard and Refitted 12−6 and 9−6 LJ Parameters[a]

			12−6 LJ		9−6 LJ	
monomer	charge model	charge parameters	SwissPARAM	*ab initio* fit	SwissPARAM	*ab initio* fit
BZN	PC	SwissPARAM	0.69	−−[b]	0.55	−−[b]
		ESPfit	0.67	0.24	0.45	0.21
		ab initio fit	−−[b]	**0.21**	−−[b]	**0.18**
	MTP	GDMA	0.65	0.25	0.47	0.24
		ESPfit	0.67	0.25	0.46	0.22
		ab initio fit	−−[b]	**0.19**	−−[b]	**0.13**
FAM	PC	SwissPARAM	0.90	−−[b]	1.06	−−[b]
		ESPfit	1.14	0.56	1.49	0.50
		ab initio fit	−−[b]	**0.30**	−−[b]	**0.24**
	MTP	GDMA	1.13	0.55	1.57	0.54
		ESPfit	1.12	0.51	1.55	0.49
		ab initio fit	−−[b]	**0.30**	−−[b]	**0.19**

[a]ESPfit denotes charge parameters that have been obtained by refitting to the ESP. *Ab initio* fit means parameters have been obtained by refitting to *ab initio* interaction energies. GDMA PCs have not been examined. SwissPARAM only gives PCs, so there is no entry for SwissPARAM MTPs. [b]Not examined.

Table 9. ε, r_{min}, and Electrostatic Parameters Obtained from SwissPARAM and the Full Fit with a 9-6 LJ Potential[a]

molecule	atomtype	ε [kcal/mol]	r_{min}[Å]	q	μ_{10}	μ_{11c}	μ_{11s}	Q_{20}	Q_{21c}	Q_{21s}	Q_{22c}	Q_{22s}
BZN SwissPARAM	CB	−0.07	1.953	−0.15								
	CBB	−0.07	1.953	0.073								
	CSP	−0.068	2.038	0.484								
	HCMM	−0.022	1.294	0.15								
	NSP	−0.20	1.813	−0.557								
BZN Full LJ 9−6	CB	−0.135	1.781	−0.018	0.0	0.0	−0.305	−1.441	0.0	0.0	−0.880	0.0
	CBB	−0.095	1.960	−0.012	0.0	0.0	0.193	−0.775	0.0	0.0	0.992	0.0
	CSP	−0.172	1.830	0.0	0.0	0.0	0.154	0.502	0.0	0.0	0.050	0.0
	HCMM	−0.0005	2.072	0.065	−0.023	0.0	0.0	0.152	0.0	0.0	−0.078	0.0
	NSP	−0.029	2.064	−0.226	−0.380	0.0	0.0	0.203	0.0	0.0	0.061	0.0
FAM SwissPARAM	C═O	−0.11	2.0	0.57								
	HC═O	−0.022	1.32	0.06								
	HNCO	−0.046	0.225	0.37								
	NC═O	−0.20	1.85	−0.80								
	O═C	−0.12	1.70	−0.57								
FAM Full LJ 9−6	C═O	−0.008	2.491	0.070	0.0	−0.215	0.059	0.097	0.0	0.0	−0.313	−0.139
	HC═O	−0.013	1.224	0.136	0.023	0.138	0.0	0.166	−0.253	0.0	0.003	0.0
	HNCO	−0.256	0.526	0.269	0.099	−0.050	0.0	−0.006	0.068	0.0	0.027	0.0
	NC═O	−0.101	2.131	−0.391	0.0	0.203	−0.005	−0.911	0.0	0.0	−0.148	0.371
	O═C	−0.015	2.078	−0.353	−0.198	−0.033	0.0	0.450	0.002	0.0	−0.242	0.0

[a]All PC/MTP parameters are given in atomic units. For BZN, the charges do not exactly add to zero due to rounding.

increases when fitting to *ab initio* interaction energies, indicating that the directionality of the H-bond is not captured sufficiently well by fitting to the monomer-ESP only. For the aromatic hydrogens (HCMM) the fitted charge and dipole remain small when going from ESP-fitted parameters to parameters fitted to interaction energies. The quadrupole parameters increase by 50−100%, although they still remain relatively small. This shows that a substantial amount of the electrostatic interaction energy of the aromatic ring has to decay with r^{-5} through the quadrupoles, rather than with r^{-1} for the point charges. Although the ensemble of point charges at the aromatic ring also generates a quadrupole moment at a far distance, this is clearly not enough to properly model interaction energies at closer distances. The fitted r_{min} of the aromatic carbon "HCMM" is quite large with 2.072 Å, but this is balanced by a small ε = −0.0005 kcal/mol. It is important that the ε parameter is not zero, because otherwise the LJ potential would be zero and "nuclear fusion" due to attractive electrostatics could occur. Some parameters are correlated, so different combinations of parameter values lead to similar qualities of the fit. Also, the parameters for HCMM—and all other atom types—might change considerably when e.g. solvation effects are included. Such correlations could be handled by using restraints in the fit, but this will most probably not have a large effect on the fit quality and therefore does not affect the general conclusions of our study.

For FAM, the atomic PCs also decrease, and a considerable proportion of the electrostatic interaction is mediated through the dipoles and quadrupoles. The Q_{20} parameter for the amide nitrogen NC═O represents the conjugated π electrons. The partial charge on NC═O decreases by ∼10% when going from ESP-fitted MTPs to *ab initio* fitted MTPs, whereas the dipole

dx.doi.org/10.1021/ct300888f | J. Chem. Theory Comput. 2013, 9, 1499−1511

and the quadrupole coefficients remain the same. This again indicates the more directional interaction that can only be captured when fitting to interaction energies. The μ_{10}, Q_{20}, and the Q_{22c} parameters for the amide oxygen O═C represent the charge distribution of the two lone pairs ("rabbit ears") and the conjugated π electrons. Compared to the ESP-fitted MTP parameters, the charge decreases by 30%, the dipole remains the same, and the quadrupole decreases by 50%. The π electrons of the amide carbons C═O are captured by the Q_{22c} parameter. Nearly all MTP parameters on C═O, including the partial charge, decrease considerably compared to the ESP fitted parameters. The surface above this carbon is much smaller than the surface of all the other FAM atoms. Therefore, the parameters of this specific atom have a lower impact on the fit. In addition, some of the interaction might be captured by MTPs on the neighboring N and O atoms. For both types of hydrogens in FAM the parameters remain largely unchanged when comparing ESP fitted MTPs to *ab initio* fitted MTPs. The Q_{20} and the Q_{21c} parameter of HC═O hydrogen model the electron density around the hydrogen which can act as a donor. This indicates that here the H-bond donation is highly directional as well. In summary, the directionality of a hydrogen bond in an H-bonded dimer increases compared to the monomer. Quantitatively, this is reflected in decreased partial charges, whereas higher MTPs on heavy atoms remain unchanged or increase. Consequently, the importance of higher MTPs for intermolecular interactions increases. The most significant changes upon fitting MTP parameters to *ab initio* energies are observed on the heavy atoms. MTP parameters on hydrogen change only slightly.

Overall, the multipole parameters found for both molecules can be related to anisotropic chemical features like lone pairs or π systems. However, it needs to be kept in mind that some parameters are correlated, and therefore the absolute values have to be interpreted with caution. In addition, PCs and MTPs are approximations, and most of the chemical questions can only be answered in condensed-phase simulations.

For both molecules, the dipole and quadrupole parameters on the hydrogens are relatively small, and they could probably be omitted for practical applications. This analysis however would go beyond the scope of the present work. In general, the full fit of all parameters with the 9−6 LJ potential yields a set of physically interpretable parameters whereas with the 12−6 LJ potential, the parameters obtained (details not shown) take on more extreme values and are physically more questionable.

4. DISCUSSION

In this contribution, we have provided a quantitative analysis of using MTPs up to quadrupoles instead of PCs in describing intermolecular interactions for two prototypical dimers in many different relative orientations. If the 12−6 LJ potential is replaced by a 9−6 LJ potential, the error in reproducing *ab initio* interaction energies by using MTPs instead of PCs can be reduced by up to 38%. If the 12−6 LJ potential is used, the maximal error reduction is ∼16%. Therefore, we reason that if MTPs are to be introduced in force fields, the repulsive r^{-12}− term in the LJ potential should be replaced by a less steep potential, such as an r^{-9} term.

We have shown that it is mandatory to refit both electrostatic and LJ parameters at the same time. Fitting the charge model to the electrostatic potential and combining it with available LJ parameters does not show clear superiority of MTPs for interaction energies. Also, refitting LJ parameters to reproduce *ab initio* energies in combination with an ESP-fitted charge model does not clearly show the superiority of the MTPs. Only when fitting both LJ and electrostatic parameters together MTPs are really capable of better reproducing *ab initio* interaction energies compared to PCs. Usually, such fits are done stepwise, but our results indicate that standard force fields, if they are fitted to interaction energies, would also benefit from a simultaneous fit of all parameters. Especially for the FAM dimer where H-bonding considerably contributes to the interaction energies, a pronounced improvement was observed when the charge model was refit to the *ab initio* energies. This might be due to the fact that a monomer-fitted ESP is not sufficient in strongly H-bonded systems, as charge-transfer and polarization may also significantly affect to the interaction energy.

For this study, we decided to focus on LJ and electrostatic parameters, because they are the most commonly and widely used representations for intermolecular interactions in a standard force field. Thus, the "quality considerations" for PC representations are directly relevant to current applications of standard force fields in all branches of the chemical, physical, and biological sciences. In particular, it is not the aim to specifically parametrize a condensed-phase force field for BZN and FAM. Rather, we want to examine in a well-controlled manner the effect of replacing PCs by MTPs starting from the observation that the ESP for both monomers can be much better represented by MTPs rather than PCs. We find that although the ESP is better captured by about one order of magnitude by MTPs, average improvements for interaction energies are only about 30% (or less than 0.1 kcal/mol) if electrostatics and LJ terms are fitted simultaneously. In other words, carefully optimizing PCs together with LJ terms along the lines proposed in the present work can yield accurate interaction energies for a wide range of applications.

The analysis presented implicitly assumes that the error introduced by ignoring polarizability is similar for both PCs and MTPs. The gain in accuracy obtained for refitting MTP parameters to the *ab initio* energies might be due to better capturing polarization because MTPs intrinsically have more degrees of freedom. We cannot control this effect in this study, but it even more emphasizes the finding that MTPs are not clearly superior to PCs in reproducing *ab initio* energies if the electrostatic parameters are fitted to the ESP and combined with either standard or refitted LJ parameters. In principle, MTPs should be superior to PCs if they are combined with the correct dispersion, repulsion, and polarization potential. However, finding an acceptable trade-off between potential complexity and computational overhead for practical MD simulations is not trivial, and our study underlines that just replacing one part of the total potential by another more complex potential does not promise to give better results.

It should be noted that there is experimental evidence that higher order (e.g., C_8, C_{10}) and anisotropic (e.g., $C_{6.0}$, $C_{6.2}$, $C_{6.4}$) dispersion terms exist,[53] and for highly accurate force fields they need to be included as well.[54]

Using a pairwise $(A/r^m)-(B/r^n)$ formulation of the LJ potential, the Mie potential, we were able to rapidly scan all possible integer (m,n) combinations. This scan showed that r^{-9} for the repulsive part of the potential best reproduces interaction energies. For FAM, the exponent-scan indicated that the charge contribution was not fully captured by the ESP, because it was beneficial to include a term $\propto r^{-1}$ for the attractive part. This was automatically adjusted in the

simultaneous fit of all electrostatic and LJ parameters. For BZN, the exponent scan showed that r^{-6} is the best exponent for the long-range attractive part of the LJ potential. The pairwise formulation of the LJ potential is very convenient for deriving parameters, since it can be fit rapidly and efficiently with a direct linear least-squares. Also, pairwise LJ parameters might be able to compensate for short-range polarization effects and allow one to additionally improve fitting the interaction energies.

The parameters obtained with the full fit of the 9−6 LJ potential and the MTPs are physically reasonable and in the same range as in other comparable force fields, except for the hydrogen parameters of HCMM, the aromatic hydrogen of BZN. Here, the optimum ε found is very small (ε = 0.0005 kcal/mol). This indicates that the contribution of the electrons around the HCMM atoms to the dispersion energy is extremely small. We cannot rule out that it could be even smaller, but since we use a LJ term to model dispersion and repulsion, $\varepsilon > 0$ is required. A small LJ-ε is also consistent with previous force field parametrizations such as OPLS where ε = 0.0 kcal/mol is used for hydrogen atoms and additional controls have to be used in the force field engine to prevent nuclear fusion.[55,56]

The main difference between optimized PCs and MTPs is that point charges in the MTP representation are considerably smaller in magnitude compared to a PC-only parametrization. In a PC-only force field they have to compensate for the anisotropic charge distribution around atoms, and therefore their magnitude increases accordingly. The aromatic carbon and the attached hydrogen in BZN are a good example for this, where PCs decrease from −0.15*e* to −0.018*e* (CB) and 0.15*e* to 0.065*e* (HCMM), if MTPs up to quadrupoles are used.

The findings of the present study are important for future force field developments that use MTP electrostatics. First, as others before us also noted,[46,57−63] the present work confirms that even with multipolar electrostatics the 12−6 LJ potential should be replaced by a 9−6 form to better reproduce interaction energies. Second, we have shown that if the charge model is changed from PCs to MTPs, LJ parameters have to be readjusted. This is in line with a recent study which showed that in polarizable force fields it is also necessary to reparametrize the LJ terms.[64] This and our findings support the notion that LJ parameters are not independent of it used and have to be consistent with the charge model. Third and most significantly, we have shown that although the monomer-ESP is a rapid and convenient starting point for a parametrization, it is insufficient if reliable dimer-interaction energies are required. Therefore, it is expected that future force field development initiatives will considerably benefit from fitting to interaction energies from ensembles of dimer structures. This finding poses new challenges, because it means that in standard biomolecular force fields all intermolecular parameters should be fitted simultaneously to balance them against one another.

5. CONCLUSIONS

In answering the question initially asked, the present study does not find a clear superiority of monomer- and ESP-fitted MTPs over PCs if combined with a 12−6 LJ potential. On the other hand, refitting MTPs to reproduce *ab initio* interaction energies leads to considerably improved force field energies compared to refitted PCs. Specifically, we find that the LJ/MTP combination reproduces *ab initio* interaction energies with up to 30% less error than the LJ/PC combination. However, this improvement can only be achieved if both electrostatic and LJ parameters are fit at the same time. We have reconfirmed[46] that a 9−6 LJ potential better reproduces the nonelectrostatic component of *ab initio* interaction energies than the standard 12−6 LJ potential. In addition, we have generalized this finding to multipolar electrostatic interactions: The benefits of using MTPs over PCs are more evident if a 9−6 LJ potential is used rather than a 12−6 form. Finally, the findings of the present contribution are important for future force-field developments. The results suggest that, at least in standard force fields without polarizability, the parameters of nonbonded interactions should be fitted simultaneously and that even replacing PCs by MTPs from ESPs together with sequential refitting of LJ parameters is not expected to yield substantial improvement.

AUTHOR INFORMATION

Corresponding Author

*E-mail: christian.kramer@novartis.com; peter.gedeck@novartis.com; m.meuwly@unibas.ch.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

C.K. thanks the Novartis Institutes for BioMedical Research for a Presidential PostDoc Fellowship. M.M. thanks the Swiss National Science Foundation (Grant 200020-132406 and the NCCR MUST) for continuous financial support.

REFERENCES

(1) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.

(2) Ponder, J. W.; Case, D. A. Force fields for protein simulations. *Adv. Prot. Chem.* **2003**, *66*, 27−85.

(3) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.

(4) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656−1676.

(5) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.

(6) Jones, J. E. On the Determination of Molecular Fields. II. From the Equation of State of a Gas. *Proc. R. Soc. London, Ser. A* **1924**, *106*, 463−477.

(7) 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*, 2549−2564.

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

(9) Neumann, M. A. Tailor-Made Force Fields for Crystal-Structure Prediction. *J. Phys. Chem. B* **2008**, *112*, 9810−9829.

(10) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J.-P. Anisotropic, Polarizable Molecular Mechanics Studies of Inter- and

Intramolecular Interactions and Ligand-Macromolecule Complexes. A Bottom-Up Strategy. *J. Chem. Theory Comput.* **2007**, *3*, 1960−1986.
(11) Piquemal, J.-P.; Jordan, K. D. From quantum mechanics to force fields: new methodologies for the classical simulation of complex systems. *Theor. Chem. Acc.* **2012**, 131.
(12) Price, S. L.; Leslie, M.; Welch, G. W. A.; Habgood, M.; Price, L. S.; Karamertzanis, P. G.; Day, G. M. Modelling organic crystal structures using distributed multipole and polarizability-based model intermolecular potentials. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8478−8490.
(13) Leslie, M. A molecular dynamics program to use distributed multipole electrostatic models to simulate the dynamics of organic crystals. *Mol. Phys.* **2008**, *106*, 1567−1578.
(14) Plattner, N.; Meuwly, M. Higher order multipole moments for molecular dynamics simulations. *J. Mol. Model.* **2009**, *15*, 687−694.
(15) Kramer, C.; Gedeck, P.; Meuwly, M. Atomic multipoles: Electrostatic potential fit, local reference axis systems, and conformational dependence. *J. Comput. Chem.* **2012**, *33*, 1673−1688.
(16) Hennemann, M.; Murray, J. S.; Politzer, P.; Riley, K. E.; Clark, T. Polarization-induced s-holes and hydrogen bonding. *J. Mol. Model.* **2012**, *18*, 2461−2469.
(17) Liem, S. Y.; Shaik, M. S.; Popelier, P. L. A. Aqueous Imidazole Solutions: A Structural Perspective from Simulations with High-Rank Electrostatic Multipole Moments. *J. Phys. Chem. B* **2011**, *115*, 11389−11308.
(18) Devereux, M.; Plattner, N.; Meuwly, M. Application of multipolar charge models and molecular dynamics simulations to study stark shifts in inhomogeneous electric fields. *J. Phys. Chem. A* **2009**, *113*, 13199−13209.
(19) Plattner, N.; Meuwly, M. The role of higher CO-multipole moments in understanding the dynamics of photodissociated carbon-monoxide in myoglobin. *Biophys. J.* **2008**, *94*, 2505−2515.
(20) Özpinar, G. A.; Beierlein, F. R.; Peukert, W.; Zahn, D.; Clark, T. A test of improved force field parameters for urea: molecular-dynamics simulations of urea crystals. *J. Mol. Model.* **2012**, *18*, 3455−3466.
(21) Devereux, M.; Meuwly, M. Force Field Optimization using Dynamics and Ensemble Averaged Data: Vibrational Spectra and Relaxation in Bound MbCO. *J. Chem. Inf. Model.* **2010**, *50*, 349−357.
(22) Jelsch, C.; Teeter, M. M.; Lamzin, V.; Pichon-Pesme, V.; Blessing, R. H.; Lecomte, C. Accurate protein crystallography at ultra-high resolution: Valence electron distribution in crambin. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 3171−3176.
(23) Mulliken, R. S. Electronic Population Analysis on LCAO[Single Bond]MO Molecular Wave Functions. I. *J. Chem. Phys.* **1955**, *23*, 1833−1841.
(24) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry Introduction to Advanced Electronic Structure Theory*, unabridged unaltered republication of the 1st ed., originally published in New York, 1989; Dover Publ.: Mineola, NY, 1996.
(25) Singh, U. C.; Kollman, P. A. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **1984**, *5*, 129−145.
(26) Breneman, C. M.; Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* **1990**, *11*, 361−373.
(27) Stone, A. Distributed multipole analysis, or how to describe a molecular charge distribution. *Chem. Phys. Lett.* **1981**, *83*, 233−239.
(28) Stone, A. *The Theory of Intermolecular Forces*; Oxford University Press: New York, 1997.
(29) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; 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.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen,W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(30) Stone, A. J. Distributed Multipole Analysis: Stability for Large Basis Sets. *J. Chem. Theory Comput.* **2005**, *1*, 1128−1132.
(31) Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. SwissParam: A fast force field generation tool for small organic molecules. *J. Comput. Chem.* **2011**, *32*, 2359−2368.
(32) Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490−519.
(33) MacKerell, A. D.; Bashford, D.; Bellott; Dunbrack, R. L.; 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. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(34) Boys, S. F.; Bernardi, F. Calculation of Small Molecular Interactions by Differences of Separate Total Energies - Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19*, 553−566.
(35) Simon, S.; Duran, M.; Dannenberg, J. J. How does basis set superposition error change the potential surfaces for hydrogen-bonded dimers? *J. Chem. Phys.* **1996**, *105*, 11024−11031.
(36) Hobza, P.; Zdenvk, H. Counterpoise-corrected potential energy surfaces of simple H-bonded systems. *Theor. Chem. Acc.* **1998**, *99*, 372−377.
(37) Schütz, M.; Hetzer, G.; Werner, H.-J. Low-order scaling local electron correlation methods. I. Linear scaling local MP2. *J. Chem. Phys.* **1999**, *111*, 5691−5705.
(38) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(39) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(40) Marchetti, O.; Werner, H.-J. Accurate calculations of intermolecular interaction energies using explicitly correlated wave functions. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3400−3409.
(41) Hrouda, V.; Florian, J.; Polasek, M.; Hobza, P. Double Proton Transfer: From the Formamide Dimer to the Adenine···Thymine Base Pair. *J. Phys. Chem.* **1994**, *98*, 4742−4747.
(42) Hargis, J. C.; Volhringer-Martinez, E.; Woodcock, H. L.; Toro-Labbet', A.; Schaefer, H. F. Characterizing the Mechanism of the Double Proton Transfer in the Formamide Dimer. *J. Phys. Chem. A* **2011**, *115*, 2650−2657.
(43) Cato, J.; Majumdar, D.; Roszak, S.; Leszczynski, J. Exploring Relative Thermodynamic Stabilities of Formic Acid and Formamide Dimers Role of Low-Frequency Hydrogen-Bond Vibrations. *J. Chem. Theory Comput.* **2013**, DOI: 10.1021/ct300889b.
(44) Hazra, M. K.; Chakraborty, T. Formamide Tautomerization: Catalytic Role of Formic Acid. *J. Phys. Chem. A* **2005**, *109*, 7621−7625.
(45) Mallajosyula, S. S.; Guvench, O.; Hatcher, E.; MacKerell, A. D. CHARMM Additive All-Atom Force Field for Phosphate and Sulfate Linked to Carbohydrates. *J. Chem. Theory Comput.* **2012**, *8*, 759−776.
(46) Warshel, A. Consistent Force Field Calculations. II. Crystal Structures, Sublimation Energies, Molecular and Lattice Vibrations, Molecular Conformations, and Enthalpies of Alkanes. *J. Chem. Phys.* **1970**, *53*, 582−594.

(47) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford Science Publications, Oxford University Press: Oxford, U. K., 1989.
(48) Baker, C. M.; Lopes, P. E. M.; Zhu, X.; Roux, B.; MacKerell, A. D. Accurate Calculation of Hydration Free Energies using Pair-Specific Lennard-Jones Parameters in the CHARMM Drude Polarizable Force Field. *J. Chem. Theory Comput.* **2010**, *6*, 1181−1198.
(49) Bondi, A. van der Waals Volumes and Radii. *J. Phys. Chem.* **1964**, *68*, 441−451.
(50) Jones, E.; Oliphant, T.; Peterson, P. *SciPy: Open source scientific tools for Python*, Version 0.9.0. http://www.scipy.org (accessed 1/11/2013).
(51) Mie, G. Zur kinetischen Theorie der einatomigen Körper. *Ann. Phy.* **1903**, *316*, 657−697.
(52) R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2011; ISBN 3-900051-07-0.
(53) Hutson, J. M.; Ernesti, A.; Law, M. M.; Roche, C. F.; Wheatley, R. J. The intermolecular potential energy surface for CO2ŰAr: Fitting to high-resolution spectroscopy of Van der Waals complexes and second virial coefficients. *J. Chem. Phys.* **1996**, *105*, 9130−9140.
(54) Buckingham, A. D.; Fowler, P. W.; Hutson, J. M. Theoretical studies of van der Waals molecules and intermolecular forces. *Chem. Rev.* **1988**, *88*, 963−988.
(55) Jorgensen, W. L.; Tirado-Rives, J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6665−6670.
(56) Jorgensen, W. L.; Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(57) Heinz, H.; Vaia, R. A.; Farmer, B. L.; Naik, R. R. Accurate Simulation of Surfaces and Interfaces of Face-Centered Cubic Metals Using 12−6 and 9−6 Lennard-Jones Potentials. *J. Phys. Chem. C* **2008**, *112*, 17281−17290.
(58) Ahmed, A.; Sadus, R. J. Solid-liquid equilibria and triple points of n-6 Lennard-Jones fluids. *J. Chem. Phys.* **2009**, *131*, 174504−174511.
(59) White, D. N. A computationally efficient alternative to the Buckingham potential for molecular mechanics calculations. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 517−521.
(60) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. Molecular mechanics. The MM3 force field for hydrocarbons. 1. *J. Am. Chem. Soc.* **1989**, *111*, 8551−8566.
(61) Engler, E. M.; Andose, J. D.; Schleyer, P. V. R. Critical evaluation of molecular mechanics. *J. Am. Chem. Soc.* **1973**, *95*, 8005−8025.
(62) Sun, H.; Mumby, S. J.; Maple, J. R.; Hagler, A. T. An ab Initio CFF93 All-Atom Force Field for Polycarbonates. *J. Am. Chem. Soc.* **1994**, *116*, 2978−2987.
(63) Sun, H. COMPASS: An ab Initio Force-Field Optimized for Condensed-Phase Applications Overview with Details on Alkane and Benzene Compounds. *J. Phys. Chem. B* **1998**, *102*, 7338−7364.
(64) Wang, J.; Cieplak, P.; Li, J.; Cai, Q.; Hsieh, M.-J.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical Calculations. 4. van der Waals Parametrization. *J. Phys. Chem. B* **2012**, *116*, 7088−7101.