Deriving Static Atomic Multipoles from the Electrostatic Potential

Christian Kramer,[*,†] Tristan Bereau,[*,‡] Alexander Spinn,[†] Klaus R. Liedl,[†] Peter Gedeck,[¶] and Markus Meuwly[‡]

[†]Institute of General, Inorganic and Theoretical Chemistry, and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, 6020 Innsbruck, Austria

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

[¶]Novartis Institutes for BioMedical Research, 4002 Basel, Switzerland

S Supporting Information

ABSTRACT: The description of molecular systems using multipolar electrostatics calls for automated methods to fit the necessary parameters. In this paper, we describe an open-source software package that allows fitting atomic multipoles (MTPs) from the ab initio electrostatic potential by adequate atom typing and judicious assignment of the local axis system. By enabling the simultaneous fit of several molecules and/or conformations, the package addresses issues of parameter transferability and lack of sampling for buried atoms. We illustrate the method by studying a series of small alcohol molecules, as well as various conformations of protonated butylamine.
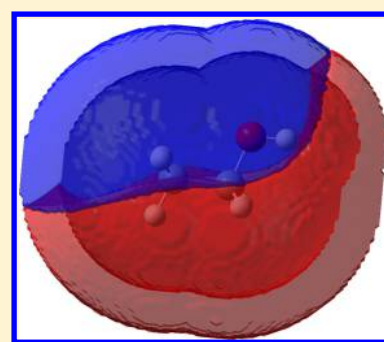
1. INTRODUCTION

For decades, standard biochemical molecular force fields have relied on point charges to describe electrostatics.[1−9] Although the deficiencies of point charges in accurately describing a molecule's electrostatic potential have long been known,[10,11] more refined descriptions—such as atomic multipole moments (MTPs)—have only gained broader attention in the past decade.[12−22] The advent of MTPs calls for automated methods to determine suitable parameters in a convenient, transparent, reliable, and reproducible way. In recent years, a number of groups have aimed at deriving MTPs from the electron density[23−33]—Stone's Gaussian Distributed Multipole Analysis (GDMA) algorithm[34] likely being the most popular method.

Point charges calculated from partitioning the electron density[35−37] can vary widely depending on the technology used to derive them, probably because the property surface is very flat. They may also have serious deficiencies in reproducing the electrostatic potential (ESP) around the molecule. As a result, techniques that directly fit point charges to the ESP have been developed (e.g., refs 38−40). We recently showed that the same holds for MTPs, where ESP-fitted MTPs can better reproduce the ESP by up to 1 order of magnitude.[41] In 1995, Koch, Popelier, and Stone showed that MTP moments can be highly conformation dependent[42]—leading to a lack of transferability of electrostatic parameters across a distribution of conformations. The specificity of MTP parameters for flexible molecules can lead to extreme cases: we showed that a PC-only ESP fit averaged over several conformations can, in some cases, perform better than the MTP parametrization of a single conformation.[41] In this context, we have developed a software package that allows the determination of atomic MTP moments by simultaneously fitting ensembles of conformations and molecules to the ESP.

There are previous attempts for more general fitting environments for MTPs including pdm97,[43] PolType,[44,45] OPEP,[46] an unnamed code by Chodkiewicz,[47] and a framework for fitting and handling Gaussian multipoles.[48] What sets the current environment apart from these efforts is the fact that the present implementation allows and demonstrates the utility of fitting MTPs to an ensemble of conformations across a number of chemically related molecules, which aims at and leads to improved transferability of the parameters.

After describing the necessary steps to fit MTP parameters to the ESP in section 2, we will discuss applications of the software package that illustrate the benefits of fitting multiple conformations and molecules at once.

2. WORKFLOW

The process of fitting atomic multipoles to the ESP consists of several steps, which are schematically illustrated in Figure 1.

The individual steps are discussed below. All scripts implementing those steps have been written in Python, making extensive use of the numpy and scipy libraries[49] and RDKit.[50]

2.1. 3D Conformer. The first step is the generation of one or several conformers of the desired molecule. There is a huge variety of programs available—both commercial and noncommercial[50−53]—to generate relevant molecular representations. Workflows and comparisons of these programs have
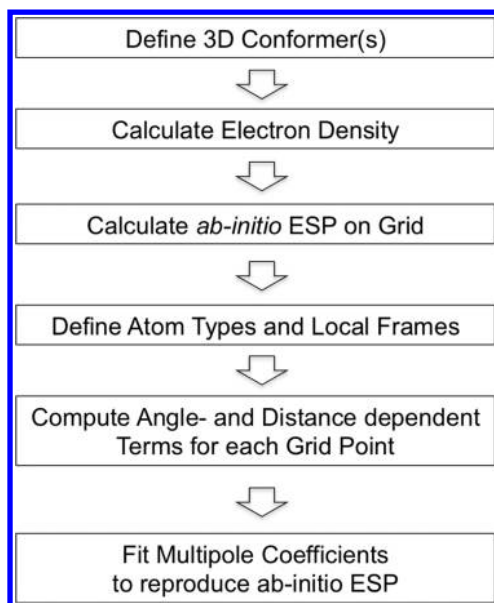
Figure 1. Workflow for fitting multipole coefficients to the ab initio ESP.

Figure 2. Automatized atom-typing-assignment scheme.

applications, we use Gaussian09[57] for density functional theory calculations at the M062X level of theory[58] and the aug-cc-PVDZ basis set.[59] Evidently, such choices should be adapted to the size and chemical properties (e.g., radicals) of the molecules. The electron density calculation step is automated in the `calc_Multipole_Gauss_MEP_comp.py` script and specifics can be controlled from the command line.

2.3. ESP on Grid. The Gaussian09 utility cubegen extracts the ESP on a rectangular grid up to 5 Å away from the molecule. The grid spacing can be specified on the command line. While the default is 0.1 Å, we have found that a grid spacing of 0.4 Å yields almost identical results when fitting MTPs, while lowering the computational cost by a factor of $4^3 = 64$. For graphical representation of the ab initio ESP, the MTP ESP, and differences between both, we recommend using the default 0.1 Å spacing. The cubegen calculation is automated in the `calc_Multipole_Gauss_MEP_comp.py` script.

2.4. Atom Types and Local Frames. Local reference-axis frames are defined following the work of Kramer et al.[41] The neighboring atoms that define the frame are automatically assigned using the `calc_LRA.py` script. This script also assigns atom types in a two-step procedure, which is schematically shown in Figure 2.

First, every atom is characterized according to its chemical element and the number of bonded neighbors (C4, for instance, is a carbon with four neighbor atoms). For terminal atoms, the number of neighbors is not assigned (e.g., "Cl" instead of "Cl1") and aromatic carbons and nitrogens are special atom types (e.g., "Car" and "Nar"). Other special atom types can be added or deleted by editing the accompanying `essential_atomtypes.txt` file. It contains a list of atom types with SMARTS definitions,[60] where definitions higher up in the list have higher priority. SMARTS matching is done using the `rdkit` Python module.[50]

Second, the full atom type is generated as a list of the atom type itself and its nearest neighbors (e.g., "C4HHN3C4" for an aliphatic carbon bound to two hydrogens, one nitrogen and another aliphatic carbon). The list is sorted with decreasing atomic number, whereas identical neighbors are moved to the front of the list (e.g., the two Hs in the above-mentioned example). For terminal atoms, the entry second to the primary element/neighbor specification is the nearest neighbor (by definition of terminal, there can only be one); and further, in the same order as described above, the second nearest neighbors (e.g., "ClC4HHF" for the chlorine in chloro-fluoromethane).

Beyond what is described in the work of Kramer et al.,[41] we have added a local reference-axis system for diatomics and linear molecules (e.g., acetylene) that relies on the nearest neighbor(s) to define a local z-axis, while x- and y-axes are arbitrary. The symmetry of the system sets all MTP moments that are *not* along z to zero.

For atom typing, the location of formal charges is read from the corresponding `.sdf` file, which has to be created in advance (e.g., using OpenBabel). All atoms that carry a formal charge or that are part of a charge-carrying conjugated system are assigned a "+" or "−" sign. For instance, the primary nitrogen of a protonated primary amine becomes "N4+", and the full atom type for the amine nitrogen on a charged aliphatic primary amine becomes "N4+HHHC4". Protonated histidine is an example of a charged conjugated system, where all heavy atoms in the ring get assigned a "+" and both nitrogens are atom-typed "Nar+Car+Car+H".
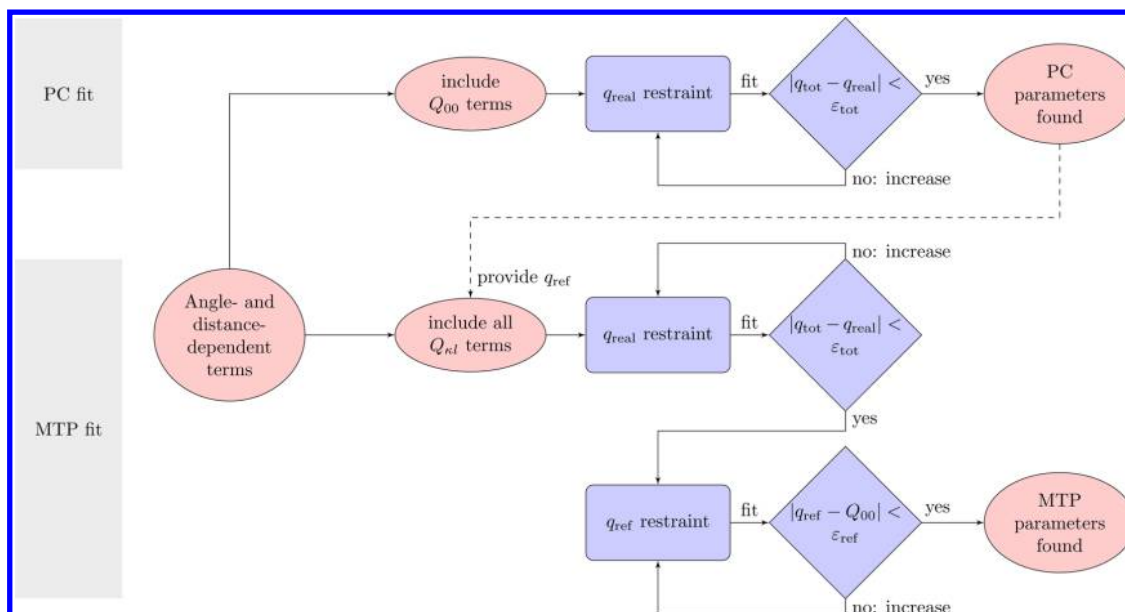
been published elsewhere.[54,55] For the present ESP fitting workflow, conformers must be generated with one of these external programs. For instance, a SMILES string can be converted to a simple 3D conformer by using the `--gen3D` option in OpenBabel.[56] When following such a strategy, we suggest to subsequently minimize the generated conformation to the next local minimum. This minimization step is automated in the `calc_Multipole_Gauss_MEP_comp.py` script (provided) using Gaussian09.

2.2. Calculate Electron Density. In order to calculate the ab initio ESP, we first determine the electron density. For the present

Figure 3. MTP-fitting workflow used in this work. First, only point charges are fitted to the ESP (top). The algorithm iteratively restrains the difference between the molecule's net charge, q_{real}, and the sum of the fitted monopoles, q_{tot}, within the threshold ε_{tot}. Second, all MTP parameters are jointly fitted to the ESP, using the reference PCs obtained. Here, two different restraints are subsequently adjusted: (i) we again control the total charge, as in step 1; and (ii) we force the MTP parameters to take small values and the PCs to remain close to the above-mentioned reference values, within the threshold ε_{ref}.

2.5. Precompute Geometrical Terms. The ESP at any grid point $\mathbf{r}^{(p)}$ can be calculated from

$$\Phi(\mathbf{r}^{(p)}) = \sum_i \sum_j Q_j^{(i)} f_j^{(i)}(\mathbf{r}^{(p)})$$

$$\approx \sum_i Q_{00}^{(i)} r^{-1} + Q_{10}^{(i)} r^{-2} \hat{r}_z + Q_{11c}^{(i)} r^{-2} \hat{r}_x + Q_{11s}^{(i)} r^{-2} \hat{r}_y$$

$$+ Q_{20}^{(i)} r^{-3} (3\hat{r}_z^2 - 1)/2 + Q_{21c}^{(i)} r^{-3} \sqrt{3} \hat{r}_x \hat{r}_z$$

$$+ Q_{21s}^{(i)} r^{-3} \sqrt{3} \hat{r}_y \hat{r}_z + Q_{22c}^{(i)} r^{-3} \sqrt{3} (\hat{r}_x^2 - \hat{r}_y^2)/2$$

$$+ Q_{22s}^{(i)} r^{-3} \sqrt{3} \hat{r}_x \hat{r}_y \tag{1}$$

where i iterates over all atoms, j iterates over all MTP coefficients, and $f_j^{(i)}(\mathbf{r}^{(p)})$ are geometrical factors, including distance- and angular-dependent terms of the MTP moment $Q_j^{(i)}$ at point p. We use the equations given by Stone[11] to calculate the ESP at $\mathbf{r}^{(p)}$ from MTPs up to quadrupoles where r denotes the vector from atom i to $\mathbf{r}^{(p)}$; $r = \|\mathbf{r}\|$; $\hat{r}_a = \mathbf{r} \cdot \hat{a}/r$, where $\hat{a}$ refers to one of the unit vectors x, y, or z; and $Q_{\kappa l}^{(i)}$ is the lth MTP moment of rank κ in spherical coordinates. In terms of Cartesian coordinates, $Q_{00}^{(i)}$ corresponds to the point charge; $\kappa = 1$ with $l = 0, 1c, 1s$ are the components of the dipole moment vector in the z-, x-, and y- directions, respectively; and $\kappa = 2$ with $l = 0, 1c, 1s, 2c, 2s$ corresponds to the z^2-, xz-, yz-, x^2-y^2-, and xy- components of the quadrupole tensor (all from the point of view of the atom's local frame), respectively.

Since the conformations are given a priori, the geometrical terms $f_j^{(i)}(\mathbf{r}^{(p)})$ in eq 1 can be precalculated, leaving a simple system of linear equations for the MTP parameters $Q_{\kappa l}^{(i)}$ to be solved. If two or more atoms in a molecule have the same atom type, the geometrical factors for any grid point can simply be added. The precomputation of these terms is done by the `mtp_fittab_maker.py` script.

By default, the script stores all grid points located in the neighborhood of the molecule: we only consider grid points that are within $2.2\sigma_i$ of *any* atom, but larger than $1.66\sigma_i$ of *all* the atoms, where σ_i represents the van-der-Waals radius of atom i (see ref 41). Though the present methodology fails to reproduce the ESP at short distances, such close intermolecular contacts are unlikely to occur in actual simulations due to the presence of van der Waals repulsion. At large distances (i.e., $r > 2.2\sigma_i$), the ESP rapidly decays to zero; this region does not influence the fit significantly and is typically well reproduced by the present methodology.

If specified, the `calc_LRA.py` script also computes contributions due to bond increments, instead of point charges. Bond increments express the charge *difference* between two bonded atoms—thus yielding a zero net molecular charge by construction. For charged molecules, however, point charges have to be used. We advise against the combination of bond increments and point charges in the fit, since they are highly correlated.

2.6. Fit MTP Coefficients. After the angle- and distance-dependent terms have been precalculated, we group eq 1 in terms of atom types t. We then fit the MTP coefficients $Q_j^{(t)}$ to the collection of ESP grid points p by optimizing the target function

$$\chi^2 = \min \sum_p (\Phi_{abinitio}(\mathbf{r}^{(p)}) - \Phi_{MTP}(\mathbf{r}^{(p)})) \tag{2}$$

which minimizes the error between the ab initio and MTP-derived ESPs. Because the problem is linear, we can rewrite the problem as $\mathbf{Xb} = \mathbf{y}$, where the matrix coefficient $X_{j,p} = f_j^{(t)}(\mathbf{r}^{(p)}) = \sum_{\tau \in t} f_j^{(\tau)}(\mathbf{r}^{(p)})$, b is the vector of MTP coefficients $\{Q_j^{(t)}\}$—one element for each MTP coefficient of every atom type—and $\mathbf{y}$ is the vector of ab initio calculated ESPs (one entry for each grid point p). The linearity of the problem guarantees us to find the global minimum. To make optimal use of the likely sparsity of $\mathbf{X}$ (i.e., $\{Q_j^{(t)}\}$ may be involved in the calculation of few grid points p), we instead solve the equivalent problem
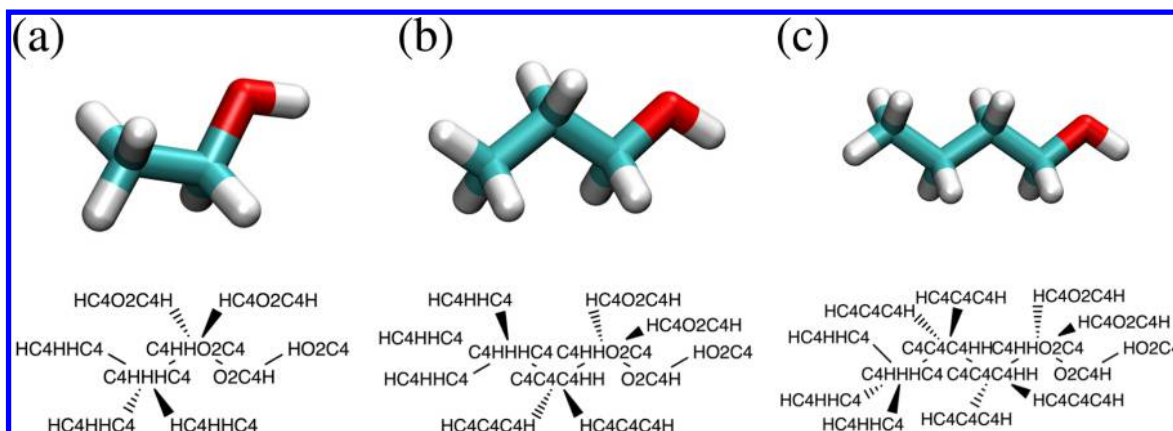
Figure 4. Molecules used in the simultaneous-molecule fit: (a) ethanol, (b) propanol, and (c) butanol. (bottom) Atom-type assignment. Figures were generated with VMD.[63]

$$\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{y} \tag{3}$$

where $\mathbf{X}^T$ is the transpose of $\mathbf{X}$. We explicitly use the formulation in eq 3, since it effectively reduces the dimensionality from p (which can be hundreds of thousands of grid points) to maximally nine (number of MTP parameters) times the number of atom types. The procedure used can be described as constrained least-squares, with the constraints discussed below. We use the `scipy.linalg.solve` solver that is based on the lapack `dgesv` routine for solution-stability reasons.

To provide realistic MTP parameters, we systematically start by fitting the simpler PC-only problem. ESP fits are known to yield "unphysical" values for buried atoms.[40] Here, we leverage the information contained in different molecules and/or conformations by (i) grouping atoms into types and (ii) simultaneously fitting multiple conformers. Technically, this is handled by joining the precomputed matrices that contain the angle- and distance-dependent terms by column names. This way, it is not necessary to match or search for equivalent grid points from different conformations, since the ESP for each conformation is sufficiently sampled with its own grid points.

During the fit, restraints are applied to address two issues: (a) the total charge of the molecule should be correctly reproduced and (b) charge and MTP parameters should have realistic (i.e., "physical") values. The first restraint, δ, penalizes deviations of the sum of the fitted point charges from the overall charge of the molecule. For each molecule, one row p_m is added to $\mathbf{X}$ that counts the number of atom types present on the molecule and scales them all by the same penalty factor: $X_{j_t p_m} = \sum_{\tau \in t} \delta$ if j_t is a monopole term and 0 otherwise. Also, we add an entry δq_{real}, where q_{real} is the molecule's net charge, to y. Given a sufficiently large value of δ, this restraint will provide an accurate net charge. The restraint is iteratively increased until the maximum deviation from the correct charge across molecules is lower than a user-defined threshold, ε_{tot}: $\max_m|q_{\text{real}}^{(m)} - q_{\text{tot}}^{(m)}| < \varepsilon_{\text{tot}}$, where m sums over molecules and $q_{\text{tot}}^{(m)} = \sum_{\tau \in m} Q_{00}^{(i)}$ is the total charge of the molecule from the PC parameters. By default, $\varepsilon_{\text{tot}} = 10^{-12}$ e is used.

Next, we apply restraints to the MTP fit that penalize deviations of the PC values from the above-mentioned PC-only fit and higher MTPs from zero. This second restraint is iteratively increased until all PCs are within a user-defined threshold, ε_{ref}, of the reference values q_{ref}: $\max_i|q_{\text{ref}}^{(t)} - q_{00}^{(t)}| < \varepsilon_{\text{ref}}$ (by default: $\varepsilon_{\text{ref}} = 0.1$ e). The restraint on higher multipoles is set to one tenth of the

restraint on PC parameters and iteratively increased with the restraint on PCs. In our experience, one tenth is sufficient since the higher MTP moments converge to reasonably low values much faster than PCs to their target value. The fitting workflow is schematically shown in Figure 3.

Note that instead of performing an initial PC-only fit, user-specified reference charges can be provided. We point out that a significantly more robust fit can be achieved from a collection of molecules/conformations for which every atom type is significantly sampled (i.e., close to the surface of the molecule). This strategy alleviates biases from a priori guesses.

2.7. Additional Programs. For an initial guess of the MTPs, the `calc_Multipole_Gauss_MEP_comp.py` script calls the Gaussian Distributed Multipole Analysis (GDMA) program.[61] We supply three additional programs that allow further visualization and numerical evaluation of the fitted parameters. The first program, `mtp_prm.to.pun.py`, applies MTPs obtained through the fit on a given molecule and rotates the parameters to the global frame. Should a molecule not be part of the parametrization set, the parameters are likely to yield an incorrect net charge—generating an excess net charge $\Delta q = q_{\text{real}} - q_{\text{tot}}$. In this case, the script can scale all charges

$$q_i^{(\text{scaled})} = \Delta q \frac{|q_i|}{\sum_j q_j} \tag{4}$$

where i runs over all atoms in the molecule.

We then use `fieldcomp` to compare the MTP-based ESP to its ab initio counterpart. For instance, it provides the means to compare the transferability of MTP parameters between different conformations or molecules (see section 3). The function `fieldcomp` also writes `.cube` files that contain ESPs calculated from the PC or MTP representation and the difference between the ab initio and the PC/MTP ESP. The `.cube` files can be visualized using for example `gaussview`.[57]

3. APPLICATIONS

The following highlights the strengths of the methodology by presenting and discussing two example applications: simultaneous fit of (i) multiple chemically related molecules and (ii) multiple conformations. A combined fit of both—a number of molecules in different conformations—has already been carried out in previous work.[62]
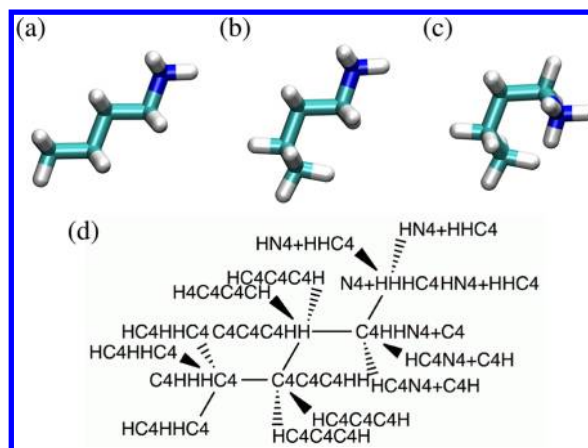
Table 1. Transferability of ESP-optimized MTPs for Ethanol, Propanol, and Butanol[a]

	molecule optimized			
probe	ethanol	propanol	butanol	all
ethanol	**0.06**	2.38	1.14	**0.07**
propanol	—	**0.05**	0.61	**0.06**
butanol	—	2.73	**0.06**	**0.06**

[a]The column refers to the molecule optimized to fit the MTP parameters, while the row corresponds to the molecule that is probed. All RMS errors between ab initio and fitted ESPs are given in kilocalories per mole. Figures below 0.1 kcal/mol appear in bold. The parameters derived for ethanol did not provide enough atom types (see Table 2) to reconstruct propanol and butanol—denoted here by long dashes.

Figure 5. Conformations of *n*-butylamine: (a) all trans, (b) gauche conformation of the methyl group, and (c) gauche conformation of both the methyl and the amine groups. (d) Atom-type assignment.

All MTP fits that follow were performed with moments up to quadrupoles, using density functional theory calculations with the M06-2X functional and the aug-cc-PVDZ basis set.

3.1. Fitting Multiple Molecules: Ethanol, Propanol, and Butanol. As a first application, MTP parameters are fitted to the ESPs of ethanol, propanol, and butanol (Figure 4)—both individually and simultaneously. Table 1 highlights the lack of transferability of fitting MTPs to one molecule and using the resulting parameters in a different, albeit chemically closely related compound. While fitting MTPs to one molecule allows to reproduce its own ESP well (root-mean-squared (RMS) error below 0.1 kcal/mol), the error increases sharply when the parameters are transferred to another molecule. It is also worth noting that ethanol does not contain a sufficient

number of atom types to describe propanol and butanol (see Table 2 and Figure 4).

Fitting all molecules simultaneously, on the other hand, allows to find MTP parameters that reproduce the ESP of all molecules with satisfying transferability across molecules (Table 1). For brevity, the present ESP-based fits shown here do include MTPs on the hydrogens, but Table 4 shows the resulting parameters of a fit with PC-only hydrogens—limiting the number of parameters to display. We point out, however, that the other MTP parameters do not vary appreciably, whether MTPs are used for hydrogens or not. Moreover, we

Table 2. ESP-Fit MTPs for the Heavy Atoms Only (i.e., PC for the Hydrogens) Fitted to Ethanol (Figure 5a), Propanol (Figure 5b), Butanol (Figure 5c), and All Conformations Simultaneously[a]

type						fitted molecule(s)			
		neighbors				ethanol	propanol	butanol	all
C4	H	H	H	C4	Q_{00}		+0.28	+0.13	−0.09
					Q_{10}		−0.15	−0.01	−0.09
					Q_{22c}		−0.03	−0.08	−0.00
					Q_{00}	−0.47	−0.56	−0.47	−0.09
					Q_{10}	+0.05	−0.03	+0.13	+0.08
					Q_{20}	+0.01	−0.10	+0.02	−0.09
C4	H	H	O2	C4	Q_{00}	+0.36	+0.12	+0.31	+0.41
					Q_{10}	+0.21	+0.16	+0.10	+0.14
					Q_{11s}	−0.01	+0.20	+0.12	+0.10
					Q_{20}	−0.05	−0.12	−0.24	−0.06
					Q_{21s}	−0.05	−0.09	−0.07	−0.12
					Q_{22c}	+0.28	+0.34	+0.27	+0.28
H	C4	C4	C4	H	Q_{00}		+0.02	−0.02	+0.05
H	C4	H	H	C4	Q_{00}	+0.12	+0.13	+0.13	+0.03
H	C4	O2	C4	H	Q_{00}	−0.01	+0.02	−0.04	−0.06
H	O2	C4			Q_{00}	+0.36	+0.41	+0.40	+0.37
O2	C4	H			Q_{00}	−0.60	−0.71	−0.70	−0.68
					Q_{11c}	−0.06	−0.03	−0.04	−0.03
					Q_{11s}	−0.07	−0.01	+0.00	−0.04
					Q_{20}	−0.47	−0.35	−0.30	−0.37
					Q_{22c}	+0.20	+0.15	+0.08	+0.15
					Q_{22s}	−0.08	+0.10	+0.05	+0.01

[a]With RMS error in the first interaction belt $\overline{\Delta E}$ = 0.28, 0.25, 0.42, and 0.43 kcal/mol, respectively. Corresponding fits with MTPs on all atoms (used for 1, but not shown in this table for brevity) yield RMS errors $\overline{\Delta E}$ = 0.06, 0.05, 0.06, and 0.06 kcal/mol, respectively. Q_{lk} denotes the coefficient of the ESP's multipole expansion with l referring to the level of expansion of the ESP when performing the fit (e.g., l = 0 corresponds to PCs).

Table 3. Transferability of GDMA-Optimized MTPs for the Conformations of *n*-Butylamine[a]

probe	molecule optimized						
	no. 1	no. 2	no. 3	no. 4	no. 5	no. 6	no. 7
no. 1	**0.40**	10.30	10.26	11.37	11.32	4.03	4.01
no. 2	10.14	**0.36**	9.38	3.84	11.82	13.11	11.25
no. 3	10.14	9.36	**0.36**	11.90	3.80	11.27	13.13
no. 4	11.27	3.99	10.99	**0.31**	11.84	13.04	10.48
no. 5	11.26	10.94	3.94	11.86	**0.31**	10.55	12.98
no. 6	4.05	12.18	11.03	12.40	10.69	**0.37**	4.07
no. 7	4.03	10.92	12.17	10.66	12.43	4.09	**0.37**

[a]The column refers to the conformation optimized to derive the GDMA parameters, while the row corresponds to the conformation that is probed. All RMS errors between ab initio and fitted ESPs are given in kilocalories per mole. Figures below 1 kcal/mol appear in bold.

note that the MTP parameters *benefit* from being fitted to multiple molecules. This can be illustrated by considering the carbon atoms (Table 2): while individual fits provide comparatively large partial charges, the simultaneous fit yields

charges that are close to 0—in agreement with, e.g., the GROMOS force field.[64] In addition, fitting several molecules at once with a consistent atom-typing scheme has obvious advantages when performing alchemical transformations, where a minimal number of electrostatic parameters are changed upon single-atom substitutions.[62]

3.2. Fitting Multiple Conformations: Protonated *n*-Butylamine. As a second application, the dependence of electrostatic parameters on conformational changes of the protonated (i.e., charged) state of *n*-butylamine (Figure 5) is studied. Seven conformations are selected. They include an all-trans conformation, and several combinations of gauche rotations for the methyl and amine groups. We probe MTP-parameter transferability across conformations and monitor the quality of the ESP using both GDMA (Table 3) and the present methodology (Table 4). The lack of transferability manifests itself through the large variation in RMS error against the ab initio ESP, i.e., up to a factor of 30 for both methods. In addition, the lack of transferability makes MTP parameters *worse* than a PC optimization which fits all conformations

Table 4. Transferability of ESP-Optimized MTPs for the Conformations of *n*-Butylamine[a]

probe	molecule(s) optimized							
	no. 1	no. 2	no. 3	no. 4	no. 5	no. 6	no. 7	all
no. 1	**0.05**	10.30	10.27	11.41	11.37	4.04	4.02	**0.41**
no. 2	10.12	**0.13**	9.41	3.85	11.82	13.12	11.27	**0.28**
no. 3	10.12	9.38	**0.14**	11.91	3.81	11.30	13.15	**0.28**
no. 4	11.29	3.99	11.03	**0.28**	11.86	13.05	10.47	**0.61**
no. 5	11.28	10.98	3.95	11.88	**0.28**	10.54	13.00	**0.61**
no. 6	4.08	12.19	11.04	12.41	10.70	**0.09**	4.08	**0.31**
no. 7	4.06	10.93	12.18	10.67	12.43	4.09	**0.09**	**0.30**

[a]The column refers to the conformation(s) used to fit the MTP parameters, while the row corresponds to the conformation that is probed. All RMS errors between ab initio and fitted ESPs are given in kilocalories per mole. Figures below 1 kcal/mol appear in bold.

Table 5. ESP-Fit MTPs for the Heavy Atoms Only (i.e., PC for the Hydrogens) Fitted to Conformation Number 4 (Figure 5b), conformation Number 6 (Figure 5c), and All Conformations Simultaneously[a]

type	neighbors					fitted conformation(s)		
						no. 4	no. 6	all
C4	C4	C4	H	H	Q_{00}	−0.09	−0.15	−0.21
					Q_{10}	+0.03	−0.03	+0.11
					Q_{22c}	−0.13	+0.01	−0.31
C4	H	H	H	C4	Q_{00}	−0.30	+0.04	−0.36
					Q_{10}	+0.05	+0.12	−0.04
					Q_{20}	−0.26	−0.05	−0.58
C4	H	H	N4+	C4	Q_{00}	+0.05	−0.07	+0.05
					Q_{10}	−0.15	+0.02	−0.10
					Q_{11s}	+0.02	+0.14	−0.03
					Q_{20}	−0.80	−0.20	−0.85
					Q_{21s}	−0.23	−0.03	−0.20
					Q_{22c}	−0.22	+0.08	−0.29
H	C4	C4	C4	H	Q_{00}	+0.09	+0.09	+0.13
H	C4	H	H	C4	Q_{00}	+0.11	+0.03	+0.14
H	C4	N4+	C4	H	Q_{00}	+0.10	+0.12	+0.12
H	N4+	H	H	C4	Q_{00}	+0.38	+0.34	+0.40
N4+	H	H	H	C4	Q_{00}	−0.59	−0.39	−0.62
					Q_{10}	+0.03	+0.09	+0.03
					Q_{20}	+0.05	−0.09	−0.07

[a]With RMS error in the first interaction belt $\overline{\Delta E}$ = 0.91, 0.50, and 0.79 kcal/mol, respectively. Corresponding fits with MTPs on all atoms (used for four, but not shown in this table for brevity) yield RMS errors $\overline{\Delta E}$ = 0.37, 0.13, 0.53 kcal/mol, respectively. Q_{lk} denotes the coefficient of the ESP's multipole expansion with l referring to the level of expansion of the ESP when performing the fit (e.g., l = 0 corresponds to PCs).

simultaneously (data not shown) and corroborates earlier results.[41] In strong contrast to MTP fits on single conformations, a simultaneous fit achieves the best overall performance (Table 4, column "all") with errors between 10 and 50 times lower compared to the individual fits. The atom-type assignments for n-butylamine are shown in Figure 5d. MTP parameters that fit (i) conformation no. 4, (ii) conformation no. 6, and (iii) all conformations simultaneously are shown in Table 5. The table clearly shows the strong variation of certain MTP parameters when fitting different conformations.

4. CONCLUSIONS

In the present paper we have presented a workflow to obtain MTP parameters for one/several molecules including multiple conformations. In particular, the code automatizes: (i) the extraction of the ESP from an ab initio calculation, (ii) the assignment of atom types and reference-axis systems that take advantage of an atom's chemical environment, and (iii) the simultaneous fit of various molecules/conformations to assign MTP parameters. We highlight and demonstrate the benefits of fitting multiple molecules and/or conformations at once, which elegantly attenuates possible conformational sampling issues of buried atoms. For the specific case of fitting multiple alcohols simultaneously a significant improvement in the transferability of parameters across molecules was found, while fitting an ensemble of representative conformations of charged n-butylamine provided good transferability across conformations, highlighting the strength of the methodology for flexible molecules. Other examples can be found in ref 41.

The MTP parameters obtained can be used in the MTPL CHARMM module[65] that uses the same reference axis system assignment. In future work, we plan to systematically investigate the transferability of MTP parameters derived from our methodology. Further future extensions include different reference axis system assignments and atomic polarizabilities in the fitting procedure.

The source code of the present package is available from https://bitbucket.org/tbereau/esp-fit.pc-mtp/ under a public license.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

Code used to fit static MTP moments and to do the analysis under an Apache 2 license. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Authors
*E-mail: christian.kramer@uibk.ac.at.
*E-mail: bereau@alumni.cmu.edu.

Notes
The authors declare no competing financial interest.

■ REFERENCES

(1) 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.

(2) 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.

(3) Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490−519.

(4) Ponder, J. W.; Case, D. A.; Daggett, V. *Protein Simulations*; Academic Press, 2003; Vol. 66; pp 27−85.

(5) 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.

(6) Mackerell, A. D., Jr. Empirical force fields for biological macromolecules: Overview and issues. *J. Comput. Chem.* **2004**, *25*, 1584−1604.

(7) Jorgensen, W. L.; Tirado-Rives, J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Nat. Acad. Sci.* **2005**, *102*, 6665−6670.

(8) 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.

(9) MacKerell, A. D.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(10) Stone, A. Distributed multipole analysis, or how to describe a molecular charge distribution. *Chem. Phys. Lett.* **1981**, *83*, 233−239.

(11) Stone, A. *The theory of intermolecular forces*; Oxford University Press: New York, 1997.

(12) 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.

(13) Stone, A. J. Intermolecular potentials. *Science* **2008**, *321*, 787−789.

(14) 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.

(15) Kdzierski, P.; Sokalski, W. A. Analysis of the transferability of atomic multipoles for amino acids in modeling macromolecular charge distribution from fragments. *J. Comput. Chem.* **2001**, *22*, 1082−1097.

(16) Plattner, N.; Meuwly, M. Higher order multipole moments for molecular dynamics simulations. *J. Mol. Model.* **2009**, *15*, 687−694.

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

(18) 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.

(19) Day, G. M.; Motherwell, W. D. S.; Jones, W. Beyond the Isotropic Atom Model in Crystal Structure Prediction of Rigid Molecules: Atomic Multipoles versus Point Charges. *Cryst. Growth Des.* **2005**, *5*, 1023−1033.

(20) Piquemal, J.; Chevreau, H.; Gresh, N. Toward a Separate Reproduction of the Contributions to the Hartree-Fock and DFT Intermolecular Interaction Energies by Polarizable Molecular Mechanics with the SIBFA Potential. *J. Chem. Theory Comput.* **2007**, *3*, 824−837.

(21) 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.

(22) Kazantsev, A. V.; Karamertzanis, P. G.; Adjiman, C. S.; Pantelides, C. C. Efficient Handling of Molecular Flexibility in Lattice Energy Minimization of Organic Crystals. *J. Chem. Theory Comput.* **2011**, *7*, 1998−2016.

(23) Whitehead, C. E.; Breneman, C. M.; Sukumar, N.; Ryan, M. D. Transferable atom equivalent multicentered multipole expansion method. *J. Comput. Chem.* **2003**, *24*, 512−529.

(24) Sokalski, W.; Keller, D.; Ornstein, R.; Rein, R. Multipole correction of atomic monopole models of molecular charge distribution. I. Peptides. *J. Comput. Chem.* **1993**, *14*, 970−976.

(25) Ritchie, J. P.; Copenhaver, A. S. Comparison of potential-derived charge and atomic multipole models in calculating electrostatic potentials and energies of some nucleic acid bases and pairs. *J. Comput. Chem.* **1995**, *16*, 777−789.

(26) Williams, D. E. Representation of the molecular electrostatic potential by atomic multipole and bond dipole models. *J. Comput. Chem.* **1988**, *9*, 745−763.

(27) Soederhjelm, P.; Krogh, J. W.; Karlstroem, G.; Ryde, U.; Lindh, R. Accuracy of distributed multipoles and polarizabilities: Comparison between the LoProp and MpProp models. *J. Comput. Chem.* **2007**, *28*, 1083−1090.

(28) Bader, R. *Atoms in molecules: a quantum theory*; Oxford University Press: New York, 1994.

(29) Sokalski, W. A.; Shibata, M.; Rein, R.; Ornstein, R. L. Cumulative atomic multipole moments complement any atomic charge model to obtain more accurate electrostatic properties. *J. Comput. Chem.* **1992**, *13*, 883−887.

(30) 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.

(31) Piquemal, J.-P.; Jordan, K. D. From quantum mechanics to force fields: new methodologies for the classical simulation of complex systems. *Theo. Chem. Acc.* **2012**, *131*, 1207.

(32) Ferenczy, G. G.; Winn, P. J.; Reynolds, C. A. Toward Improved Force Fields. 2. Effective Distributed Multipoles. *J. Phys. Chem. A* **1997**, *101*, 5446−5455.

(33) Vigne-Maeder, F.; Claverie, P. The exact multicenter multipolar part of a molecular charge distribution and its simplified representations. *J. Chem. Phys.* **1988**, *88*, 4934−4948.

(34) Stone, A. J. Distributed Multipole Analysis: Stability for Large Basis Sets. *J. Chem. Theory Comput.* **2005**, *1*, 1128−1132.

(35) Mulliken, R. S. Electronic Population Analysis on LCAO[Single Bond]MO Molecular Wave Functions. I. *J. Chem. Phys.* **1955**, *23*, 1833−1841.

(36) Gasteiger, J.; Marsili, M. A new model for calculating atomic charges in molecules. *Tetrahedron Lett.* **1978**, *19*, 3181−3184.

(37) Foster, J. P.; Weinhold, F. Natural hybrid orbitals. *J. Am. Chem. Soc.* **1980**, *102*, 7211−7218.

(38) Singh, U. C.; Kollman, P. A. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **1984**, *5*, 129−145.

(39) 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.

(40) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.

(41) 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.

(42) Koch, U.; Popelier, P. L. A.; Stone, A. J. Conformational dependence of atomic multipole moments. *Chem. Phys. Lett.* **1995**, *238*, 253−260.

(43) Williams, D. E. Pdm97, Least-Squares Fitting of the Molecular Electrostatic Potential With Net Atomic Charges and/or Multipoles. 1997; http://www.netsci.org/Resources/Software/Modeling/QM/pdm97.html.

(44) 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.

(45) Wu, J. C.; Chattree, G.; Ren, P. Automation of AMOEBA polarizable force field parameterization for small molecules. *Theo. Chem. Acc.* **2012**, *131*, 1138.

(46) Angyan, J. G.; Chipot, C.; Dehez, F.; Haettig, C.; Jansen, G.; Millot, C. OPEP: A tool for the optimal partitioning of electric properties. *J. Comput. Chem.* **2003**, *24*, 997−1008.

(47) Chodkiewicz, M. L. Compact multipolar representation of the electrostatic potential for flexible molecules. *J. Chem. Phys.* **2005**, *122*, 224107−224107−8.

(48) Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. Generalization of the Gaussian electrostatic model: Extension to arbitrary angular momentum, distributed multipoles, and speedup with reciprocal space methods. *J. Chem. Phys.* **2006**, *125*, 184101.

(49) Jones, E.; Oliphant, T.; Peterson, P. SciPy: Open source scientific tools for Python. 2001; http://www.scipy.org/.

(50) Landrum, G.; et al., RDKit. 2013; www.rdkit.org.

(51) Miteva, M. A.; Guyon, F.; Tufféry, P. Frog2: Efficient 3D conformation ensemble generator for small compounds. *Nucleic Acids Res.* **2010**, *38*, W622−W627.

(52) O'Boyle, N. M.; Vandermeersch, T.; Flynn, C. J.; Maguire, A. R.; Hutchison, G. R. Confab - Systematic generation of diverse low-energy conformers. *J. Cheminf.* **2011**, *3*, 8.

(53) Vainio, M. J.; Johnson, M. S. Generating Conformer Ensembles Using a Multiobjective Genetic Algorithm. *J. Chem. Inf. Model.* **2007**, *47*, 2462−2474.

(54) Bolton, E. E.; Kim, S.; Bryant, S. H. PubChem3D: Conformer generation. *J. Cheminf.* **2011**, *3*, 4.

(55) Ebejer, J.-P.; Morris, G. M.; Deane, C. M. Freely Available Conformer Generation Methods: How Good Are They? *J. Chem. Inf. Model.* **2012**, *52*, 1146−1158.

(56) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *J. Cheminf.* **2011**, *3*, 33.

(57) Frisch, M. J. et al. *Gaussian 09*, Revision A.1. 2009; Gaussian Inc.: Wallingford CT, 2009.

(58) Zhao, Y.; Truhlar, D. G. Exploring the Limit of Accuracy of the Global Hybrid Meta Density Functional for Main-Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2008**, *4*, 1849−1868.

(59) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(60) *SMARTS Theory Manual*; Daylight Chemical Information Systems: Santa Fe, NM, 2007.

(61) Stone, A. J. GDMA 2.2.06. 2012; http://www.stone.ch.cam.ac.uk/pub/gdma/download.php.

(62) Bereau, T.; Kramer, C.; Monnard, F. W.; Nogueira, E. S.; Ward, T. R.; Meuwly, M. Scoring Multipole Electrostatics in Condensed-Phase Atomistic Simulations. *J. Phys. Chem. B* **2013**, *117*, 5460−5471.

(63) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33−38.

(64) Daura, X.; Mark, A. E.; Van Gunsteren, W. F. Parametrization of aliphatic CHn united atoms of GROMOS96 force field. *J. Comput. Chem.* **1998**, *19*, 535−547.

(65) Bereau, T.; Kramer, C.; Meuwly, M. Leveraging Symmetries of Static Atomic Multipole Electrostatics in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, DOI: 10.1021/ct400803f.

3417

dx.doi.org/10.1021/ci400548w | *J. Chem. Inf. Model.* 2013, 53, 3410−3417