# Evaluation of Representations and Response Models for Polarizable Force Fields

Amanda Li,[†] Alexey Voronin,[‡,§] Andrew T. Fenley,[‡] and Michael K. Gilson[*,‡]

[†]Department of Bioengineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0419, United States

[‡]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0736, United States

[§]Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, California 94550, United States

**ABSTRACT:** For classical simulations of condensed-phase systems, such as organic liquids and biomolecules, to achieve high accuracy, they will probably need to incorporate an accurate, efficient model of conformation-dependent electronic polarization. Thus, it is of interest to understand what determines the accuracy of a polarizable electrostatics model. This study approaches this problem by breaking polarization models down into two main components: the *representation* of electronic polarization and the *response model* used for mapping from an inducing field to the polarization within the chosen representation. Among the most common polarization representations are redistribution of atom-centered charges, such as those used in the fluctuating charge model, and atom-centered point dipoles, such as those used in a number of different polarization models. Each of these representations has been combined with one or more response models. The response model of fluctuating charge, for example, is based on the idea of electronegativity equalization in the context of changing electrostatic potentials (ESPs), whereas point-dipole representations typically use a response model based on point polarizabilities whose induced dipoles are computed based on interaction with other charges and dipoles. Here, we decouple polarization representations from their typical response models to analyze the strengths and weaknesses of various polarization approximations. First, we compare the maximal possible accuracies achievable by the charge redistribution and point-dipole model *representations*, by testing their ability to replicate quantum mechanical (QM) ESPs around small molecules polarized by external inducing charges. Perhaps not surprisingly, the atom-centered dipole model can yield higher accuracy. Next, we test two of the most commonly used *response functions* used for the point-dipole representations, self-consistent and direct (or first-order) inducible point polarizabilities, where the polarizabilities are optimized to best fit the full set of polarized QM potentials for each molecule studied. Strikingly, the induced-dipole response model markedly degrades accuracy relative to that obtainable with optimal point dipoles. In fact, the maximal accuracy achievable with this response model is even lower than that afforded by an optimal charge-redistribution representation. This means that, if coupled with a sufficiently accurate response function, the point-charge representation could outperform the standard induced-dipole model. Furthermore, although a key advantage of the point-dipole representation, relative to charge redistribution, is its ability to capture out-of-plane polarization, the inducible dipole response model causes it to be less accurate than optimal charge redistribution for out-of-plane induction of the planar nitrobenzene molecule. Thus, the widely used inducible dipole response function falls short of the full potential accuracy achievable with the point-dipole representation it employs. Additional results reported here bear on the relative accuracy of self-consistent inducible dipoles versus that of the first-order, or direct, approximation and on methods for assigning partial atomic charges for use in conjunction with inducible dipole models. In sum, these results point to the improvement of polarization response models as an important direction for future research aimed at improving the accuracy of molecular simulations.
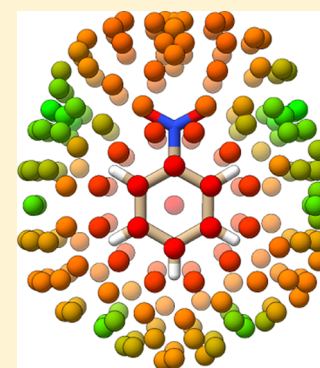
## INTRODUCTION

Classical simulations of condensed-phase molecular systems rely on potential functions, or force fields, to map the spatial coordinates of the atomic nuclei to the potential energy of the system and the force on each atom. The commonly used force fields model electrostatic interactions in terms of Coulombic interactions among atom-centered point charges with fixed values.[1−8] This functional form strikes a practical balance between computational efficiency and accuracy and thus has found wide applications. Indeed, we have argued that, given the

modest scope of experimental data used so far to adjust force-field parameters, force fields using this functional form can become even more accurate.[9] On the other hand, their accuracy must ultimately be limited by their neglect of configuration-

dependent changes in electronic polarization, that is, shifts in molecular electron densities induced by variations in the electrical field felt by a molecule, as the system evolves over time. This limitation of the fixed-charge model has motivated the development of force fields that incorporate configuration-dependent electronic polarization.[10−15]

The functional form used to model electronic polarization keeps the partial charges constant and adds atom-centered dipoles, whose moments vary with the local electric field. A well-regarded version of this functional form uses atom-centered point polarizabilities, where the field felt by each point polarizability is that generated by the point charges and the other induced dipoles in the system. This implementation is computationally burdensome, however, because it requires solving a matrix equation for the self-consistent set of induced dipoles. This problem may be moderated using improved mathematical approaches[16] or solved entirely using variants in which the field felt by each point polarizability omits any contribution from the other induced dipoles.[17−20] These direct, or first-order, methods are fast because there is no self-consistent matrix problem to solve, but the lack of physical consistency between fields and dipoles might lead to reduced accuracy. On the other hand, even the full, self-consistent point-polarizability model is itself a simplified representation of relatively complex electron population shifts and going from self-consistency to the direct approximation may not add much more error. Another way to avoid solving the self-consistent induced-dipole problem is to use the Drude oscillator[21−23] or the charge-on-spring[24] model. This approximates atom-centered point polarizabilities by attaching an artificial particle, with a small mass and a point charge, to each atom treated as polarizable and including the motions of these charges as part of the dynamical system. It is also important to mention more detailed models, such as SIBFA,[25] which place anisotropic polarizabilities off atom centers and carry out treatments of electronic polarization based on a continuum dielectric representation.[26−28]

Another functional form used to model electronic polarization does not use point dipoles but instead allows partial atomic charges to vary in response to time-varying electric fields. For example, the fluctuating charge implementation of this approach[29,30] uses an electronegativity equalization[31] ansatz to demonstrate how charges vary with field and treats the changes in charge as additional dynamical variables via an extended Lagrangian method. However, other polarization implementations based on variable partial charges would also be possible.[32]

The applicability of these two basic models, inducible point dipoles and variable point charges, raises the question whether one is fundamentally better suited than the other to model shifts in the electron density of a molecule induced by external fields. That is, setting aside how these quantities are assigned in any given implementation, is there any difference in the abilities of atom-centered point charges and point dipoles to model the changes in electrical field due to induced polarization? This question originally arose in discussions about aqueous nitrobenzene: perhaps a configuration with a water molecule hydrogen-bonded to only one nitro oxygen would lead to a redistribution of electrons between the two oxygens that would not be readily captured by atom-centered dipoles (William Swope, personal discussion). On the other hand, it is well known that atom-centered point charges are not well suited to capture out-of-plane polarization of a planar molecule, such as

nitrobenzene. Additionally, given the greater computational cost of the self-consistent induced-dipole model compared to that of the direct approximation, it is worth asking how much accuracy is lost in moving to the direct model. Finally, while formulating a polarization model based on atom-centered polarizabilities, one must choose between simply overlaying the new polarizabilities on an existing set of fixed partial atomic charges (e.g., restrained electrostatic potential (RESP) charges[33]) or optimizing a new set of charges for use with the polarizability model.

Here, we address these and related issues by computing changes in molecular electrostatic potentials (ESPs) induced by point charges at multiple locations around small molecules in vacuum, using electronic structure methods and then testing how well-optimized implementations of various polarizability models replicate these changes. The results have implications for the formulation of force fields that account for configuration-dependent electronic polarization.

## ■ METHODS

The basic approach taken here is to assess how well various polarization models can replicate ESPs, computed by quantum mechanical (QM) methods, around molecules polarized by artificial inducing charges. Our use of polarized QM ESPs as a reference for polarizability models follows the work of others.[23,34] We regard each polarization model as consisting of a *representation* of polarization, in terms of shifts in atom-centered point charges or in terms of added atom-centered point dipoles, and a *response model*, which is a recipe for computing these charges or dipoles. For example, fluctuating charge[29,30] is a model that represents polarization via changes in point charges and that uses an electronegativity-equalization response model to derive these charges.[31] Similarly, both the self-consistent and direct polarization models represent polarization in terms of atom-centered point dipoles, but they use somewhat different response models to compute the induced dipoles. We studied the following polarization models:

**Model 0: Global optimized point charges.** A single charge set is optimized to best replicate the full set of polarized ESPs. Chemically equivalent atoms are constrained to have equal charges. It should be emphasized that the charges in this model are constant across all locations of the inducing charge. In addition, there is no response model for computing the charges beyond fitting to QM results, so this model could not be used in an actual simulation. However, this model is informative as it reveals the greatest accuracy attainable using a single set of permanent point charges to capture the various polarized states of a molecule.

**Model 1: Optimal point charges.** Polarization is represented by changes in atom-centered point charges, and, unlike model 0, a different charge set is optimized to best replicate each polarized ESP. There is no response model for computing the charges beyond fitting to QM results, so this model could not be used in an actual simulation. However, this model demonstrates the greatest accuracy attainable with any model using the point-charge representation of polarization.

**Model 2: RESP charges and optimal point dipoles.** RESP charges[33] are assigned on the basis of the unpolarized ESP and then held constant, whereas a different set of atom-centered point dipoles is optimized
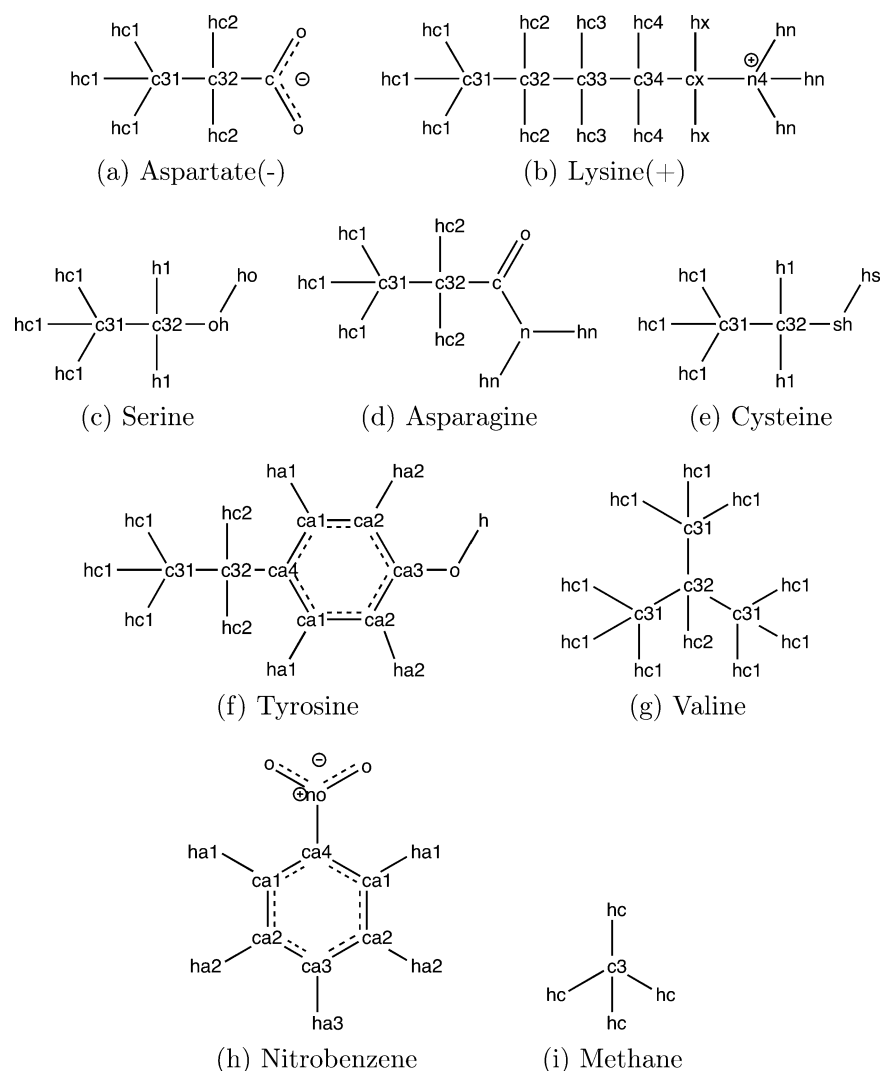
**Figure 1.** Molecular structures with atom types. Where multiple nonequivalent atoms have the same force field type, such as "c3", an additional integer is added so that only chemically equivalent atoms have the same type for the present study. Note that (a)−(g) are amino acid side-chain analogues rather than full amino acids and that atom types are molecule-specific.

to best replicate each polarized ESP. As for optimal point charges (above), there is no response model for computing the dipoles beyond fitting to QM results; this model could not be used in a simulation. However, it reveals the greatest accuracy attainable by the point-dipole representation of polarizability, in the context of the baseline RESP charges.

**Model 3: RESP charges and direct polarizabilities.** RESP charges are assigned on the basis of the unpolarized ESP and then held constant; then a single set of atom-centered point polarizabilities, modeled as not interacting with each other, is adjusted for each molecule, so as to allow optimal replication of the full set of polarized QM ESPs. Chemically equivalent atoms are constrained to have equal polarizabilities.

**Model 4: Co-optimized charges and direct polarizabilities.** This is the same as model 3, except that a single, fixed set of atomic charges is optimized, along with the point polarizabilities, to best replicate the full set of polarized ESPs. Thus, the final set of point charges differs from the RESP charges as it is chosen to best model not only the baseline unpolarized potential but

also all induced potentials, in conjunction with the inducible dipoles. Chemically equivalent atoms are constrained to have equal polarizabilities and charges.

**Model 5: RESP charges and self-consistent polarizabilities.** This is the same as model 3, except that the induced dipoles interact with each other and their induced dipole moments are solved self-consistently.

**Model 6: Co-optimized charges and self-consistent polarizabilities.** This is the same as model 5, except that a single, optimal set of point charges is obtained along with the optimal point polarizabilities.

For completeness, we also examined how much the polarized ESPs deviated from the ESPs computed with standard RESP charges optimized to the baseline, unpolarized potentials. On the other hand, we did not examine the Drude model because it is essentially an implementation of the self-consistent, atom-centered, point-polarizability model, which is analyzed here (models 4 and 6). In addition, due to its dynamical nature, the Drude model does not appear to provide an unambiguous mapping between a molecular conformation and a set of induced dipoles.

**Molecular Structures and Atom Types.** To examine polarization in the context of biologically relevant molecules, we studied nonpolar, neutral polar, and ionized amino acid side-chain analogues (Figure 1). The nitrobenzene molecule was also included to test the ability of point dipoles to model any charge redistribution between the two nitro oxygens when an inducing charge is near one of them and also to probe how well the two polarization representations handle out-of-plane induction. In addition, the gas-phase dimers of the valine analogue and of methane were studied. These are informative because they increase the number of dipole−dipole interactions among point polarizabilities relative to the small number of such interactions present in the monomers. Therefore, they are useful to further probe the accuracy of the direct model versus that of the self-consistent polarizability model. Multiple conformations were included for the aspartate, tyrosine and valine analogues. The coordinates of all molecular systems are energy-minimized structures drawn from the Benchmark Energy and Geometry Database.[35]

In point-polarizability models, whose parameters are optimized across all polarized ESPs, chemically equivalent atoms are assigned identical polarizabilities and atomic charges. In particular, using the fact that the RESP software available through the Antechamber program[36] automatically forces chemically equivalent atoms to have identical RESP charges, we assigned equivalent parameters to groups of atoms with identical RESP charges. To do this, we assigned a type to each group of chemically equivalent atoms in a molecule; for example, the valine side-chain analogue has two carbon types and two hydrogen types (Figure 1g). Note that these types do not necessarily correspond to force-field atom types and that no equivalence was enforced between molecules. For example, c31 can have different parameters in each molecule where it occurs.

**Calculations of Reference QM ESPs.** For each molecule or dimer, gas-phase QM ESPs were computed at the HF/6-31G* level with Gaussian 09 (RevD.01).[37] In addition to baseline ESPs, which represent unpolarized states, polarized ESPs were generated by solving the wave equation with an inducing point charge, ±1.0$e$, at locations around the molecule determined with the dot molecular surface program dms included with MIDAS.[38] To position the inducing charges roughly one heavy-atom diameter away from the atom-centers, the surfaces were generated with all default atom radii incremented by 1.9 Å. The default probe radius of 1.4 Å was used, and the point density was set to 0.1 points per Å$^2$, leading to roughly 150 points per molecule. The resulting baseline and polarized ESPs were treated as reference data to optimize the various polarization models, as detailed below.

**Description of Systems and Characterization of Errors.** We consider a molecule with $N$ atoms so that for $i = 1,...,N$, $q_i$ are atom-centered point charges, $\boldsymbol{\mu}_i$ are atom-centered point dipoles, and $\alpha_i$ are atom-centered polarizabilities, where the atom centers are at locations $\boldsymbol{r}_i$. The molecule has $T \leq N$ atom types, so that for $t_i \in [1,...T]$, $q_{t_i}$ are atom-typed point charges and $\alpha_{t_i}$ are atom-typed polarizabilities. The ESP values are computed at $M$ locations $\boldsymbol{r}_m$, $m = 1,...,M$, where $M$ is determined by Gaussian 09 and varies from one molecule to another. The unpolarized reference quantum ESP values are obtained in the absence of an inducing charge, whereas the polarized reference quantum ESP values are computed in the presence of an inducing charge $q_k = \pm 1$ at a dms-assigned location $\boldsymbol{r}_k$. The inducing charge locations are repeated for both

positive and negative point charges, so that the reference set consists of a total of 2$K$ polarized QM ESPs. At an ESP point $\boldsymbol{r}_m$, the unpolarized reference potential is $\phi_{m0}^0$, whereas the QM potential computed in the presence of an inducing charge $q_k = 1$ is $\phi_{mk}^0$ for $k = 1,...,K$, and the potential computed in the presence of an inducing charge $q_k = -1$ is $\phi_{mk}^0$ for $k = K + 1,...,2K$. The corresponding potentials from a given polarization model are $\phi_{mk}$, for $k = 0,...,2K$. Note that all ESP values considered here omit the direct Coulombic contribution from the inducing charges.

Optimization of a polarization model means minimizing errors computed in terms of sums of squared potential differences $(\phi_{mk} - \phi_{mk}^0)^2$. We report the overall error of any model when it was applied to a given molecule in the presence of inducing charges at positions corresponding to $k = 1, ..., K$ as the root-mean-square error (RMSE) of the potential (kcal/mol-$e$) across all ESP points for both the positive ($k \in [1, K]$) and negative ($k \in [K + 1, 2K]$) inducing charges located at $\boldsymbol{r}_k$

$$R_k = \left( \frac{\chi_k^2 + \chi_{k+K}^2}{2M} \right)^{1/2} \tag{1}$$

$$\chi_k^2 = \sum_{m=1}^{M} (\phi_{mk} - \phi_{mk}^0)^2 \tag{2}$$

The error across all inducing charges for a given molecule can then be written as

$$R = \left( \frac{1}{K} \sum_{k=1}^{K} R_k^2 \right)^{1/2} \tag{3}$$

**Implementation of Models.** The following sections detail the calculation of RESP charges for the baseline, unpolarized potential and the implementation of each polarization model listed above.

*Baseline RESP Charges.* Atom-centered point charges were fitted to best replicate the baseline, unpolarized QM ESP of each molecule, $\phi_{0m}^0$, using RESP as implemented in the Antechamber program.[36] The two-stage fitting process uses the default assisted model building with energy refinement (AMBER) force-field charge restraint weights of 0.0005 for the first stage and 0.001 for the second.

*Model 0: Global Optimal Point Charges.* For each molecule, atom-centered point charges were optimized for the best simultaneous fit to all polarized ESPs. The ESP at site $\boldsymbol{r}_m$ due to the full set of atomic charges, $q_i$, $i = 1, ..., N$, is calculated by

$$\phi_{mk} = \sum_{i}^{N} \frac{q_i}{|\boldsymbol{r}_{im}|} \tag{4}$$

where well-known physical constants are omitted for simplicity and

$$\boldsymbol{r}_{ij} \equiv \boldsymbol{r}_i - \boldsymbol{r}_j \tag{5}$$

The indices $i$ and $j$ are used for atomic centers, whereas $m$ refers to the ESP locations. Note that the same set of atomic charges is used for every inducing charge $k$, so that the potential $\phi_{mk}$ is actually independent of $k$. The optimization procedure, which minimizes the quantity $R$ in eq 3, is described in the Optimization of Parameters section.

*Model 1: Optimal Point Charges.* Atom-centered point charges were fitted separately to each polarized QM ESP using the RESP software and the same set of ESP points as for the baseline, unpolarized, potential but with the restraint weights set to 0 and with atom equivalency disabled to allow free optimization of point charges. Thus, this model reports on the maximal accuracy attainable by any polarization model that uses the atom-centered point-charge representation of polarization. However, it cannot be employed in simulations because it requires a new quantum calculation for each molecular configuration.

*Model 2: RESP Charges and Optimal Point Dipoles.* Atom-centered point charges were first fitted to the baseline, unpolarized ESP using the standard RESP, then atom-centered point dipoles, superimposed on the point charges, were optimized separately for each polarized QM ESP. This model reports on the maximal accuracy attainable by any polarization model that uses the atom-centered point-dipole representation of polarization along with RESP baseline point charges. However, like the optimal-point-charge model above, it cannot be employed in simulations because it requires a new quantum calculation for each molecular configuration. The procedure for computing optimal point dipoles is described below.

The ESP at $r_m$ due to the baseline RESP partial charges, $q_i$, and the atom-centered point dipoles, $\mu_i$, is given by

$$\phi_m = \sum_i \frac{q_i}{|r_{im}|} + \sum_i \mu_i A_{im} \tag{6}$$

where well-known physical constants are omitted for simplicity and $r_{ij}$ is as given in eq 5 and

$$A_{ij} \equiv \frac{r_{ij}}{|r_{ij}^3|} \tag{7}$$

For each external charge position $r_k$, where index $k$ refers to the inducing charge, the error metric $\chi_k^2$ (eq 2) may be written as

$$\chi_k^2 = \sum_m \left( \phi_{mk}' - \sum_i \mu_i A_{im} \right)^2$$
$$\phi_{mk}' = \phi_{mk}^0 - \sum_i \frac{q_i}{|r_{im}|} \tag{8}$$

where $\phi_{mk}'$ contains all quantities independent of the dipoles.

Optimal dipoles are obtained by setting $\frac{\partial(\chi_k^2)}{\partial\mu_i} = 0$ for all $\mu_i$. This yields the following system of linear equations

$$\begin{bmatrix} \sum_m A_{1m}\phi_m' \\ \vdots \\ \sum_m A_{Nm}\phi_m' \end{bmatrix}_k = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_N \end{bmatrix}_k \begin{bmatrix} \sum_m A_{1m}^2 & \cdots & \sum_m A_{1m}A_{Nm} \\ \vdots & \ddots & \vdots \\ \sum_m A_{Nm}A_{1m} & \cdots & \sum_m A_{Nm}^2 \end{bmatrix} \tag{9}$$

Solving this matrix equation yields the desired atom-centered point dipoles optimized for the external charge position $r_k$. The NumPy *linalg.norm*[39] function was used to find the solution to the matrix equation.

*Model 3: RESP Charges and Direct Polarizabilities.* First, atom-centered point charges were fitted to the baseline, unpolarized ESP, using the standard RESP. Then, a single set of atom-typed, atom-centered, isotropic polarizabilities was

optimized for best simultaneous fit to all polarized ESPs ($k = 1$, ..., $2K$), where the model ESPs are computed as sums of the potentials from the baseline RESP charges and those from the induced dipoles. In this first-order, or direct, polarizability model, the field at each atom $i$, $E_i$, is that due to the inducing charge, $E_i^{q_k}$, plus that due to the fixed RESP charges, $E_i^q$; fields due to induced dipoles are ignored. Thus

$$E_i = E_i^{q_k} + E_i^q = q_k A_{ik} + \sum_{j \neq i}^N S_{ij} q_j A_{ij} \tag{10}$$

Here, $S_{ij}$ is a screening function which excludes 1−2 and 1−3 interactions and scales 1−4 interactions by a factor of 0.5, in keeping with common practice in computing Coulombic interactions with empirical force fields. The induced dipoles are proportional to the inducing field $E_i$

$$\mu_i = \alpha_{t_i} E_i = \alpha_{t_i} [E_i^{q_k} + E_i^q] \tag{11}$$

Atom-typed point polarizabilities, $\alpha_{t_i}$, are parameters which must be adjusted to minimize the mean-squared error, $R^2$, across all polarization states $k = 1$, ..., $2K$ (see eq 3). The optimization procedure is described in the Optimization of Parameters section. Note that the set of charges and point polarizabilities derived in this way is applicable to all inducing charges and charge positions as the inducible dipoles respond to the inducing field according to eq 11. Thus, no additional QM calculations are needed, and this model could be used in a simulation.

*Model 4: Co-optimized Charges and Direct Polarizabilities.* This model is identical to model 3, except that a single set of atom-centered, atom-typed, point charges is co-optimized with the point polarizabilities across all polarization states, to minimize the global mean-squared error $R^2$. Again, this yields a model which can be used in a simulation because no additional QM calculations are needed.

*Model 5: RESP Charges and Self-Consistent Polarizabilities.* This model is the same as model 3, except that the field at each atom $i$ now includes a contribution from other induced dipoles, $E_i^\mu$, and the system of induced dipoles is solved self-consistently. For a given inducing charge $k$, the total electric field at atom $i$ becomes

$$E_i = E_i^{q_k} + E_i^q + E_i^\mu = q_k A_{ik} + \sum_{j \neq i}^N S_{ij} q_j A_{ij} + \sum_{j \neq i}^N S_{ij} \mu_j B_{ij} \tag{12}$$

(Because the RESP charges are atom-typed, one could properly replace $q_j$ by $q_{t_j}$.) We use the Applequist dipole interaction model,[40] for which $B_{ij}$ depends only on separation $r_{ij}$

$$B_{ij}^{\beta\gamma} = \frac{3r_i^\beta r_j^\gamma}{r_{ij}^5} - \frac{\delta_{\beta\gamma}}{r_{ij}^3} \tag{13}$$

where $\beta, \gamma \in \{x, y, z\}$.

The induced dipoles are, again, computed from these fields and the atom-typed point polarizabilities

$$\mu_i = \alpha_i E_i = \alpha_{t_i} [E_i^{q_k} + E_i^q + E_i^\mu] \tag{14}$$
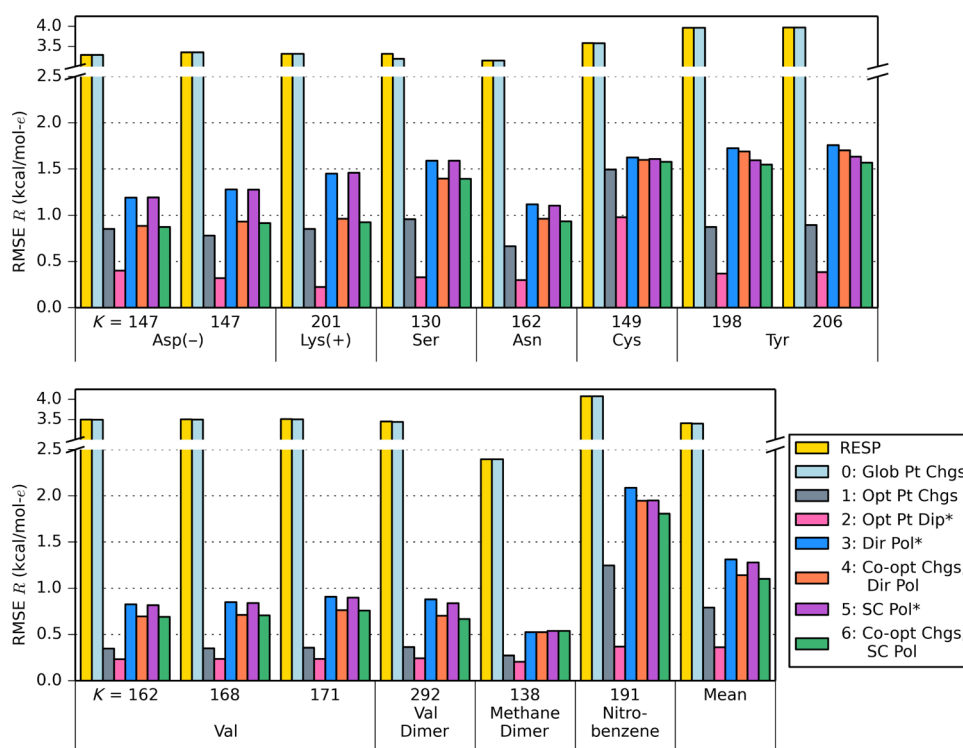
These equations can be rewritten in a matrix form as

**Figure 2.** RMSE ($R$) of polarization models. The asterisk in the legend indicates that RESP charges were used.

$$
\begin{bmatrix} E_1^{q_k} + E_1^{q} \\ \vdots \\ \vdots \\ E_N^{q_k} + E_N^{q} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \vdots \\ \mu_N \end{bmatrix} \begin{bmatrix} \alpha'_{t_1} & -B_{12} & \cdots & -B_{1N} \\ -B_{12} & \alpha'_{t_2} & \cdots & -B_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ -B_{N1} & -B_{N2} & \cdots & \alpha'_{t_N} \end{bmatrix}
\tag{15}
$$

where $\alpha'_{t_i} = \alpha_{t_i}^{-1} I$ and $B_{ij}$ are $3 \times 3$ matrices. Solving this matrix equation yields a set of self-consistent induced dipoles, $\mu_i$, particular to a set of point charges, $q_i$, polarizabilities $\alpha_{t_i}$, and inducing charge $q_k$ at $r_k$.

Using baseline, atom-typed RESP charges for $q_t$, the atom-typed, isotropic polarizabilities $\alpha_{t_i}$ are solved by the same optimization process as used in the direct polarization models (see the Optimization of Parameters section). The resulting polarization model is particular to the set of baseline RESP charges but applicable to all external charge locations $r_k$. Again, this is a model which could be used in a simulation because no further QM calculations are required once the polarizabilities are established.

The treatment of close-ranged dipole−dipole interactions adopted here deserves comment. As noted above, we exclude 1−2 and 1−3 interactions, and scale 1−4 interactions by 0.5. This approach is consistent with many existing polarizable force fields, especially those also utilizing the Applequist model, such as AMBER ff02.[41,42] Some dipole−dipole screening models, such as that of Thole[43] and its variations,[44] also modify the $B_{ij}$ term (eq 13) to prevent the polarization catastrophe to which the Applequist model is susceptible. However, excluding 1−2 and 1−3 short-range intramolecular interactions and having 1−4 interactions, as done here, prevents this problem, even with the Applequist model, and our test calculations demonstrated that adding Thole or Thole-like screening terms to the present models produced negligible changes in the results (data not

shown). Indeed, Cieplak et al.[11] note that other screening models deviate significantly from the Applequist model only at ranges shorter than about 3.0 Å but the 1−2 and 1−3 exclusions already eliminate many interactions within this range, so the screening models become virtually equivalent. This observation is corroborated by the similarity of fitting results and parameters in Wang et al.[45] for the CL, CE, CT, and CA models, where C refers to the same short-range scaling as detailed above and the second letter refers to the added screening functions tested.

*Model 6: Co-optimized Charges and Self-Consistent Polarizabilities.* This model is the same as model 5, except that a fixed set of atom-typed partial charges is co-optimized with the polarizabilities against the full set of polarized ESPs. Thus, the atom-typed parameters $q_{t_i}$ and $\alpha_{t_i}$ are simultaneously adjusted, using the same optimization process and error function as used for the other inducible dipole models. Again, this model can be used in a simulation as no additional QM calculations are needed.

**Optimization of Parameters.** All models require global parameter optimization. For model 0 and models 3−6, the parameters were optimized to minimize $R^2$, the mean of the squared potential deviations across all inducing charge sites $k$, for each molecule of interest (eq 3). For models 1 and 2, the parameters were optimized to minimize $R_k^2$, the mean-squared potential deviations for each separate inducing charge position (eq 1). All optimizations were performed with a SciPy implementation of L-BFGS-B,[46,47] a gradient-based constrained minimization method. Charges are left unconstrained, whereas polarizabilities are restricted to positive values. For the inducible dipole models with fixed RESP charges (models 3 and 5), only the atom-typed polarizabilities, $\alpha_{t_i}$, require adjustment; for those with co-optimized point charges, the

atom-typed charges, $q_t$, are adjusted along with the polarizabilities.

For each molecule, multiple optimizations were run with initial parameter values drawn from a uniform distribution using *numpy.random.rand()*.[39] For model 0, five optimizations were run using initial charges randomly drawn from the range $-1.0e$ to $1.0e$. For models 3–6, 50 optimizations were run using initial polarizabilities randomly drawn from 0 to 10 bohr³ (0– 1.482 Å³). The parameter set with the lowest value of $R^2$ was selected as the optimum. When charges were co-optimized (models 4 and 6), the baseline RESP charges were used as their starting values. Only 10 optimizations were run for tyrosine as the calculations became time-consuming for this relatively large molecule.

**Calculation of Isotropic Molecular Polarizabilities.** To check the plausibility of the optimized atomic polarizabilities for the inducible dipole models, we computed the corresponding molecular polarizabilities and compared them with the molecular polarizabilities for the same compounds computed with the Gaussian 09 software. The isotropic molecular polarizability of each molecule was calculated from the optimized isotropic atomic polarizabilities as

$$\alpha_{\mathrm{mol}} = \sum_i \alpha_{t_i} \tag{16}$$

The molecular polarizabilities from Gaussian 09 were calculated at the HF/6-31G* level, for consistency with the ESP calculations used to derive the atomic polarizabilities.

## ■ RESULTS AND DISCUSSION

This section reports on the accuracy of the various representations and models of polarization and then appraises the charges and polarizabilities obtained by optimizing these parameters for the four inducible dipole models.

**Accuracy of Polarization Models.** *Overall Accuracy.* The errors of the models are reported as $R$ (eq 3), which is the RMSE of the ESPs approximated by each model for all inducing charges. Note that these errors will vary with the magnitudes of the inducing fields, with larger fields generating larger errors. Here, the inducing fields were generated by unit monopoles placed roughly one atomic diameter away from the nearest atom-center of the molecule, in vacuo. Thus, roughly similar errors might be anticipated in simulations of systems with univalent ions near organic compounds in a low-dielectric environment, like a lipid membrane; the overall error in such a system will then accumulate and/or cancel across a multitude of complex interatomic interactions.

To establish a baseline for comparing the various models, it is of interest to summarize the errors incurred by *not* using any polarizable model at all. This baseline $R$ measures the RMSE of the approximate ESPs computed using standard RESP point charges, which were fit to the unpolarized QM ESP, for all of the inducing charges. As shown in Figure 2 (yellow columns), these errors range from about 3 to 4 kcal/mol-$e$. No significant reduction in error is obtained by using a single set of point charges simultaneously optimized to the ESPs associated with all inducing charges (model 0), as evident in Figure 2. In fact, the values of $R$ from model 0 are the same as those from plain RESP to within 0.01 kcal/mol-$e$. The validity of this result was confirmed by running five different optimizations of the model 0 charges from different randomized starting values; the standard deviation of the optimized atomic charges for each

atom across the five runs was within 0.0026$e$. Interestingly, the model 0 charges strongly resemble the RESP charges: a linear regression of model 0 versus RESP charges across all molecules gives a slope of 0.99, $y$-intercept of 0.0017 and Pearson correlation coefficient of 0.997. Due to the similarity of the model 0 and RESP charges and errors, subsequent references to the results when polarization is neglected may be considered to reference either the baseline RESP results or model 0.

As detailed in the Methods section, we tested two representations of polarization tuned to fit each individual polarized QM ESP: adjustable atom-centered point charges (model 1) and adjustable atom-centered point dipoles (model 2). Both representations lead to errors in the induced ESPs well below the 3–4 kcal/mol-$e$ errors obtained when polarization is neglected (prior paragraph). Optimal point charges reduce the error, on average, to about 0.75 kcal/mol-$e$ (gray columns, Figure 2), whereas optimal point dipoles lead to even lower average errors of about 0.4 kcal/mol-$e$ (magenta columns, Figure 2). The advantage of point dipoles over point charges is consistent across all molecules studied, though the difference varies from case to case and is lowest for methane and for the simple alkane model of valine. It should be emphasized, however, that these results bear only on the ability of adjustable point charges and point dipoles to capture induced changes in molecular ESPs.

We investigated whether model 2 provides greater accuracy than model 1 because it offers a better description of the baseline, unpolarized electrostatic field. We used nitrobenzene as a test case to check for this possibility, creating a model with RESP charges supplemented by point dipoles optimized to replicate the baseline, unpolarized ESP of this molecule. These baseline optimized dipoles, $\boldsymbol{\mu}_i^0$, may be viewed as optimized permanent dipoles for the unpolarized molecule, and they help correct for the errors of RESP charges in replicating the unpolarized QM ESP. Then, we computed the RMSE of the ESP generated by this permanent-charge plus permanent-dipole model against the full set of induced ESPs for all 382 inducing charges ($\pm 1.0e$ at each of 191 locations). The resulting RMSE of 4.06 kcal/mol-$e$ is essentially the same as that for baseline RESP charges alone, 4.08 kcal/mol-$e$. Thus, adding permanent dipoles to the baseline RESP charges produces minimal improvement in the ESP fits across the set of inducing charges. In contrast, the RMSE is 0.37 kcal/mol-$e$ for model 2, in which a new set of point dipoles is optimized for each inducing charge. Thus, the lower errors observed for model 2 versus those for baseline RESP charges are attributable to the improved description of induced polarization rather than to that of the baseline potential.

The four polarizability models based on inducible, atom-centered point dipoles (models 3–6) were then examined. In each case, a single set of polarizabilities and, where applicable, point charges, was optimized to best replicate the full set of QM polarized ESPs for each molecule. Figure 2 summarizes the accuracy of the resulting parameterized models.

Both models 3 and 4 use the efficient direct polarization approximation, but model 3 keeps the baseline RESP point charges (first paragraph of the Results section), whereas model 4 uses a new set of point charges optimized along with the polarizabilities. Both of these models yield errors ($R$) averaging about 1.25 kcal/mol-$e$ (blue and orange columns, respectively; Figure 2). This is substantially better than the baseline errors of 3–4 kcal/mol-$e$ associated with unpolarized RESP charges (above) but about 3-fold worse than the theoretical optimum of
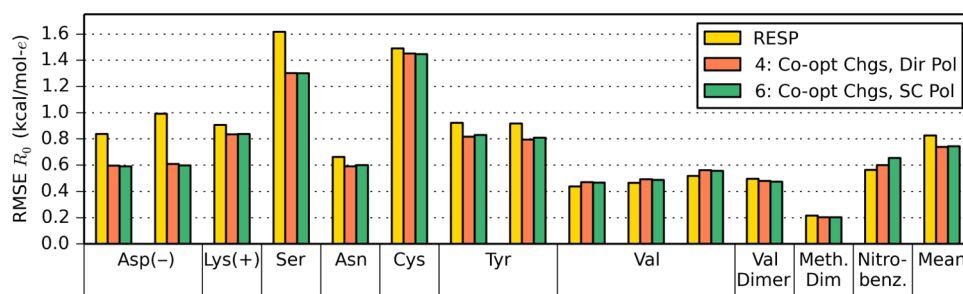
**Figure 3.** RMSE ($R_0$) of polarization models for unpolarized states.

about 0.4 kcal/mol-$e$ for the point-dipole representation (model 2). It is also nearly 2-fold worse than the theoretical optimum of about 0.75 kcal/mol-$e$ achievable with the variable-point-charge model (model 1).

One possible source of error in models 3 and 4 is their nonphysical neglect of interactions among the induced dipoles. However, models 5 and 6, which are the same, respectively, except that they include these interactions and thus require solving a matrix equation, yield at best marginal improvements in accuracy across all cases tested (purple and green columns, Figure 2); the differences in $R$ are all <0.15 kcal/mol-$e$. A possible concern regarding this observation is that the exclusion of 1−2 and 1−3 interactions and the scaling of 1−4 interactions eliminate or weaken many dipole−dipole interactions that would otherwise have been fully included in the self-consistent calculations; thus, the direct model may not be very different from the self-consistent model. We addressed this concern by including two dimers (valine analogue and methane), reasoning that no intermolecular dipole−dipole interactions are excluded; thus, these cases should better probe the differences between the direct and self-consistent approaches. Nonetheless, the self-consistent models do not particularly outperform the direct models for the dimers either (Figure 2); in fact, they are marginally worse than the direct models for the methane dimer.

Note that the highest-level polarizability model tested here, model 6, which uses self-consistent induced dipoles and co-optimized point charges, does not reach the maximal accuracy achievable for the point-charge representation of polarization (model 1). This means that a sufficiently well designed point-charge model of polarizability could outperform model 6, which is the standard self-consistent induced-dipole model. Given the failure of model 6 to approach the accuracy of the theoretical optimal-point-dipole model (model 2), there is even more room to improve the accuracy of polarization models based on the point-dipole representation.

We examined whether the greater accuracy of models 1 and 2 relative to the linear response models, models 3−6, perhaps resulted from their ability to account for nonlinearity in the QM polarization response. We again used nitrobenzene as a test case, rerunning quantum calculations with charges of ±0.2, ±0.4, ±0.6, and ±0.8 $e$ at each of the 191 inducing charge locations and testing for linearity by several criteria. First, for each inducing charge location, we ran a linear regression of the molecular dipole moment provided by Gaussian against the value of the inducing charge: the Pearson correlation coefficients were found to be at least 0.9999 for all locations of the inducing charge, confirming the linearity of the overall polarization response. Second, for each inducing charge location, we carried out linear regressions of the magnitudes of the optimized (model 2) point dipoles at each atom against

the magnitude of the inducing charges. (As nitrobenzene has 14 atoms, there are 382 × 14 regression fits.) We averaged the Pearson correlation coefficients for each atom across charge locations to provide summary statistics, and the lowest mean correlation coefficient is found to be 0.9999, again indicating a linear response. The corresponding analysis of the optimal-point-charge model, model 1, similarly yielded correlation coefficients ≥0.9990. Finally, we tested how well the accuracy of a linear extrapolation of the atom-centered point dipoles fitted to inducing charges of ±0.2$e$ could replicate quantum ESPs induced with a charge of ±1.0$e$. For each inducing charge location (not indexed, for simplicity) and atom ($i$), the dipole moment for a unit inducing charge is extrapolated as follows: $\mu_i^{1.0} \approx 5(\mu_i^{0.2} - \mu_i^0) + \mu_i^0$, where $\mu$ indicates a dipole moment and the superscripts indicate the value of the inducing charges to which they pertain. (The expression for a negative inducing charge is analogous.) The ESPs from these extrapolated dipoles were compared with the quantum ESPs for the corresponding inducing charge locations but obtained with inducing unit charges. The RMSE across all inducing charge positions is 0.372 kcal/mol-$e$, which is essentially the same as the result of model 2 obtained by optimizing dipoles for inducing charges of unit magnitude, 0.369 kcal/mol-$e$. The analogous extrapolation from inducing charges of 0.2 for the optimal-point-charge model, model 1, yields an RMSE of 1.247 kcal/mol-$e$, which is essentially the same as that for model 1 charges optimized directly against ESPs for inducing charges of unit magnitude, 1.246 kcal/mol-$e$. Thus, the greater accuracy of models 1 and 2 relative to the linear polarization models, models 3−6, is not related to their ability to capture a nonlinear polarization response.

Another dimension of the inducible dipole models studied here is their use of either baseline unpolarized RESP charges or a new set of point charges optimized along with the polarizabilities to best replicate the QM polarized ESPs. The use of co-optimized charges yields somewhat greater accuracy in all cases; compare model 4 with model 3 and model 6 with model 5 (Figure 2). The improvements are greatest (0.3−0.5 kcal/mol-$e$) for the ionized systems and least for the cysteine analogue (∼0.03 kcal/mol-$e$). Importantly, when models 4 and 6, with co-optimized charges and either direct or self-consistent polarizabilities, are used to compute the baseline, unpolarized ESP, the agreement is somewhat better, on average, than that provided by standard (unpolarized) RESP charges (Figure 3). Thus, optimization of charges and polarizabilities against sets of polarized ESPs results in parameters that are also applicable to the unpolarized state.

*Accuracy at Each Inducing Charge Location.* Further insight into the strengths and weaknesses of the various polarization models can be obtained by visualizing the errors

(a) Baseline RESP



(b) Optimal point charges



(c) Optimal point dipoles



(d) Co-optimized charges and self-consistent polarizabilities
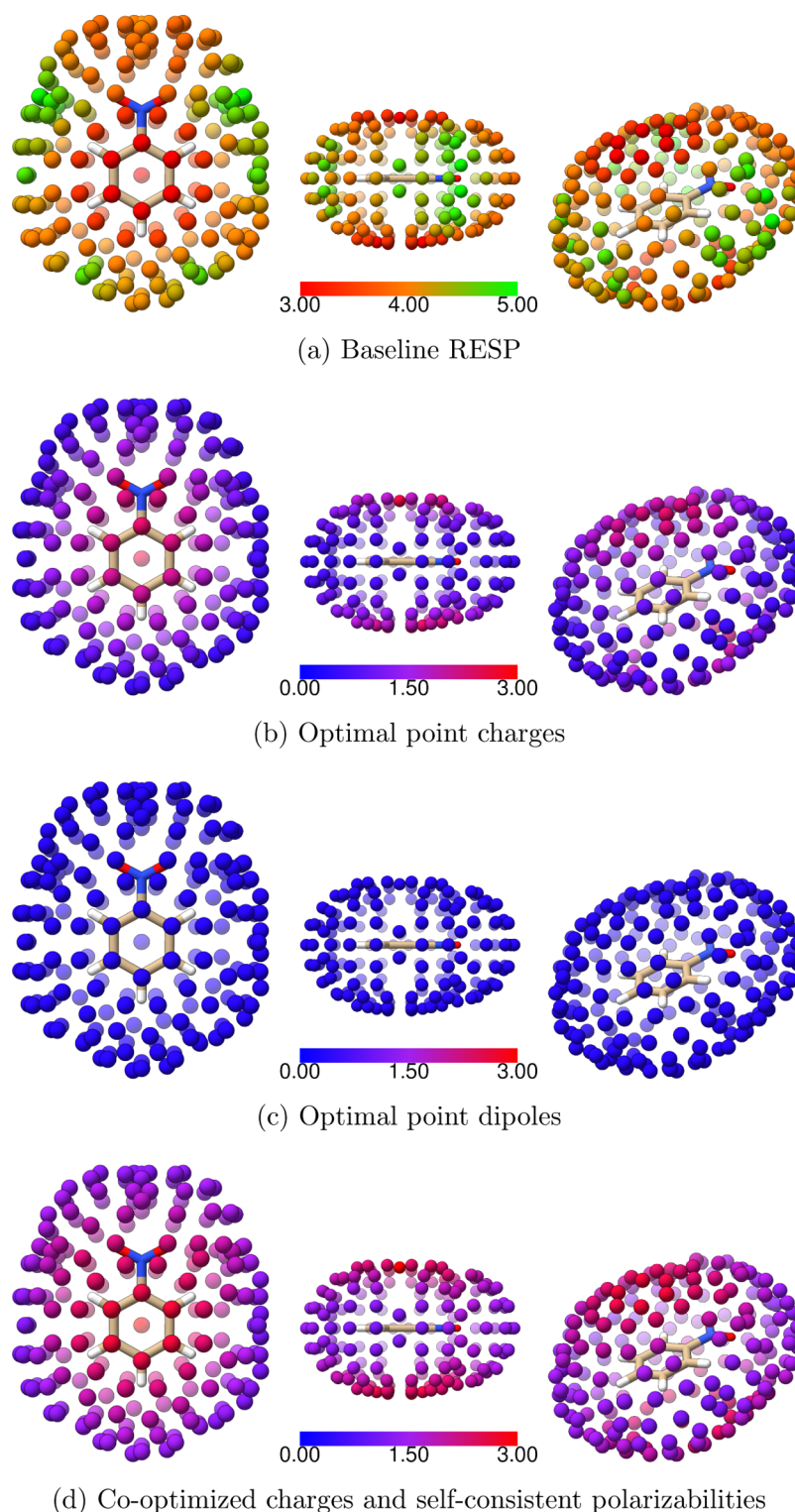
**Figure 4.** RMSE ($R_k$, kcal/mol-$e$) for nitrobenzene as a function of inducing charge location $r_k$.

associated with different positions, $r_k$, of the inducing charge. Figures 4 and 5 depict nitrobenzene and the valine analogue, respectively, each surrounded by colored spheres at representative positions, $r_k$, of the inducing charge. Each sphere is colored according to the overall error $R_k$ (eq 1) associated with the inducing charge at the location of the sphere. Results are presented for the baseline RESP charges, optimal point charges, optimal point dipoles, co-optimized charges and self-consistent

polarization. We used different color scales for different panels to bring out the geometric variations associated with each model and molecule.

Not surprisingly (see Figure 2), baseline RESP charges do a poor job of replicating the induced potentials. For nitrobenzene, the errors are less for out-of-plane than for in-plane inducing charges (Figure 4a), whereas for the valine analogue, the errors are least for inducing charges near the axes of the

(a) Baseline RESP



(b) Optimal point charges



(c) Optimal point dipoles



(d) Co-optimized charges and self-consistent polarizabilities
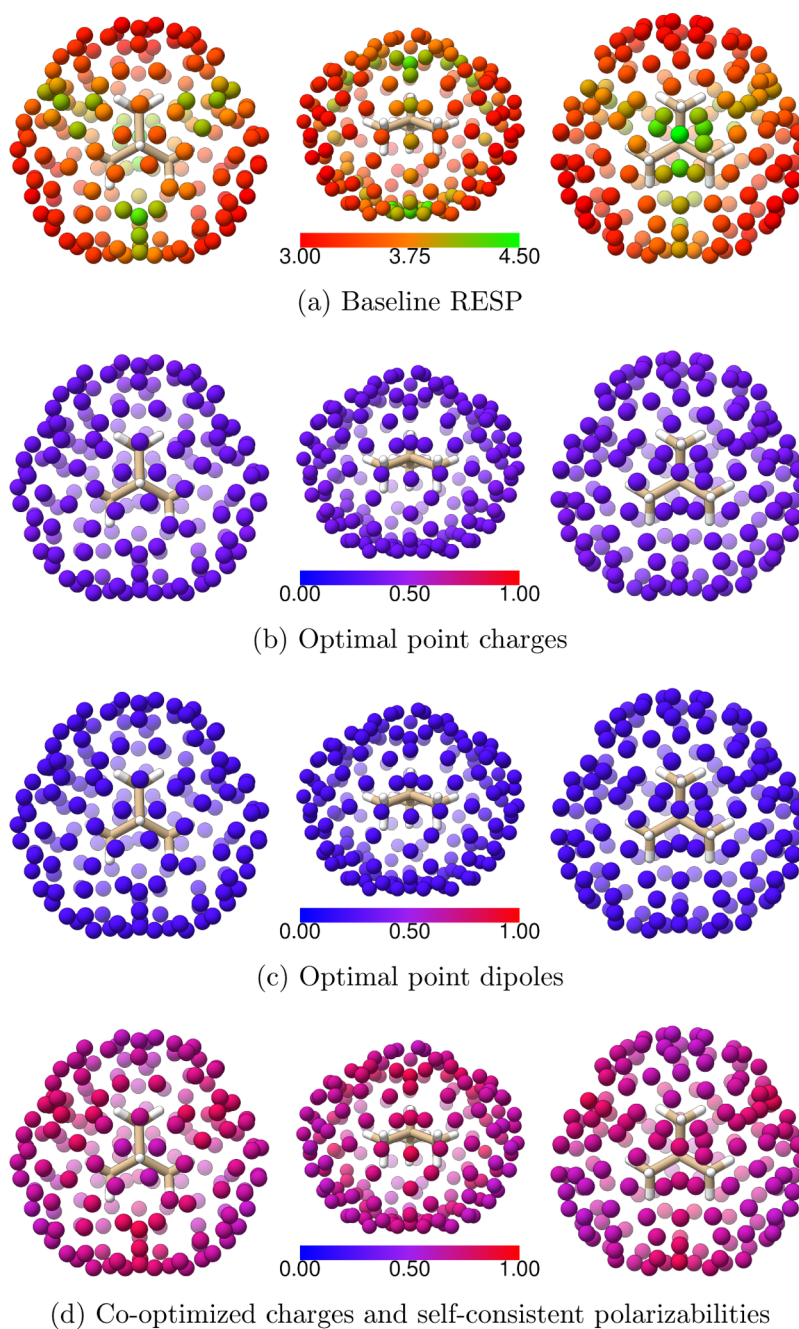
**Figure 5.** RMSE ($R_k$, kcal/mol-$e$) for the valine analogue as a function of inducing charge location $r_k$.

carbon−carbon bonds (Figure 5a). All of the polarization models yield lower errors than the baseline RESP at all external locations, so accounting for polarization is consistently advantageous. The relative ranking of the methods is also consistent: optimal point dipoles (model 2) yield the lowest errors, followed by optimal point charges (model 1) and then co-optimized charges with self-consistent polarizabilities (model 6). In the case of valine, which is nonplanar, the errors for all three models vary little with the position of the inducing charge as the values of $R_k$ remain within a 0.7 kcal/mol-$e$ range for each model. However, the pattern is more complex for the planar nitrobenzene molecule. Here, the errors are greater for out-of-plane inducing charges than for in-plane ones for both models 1 and 6 (Figure 4b,d). (A similar pattern is seen for the aromatic ring of the tyrosine analogue; data not shown.) It is

necessarily the case when the point-charge representation of polarization in model 1 is unable to account for out-of-plane polarization for a planar molecule. It is harder to explain why model 6 performs worse for out-of-plane than for in-plane charges and, indeed, worse for out-of-plane charges than model 1; compare Figure 4b,d.

Optimal point dipoles (model 2) can yield out-of-plane induced moments, and they afford good accuracy for inducing charges at in-plane and out-of-plane positions around nitrobenzene, as well as for all positions around the valine analogue. Indeed, the errors for optimal point dipoles are uniform, to within 0.0035 kcal/mol-$e$, for inducing charges around both molecules. This result supports a view that the point dipole representation can provide consistent performance in modeling

**Table 1. Isotropic Molecular Polarizabilities ($Å^3$) from eq 16 Using Optimized Parameters, QM Calculations, and Experiment[a]**

|  | 3: RESP, Dir Pol | 4: Co-opt Chgs, Dir Pol | 5: RESP SC Pol, | 6: Co-opt Chgs, SC Pol | HF/6-31G* | expt[51] |
|---|---|---|---|---|---|---|
| Asp(−) | 4.806 | 5.050 | 4.813 | 5.082 | 5.294 |  |
| Asp(−) | 4.820 | 5.048 | 4.833 | 5.087 | 5.306 |  |
| Lys(+) | 7.258 | 8.004 | 7.313 | 8.249 | 8.589 |  |
| Ser | 3.666 | 3.641 | 3.674 | 3.650 | 3.763 | 5.41 |
| Asn | 5.087 | 5.256 | 5.128 | 5.320 | 5.625 |  |
| Cys | 5.073 | 5.113 | 5.128 | 5.178 | 5.414 | 7.41 |
| Tyr | 10.312 | 10.443 | 10.812 | 10.931 | 11.734 |  |
| Tyr | 10.199 | 10.339 | 10.725 | 10.893 | 11.826 |  |
| Val | 6.064 | 6.121 | 6.108 | 6.166 | 6.302 | 8.14 |
| Val | 6.065 | 6.123 | 6.110 | 6.169 | 6.315 | 8.14 |
| Val | 6.063 | 6.129 | 6.105 | 6.169 | 6.314 | 8.14 |
| Val dimer | 11.966 | 12.034 | 12.241 | 12.285 | 12.829 | 16.28 |
| methane dimer | 3.571 | 3.571 | 3.609 | 3.609 | 3.618 | 5.186 |
| nitrobenzene | 8.218 | 8.579 | 8.666 | 9.043 | 9.808 |  |

[a]The experimental values listed for the dimers are merely twice the experimental monomer results.

all polarization phenomena. However, optimal performance is clearly not attained by the customary inducible dipole models.

**Optimized Polarizabilities and Charges.** A basic check of the plausibility of the optimized polarizabilities is that these should be consistent with independently determined molecular polarizabilities. Isotropic molecular polarizabilities were computed from atomic polarizabilities using eq 16, and Table 1 compares these values with molecular polarizabilities computed independently with Gaussian 09 QM calculations at the same HF/6-31G* level used to compute the molecular ESPs, as well as with available experimental molecular polarizabilities. The molecular polarizabilities from the inducible dipole models agree well with the independently computed QM molecular polarizabilities and hence with each other. However, the QM molecular polarizabilities underestimate the experimental results by on the order of 30%. This is consistent with a broader set of data showing that polarizabilities computed at the 6-31G* level underestimate experimental results, whereas more complete basis sets, such as aug-cc-pVTZ,[48,49] yield more accurate results.[50]

All optimized parameters for the four inducible dipole models are listed in Table 2; the unpolarized RESP charges are also shown, as these are used in models 3 and 5 and may be compared with the co-optimized charges of models 4 and 6. The values of the optimized polarizabilities range from 0.000 to 3.643 $Å^3$. The zeroes are, arguably, nonphysical, but we note that in many of these cases there appears to be compensation by neighboring atoms with particularly large polarizabilities, consistent with the fact that the overall molecular polarizabilities are physically reasonable (above). For example, in nitrobenzene, in which the direct and self-consistent polarizabilities, optimized with RESP charges, are 0 for the *no* atom (Figure 1), the polarizabilities of the immediately neighboring ca4 atoms are relatively large, 2.964 or 3.643 $Å^3$, respectively. Similarly, whereas the c31 atoms in the valine analogue are consistently assigned a polarizability of 0.000 $Å^3$, the central c32 atom type has a relatively large polarizability (>2 $Å^3$). Cases like these could be avoided by adding further restraints during optimization, such as by forcing all carbon atoms in the valine analogue to have equal polarizabilities. However, adding restraints would presumably lessen the accuracy of the agreement of the models with the QM polarized ESPs, and the present minimally restrained optimizations have the merit of revealing the best possible performance of each model.

The optimized polarizabilities may also be analyzed by element. As shown in Table 3, carbon consistently emerges as the most polarizable, followed by oxygen and with nitrogen and hydrogen approximately tied for third place. (The solitary sulfur atoms in the models studied are assigned polarizabilities of 1.7−1.9 $Å^3$, well above the mean of about 1.0 for carbon.) This ranking is in rough agreement with that of the polarizabilities assigned to these elements in a prior study, which used only experimental molecular polarizabilities as target data.[45] Perhaps, the largest discrepancy is for nitrogen, whose optimized polarizabilities are significantly lower here. However, when appropriately atom-typed polarizabilities from the prior study are substituted for the optimized ones developed here, they yield polarized ESPs that deviate about twice as much from the reference polarized QM ESPs, regardless of the point-charge model used (data not shown). About 30% of this increased deviation probably traces to the tendency of the present quantum calculations to underestimate the experimental molecular polarizabilities used (see above).

It is also instructive to examine how the optimized polarizabilities can change with molecular conformation. For the simple valine analogue, all three conformations are quite similar to each other, and the three sets of polarizabilities agree to within 0.015 $Å^3$. However, the picture is more complicated for the two conformations apiece of the aspartate and tyrosine analogues (Figure 6). The molecular conformations tested for the aspartate analogue differ in the torsion of the side chain relative to the ethyl terminus. Thus, in one conformation, the shortest oxygen-to-methyl carbon (o−c31) distance is 2.8 Å, whereas in the other conformation, it is 3.2 Å. The optimal polarizability of c31 is larger for the structure with the shorter o−c31 distance, but the polarizabilities of the other atoms are quite similar. The two conformations tested for the tyrosine analogue are the same to within 0.078 Å RMSD. Nonetheless, the optimized polarizabilities for certain atoms differ substantially between the two conformations; notably, the c32 atom type has polarizability 0.000 $Å^3$ in one structure and 1.442 $Å^3$ in the other. However, this change is partly compensated by opposite shifts in the polarizabilities of atom types ca1 and c31. Interestingly, when the optimized parameters are swapped between conformers, the ESP errors, *R*, change by <0.05 kcal/mol-*e*; thus, the optimal solutions found here are degenerate.

As detailed in the previous section, all four inducible dipole models (models 3−6) yield rather similar levels of accuracy

**Table 2. Optimized Parameters for Inducible Dipole Models (Polarizabilities ($\alpha$) in Å$^3$ and Charges ($q$) in $e$)**

| | | RESP | 3: RESP, Dir Pol | | 4: Co-opt Chgs, Dir Pol | | 5: RESP, SC Pol | 6: Co-opt Chgs, SC Pol | |
|---|---|---|---|---|---|---|---|---|---|
| | | $q$ | $\alpha$ | $q$ | $\alpha$ | $\alpha$ | $q$ | $\alpha$ |
| Asp(−) | c | 0.819 | 0.127 | 0.864 | 0.304 | 0.000 | 0.868 | 0.251 |
| | c31 | −0.062 | 0.000 | −0.145 | 0.446 | 0.000 | −0.137 | 0.490 |
| | c32 | 0.018 | 2.884 | −0.267 | 2.058 | 2.938 | −0.269 | 2.042 |
| | hc1 | −0.002 | 0.162 | 0.049 | 0.306 | 0.160 | 0.048 | 0.305 |
| | hc2 | −0.050 | 0.000 | 0.041 | 0.000 | 0.000 | 0.041 | 0.000 |
| | o | −0.835 | 0.655 | −0.841 | 0.661 | 0.698 | −0.844 | 0.693 |
| Asp(−) | c | 0.808 | 0.186 | 0.884 | 0.240 | 0.116 | 0.889 | 0.094 |
| | c31 | −0.103 | 0.652 | 0.004 | 1.330 | 0.655 | −0.004 | 1.356 |
| | c32 | 0.213 | 2.695 | −0.206 | 1.842 | 2.713 | −0.203 | 1.854 |
| | hc1 | −0.008 | 0.000 | 0.001 | 0.081 | 0.000 | 0.004 | 0.085 |
| | hc2 | −0.110 | 0.000 | 0.007 | 0.000 | 0.000 | 0.007 | 0.000 |
| | o | −0.837 | 0.643 | −0.850 | 0.696 | 0.674 | −0.853 | 0.764 |
| Lys(+) | c31 | −0.348 | 0.000 | −0.360 | 0.000 | 0.000 | −0.364 | 0.011 |
| | c32 | 0.242 | 0.000 | 0.297 | 2.155 | 0.000 | 0.306 | 1.929 |
| | c33 | −0.102 | 1.574 | −0.221 | 0.007 | 1.473 | −0.210 | 0.133 |
| | c34 | −0.096 | 0.000 | 0.022 | 0.003 | 0.000 | 0.019 | 0.000 |
| | c3x | 0.283 | 2.920 | 0.441 | 1.990 | 3.007 | 0.432 | 1.972 |
| | hc1 | 0.094 | 0.481 | 0.087 | 0.377 | 0.475 | 0.087 | 0.393 |
| | hc2 | −0.024 | 0.376 | −0.047 | 0.044 | 0.411 | −0.052 | 0.123 |
| | hc3 | 0.035 | 0.000 | 0.061 | 0.502 | 0.000 | 0.056 | 0.524 |
| | hc4 | 0.041 | 0.000 | 0.013 | 0.444 | 0.000 | 0.015 | 0.457 |
| | hn | 0.373 | 0.190 | 0.390 | 0.246 | 0.196 | 0.389 | 0.253 |
| | hx | 0.043 | 0.000 | −0.000 | 0.000 | 0.000 | 0.003 | 0.000 |
| | n4 | −0.568 | 0.000 | −0.663 | 0.004 | 0.000 | −0.654 | 0.057 |
| Ser | c | 0.866 | 1.992 | 0.966 | 0.683 | 2.018 | 0.968 | 0.634 |
| | c31 | −0.135 | 0.943 | −0.204 | 0.678 | 0.550 | −0.199 | 0.597 |
| | c32 | −0.086 | 1.191 | −0.214 | 1.972 | 1.249 | −0.213 | 1.918 |
| | hc1 | 0.045 | 0.106 | 0.065 | 0.162 | 0.221 | 0.064 | 0.192 |
| | hc2 | 0.029 | 0.000 | 0.070 | 0.047 | 0.000 | 0.069 | 0.088 |
| | n | −1.087 | 0.644 | −1.094 | 0.507 | 0.649 | −1.108 | 0.322 |
| Asn | c31 | −0.282 | 0.014 | −0.255 | 0.000 | 0.019 | −0.257 | 0.000 |
| | c32 | 0.526 | 1.888 | 0.468 | 1.929 | 1.897 | 0.466 | 1.931 |
| | h1 | −0.062 | 0.000 | −0.053 | 0.000 | 0.000 | −0.053 | 0.000 |
| | hc1 | 0.061 | 0.473 | 0.059 | 0.452 | 0.474 | 0.059 | 0.455 |
| | hn | 0.443 | 0.000 | 0.451 | 0.190 | 0.000 | 0.459 | 0.281 |
| | ho | 0.419 | 0.279 | 0.376 | 0.192 | 0.312 | 0.376 | 0.229 |
| | o | −0.637 | 0.000 | −0.691 | 0.455 | 0.000 | −0.697 | 0.535 |
| | oh | −0.722 | 0.066 | −0.659 | 0.164 | 0.023 | −0.658 | 0.126 |
| Cys | c31 | −0.125 | 0.000 | −0.161 | 0.000 | 0.000 | −0.178 | 0.000 |
| | c32 | 0.050 | 1.211 | 0.101 | 1.174 | 0.973 | 0.114 | 0.820 |
| | h1 | 0.038 | 0.271 | 0.026 | 0.250 | 0.334 | 0.022 | 0.348 |
| | hc1 | 0.057 | 0.455 | 0.065 | 0.464 | 0.483 | 0.069 | 0.499 |
| | hs | 0.189 | 0.254 | 0.177 | 0.276 | 0.279 | 0.177 | 0.294 |
| | sh | −0.360 | 1.702 | −0.364 | 1.771 | 1.758 | −0.365 | 1.870 |
| Tyr | c31 | −0.128 | 1.358 | −0.116 | 0.925 | 2.146 | −0.012 | 2.138 |
| | c32 | 0.136 | 0.552 | 0.063 | 1.911 | 0.000 | 0.079 | 0.000 |
| | ca1 | −0.047 | 1.724 | −0.120 | 0.957 | 1.834 | −0.067 | 1.609 |
| | ca2 | −0.451 | 0.777 | −0.495 | 1.775 | 0.994 | −0.525 | 1.601 |
| | ca3 | 0.542 | 1.183 | 0.684 | 0.000 | 0.421 | 0.693 | 0.000 |
| | ca4 | −0.127 | 0.000 | 0.007 | 0.063 | 0.000 | −0.082 | 0.000 |
| | ha1 | 0.143 | 0.251 | 0.162 | 0.350 | 0.320 | 0.155 | 0.272 |
| | ha2 | 0.208 | 0.034 | 0.214 | 0.000 | 0.097 | 0.215 | 0.000 |
| | hc1 | 0.032 | 0.195 | 0.030 | 0.223 | 0.000 | 0.002 | 0.000 |
| | hc2 | −0.009 | 0.270 | −0.007 | 0.000 | 0.462 | −0.003 | 0.523 |
| | ho | 0.376 | 0.000 | 0.389 | 0.000 | 0.056 | 0.398 | 0.199 |
| | oh | −0.581 | 0.520 | −0.626 | 0.711 | 0.773 | −0.632 | 0.585 |
| Tyr | c31 | −0.123 | 1.419 | −0.184 | 0.473 | 2.164 | −0.116 | 1.538 |
| | c32 | 0.103 | 0.444 | 0.011 | 2.169 | 0.000 | 0.010 | 1.442 |
| | ca1 | −0.271 | 1.702 | −0.425 | 0.828 | 1.643 | −0.424 | 1.059 |

**Table 2. continued**

| | | RESP | 3: RESP, Dir Pol | 4: Co-opt Chgs, Dir Pol | | 5: RESP, SC Pol | 6: Co-opt Chgs, SC Pol | |
|---|---|---|---|---|---|---|---|---|
| | | $q$ | $\alpha$ | $q$ | $\alpha$ | $\alpha$ | $q$ | $\alpha$ |
| | ca2 | −0.241 | 0.662 | −0.212 | 1.803 | 0.727 | −0.200 | 1.860 |
| | ca3 | 0.335 | 1.407 | 0.400 | 0.072 | 1.219 | 0.373 | 0.000 |
| | ca4 | 0.103 | 0.000 | 0.315 | 0.001 | 0.000 | 0.301 | 0.000 |
| | ha1 | 0.182 | 0.204 | 0.213 | 0.343 | 0.388 | 0.209 | 0.398 |
| | ha2 | 0.172 | 0.070 | 0.179 | 0.000 | 0.121 | 0.177 | 0.000 |
| | hc1 | 0.030 | 0.177 | 0.048 | 0.316 | 0.000 | 0.030 | 0.087 |
| | hc2 | −0.010 | 0.322 | −0.002 | 0.000 | 0.475 | −0.001 | 0.134 |
| | ho | 0.372 | 0.088 | 0.389 | 0.015 | 0.139 | 0.401 | 0.399 |
| | oh | −0.546 | 0.388 | −0.581 | 0.712 | 0.494 | −0.578 | 0.353 |
| Val | c31 | −0.510 | 0.000 | −0.412 | 0.000 | 0.000 | −0.416 | 0.000 |
| | c32 | 0.572 | 2.858 | 0.557 | 2.291 | 2.809 | 0.563 | 2.279 |
| | hc1 | 0.114 | 0.356 | 0.087 | 0.410 | 0.367 | 0.088 | 0.416 |
| | hc2 | −0.068 | 0.000 | −0.107 | 0.138 | 0.000 | −0.108 | 0.143 |
| Val | c31 | −0.514 | 0.000 | −0.405 | 0.000 | 0.000 | −0.409 | 0.000 |
| | c32 | 0.564 | 2.877 | 0.551 | 2.297 | 2.827 | 0.558 | 2.282 |
| | hc1 | 0.116 | 0.354 | 0.086 | 0.409 | 0.365 | 0.086 | 0.415 |
| | hc2 | −0.066 | 0.000 | −0.108 | 0.142 | 0.000 | −0.109 | 0.149 |
| Val | c31 | −0.526 | 0.000 | −0.386 | 0.000 | 0.000 | −0.392 | 0.000 |
| | c32 | 0.555 | 2.931 | 0.522 | 2.312 | 2.874 | 0.529 | 2.284 |
| | hc1 | 0.120 | 0.348 | 0.082 | 0.408 | 0.359 | 0.083 | 0.414 |
| | hc2 | −0.061 | 0.000 | −0.103 | 0.145 | 0.000 | −0.104 | 0.157 |
| Val dimer | c31 | −0.523 | 0.000 | −0.371 | 0.000 | 0.000 | −0.379 | 0.000 |
| | c32 | 0.567 | 3.148 | 0.495 | 2.540 | 3.051 | 0.503 | 2.519 |
| | hc1 | 0.118 | 0.315 | 0.079 | 0.374 | 0.341 | 0.080 | 0.390 |
| | hc2 | −0.062 | 0.000 | −0.090 | 0.109 | 0.000 | −0.090 | 0.113 |
| methane dimer | c3 | −0.500 | 0.980 | −0.496 | 0.983 | 1.020 | −0.497 | 1.018 |
| | hc | 0.125 | 0.201 | 0.124 | 0.201 | 0.196 | 0.124 | 0.196 |
| nitrobenzene | ca1 | −0.194 | 0.000 | −0.286 | 0.598 | 0.000 | −0.291 | 1.670 |
| | ca2 | −0.111 | 0.779 | −0.047 | 1.867 | 0.498 | −0.005 | 1.523 |
| | ca3 | −0.127 | 1.573 | −0.133 | 0.000 | 1.914 | −0.156 | 0.000 |
| | ca4 | 0.092 | 2.964 | 0.170 | 1.099 | 3.643 | 0.158 | 0.000 |
| | ha1 | 0.183 | 0.000 | 0.196 | 0.145 | 0.000 | 0.188 | 0.062 |
| | ha2 | 0.146 | 0.443 | 0.149 | 0.041 | 0.505 | 0.145 | 0.093 |
| | ha3 | 0.151 | 0.000 | 0.152 | 0.328 | 0.000 | 0.151 | 0.436 |
| | no | 0.751 | 0.000 | 0.715 | 0.862 | 0.000 | 0.707 | 0.062 |
| | o | −0.458 | 0.619 | −0.465 | 0.495 | 0.552 | −0.467 | 0.925 |

**Table 3. Means and Standard Deviations of Atomic Polarizabilities ($\text{Å}^3$) by Element, for Models 3−6**

| | 3: RESP, Dir Pol | | 4: Co-opt Chgs, Dir Pol | | 5: RESP, SC Pol | | 6: Co-opt Chgs, SC Pol | |
|---|---|---|---|---|---|---|---|---|
| | mean | std dev. | mean | std dev. | mean | std dev. | mean | std. dev. |
| C | 1.107 | 1.070 | 0.972 | 0.893 | 1.102 | 1.156 | 0.950 | 0.903 |
| H | 0.163 | 0.164 | 0.198 | 0.164 | 0.184 | 0.188 | 0.232 | 0.171 |
| O | 0.413 | 0.276 | 0.556 | 0.203 | 0.459 | 0.320 | 0.569 | 0.266 |
| N | 0.215 | 0.372 | 0.458 | 0.431 | 0.216 | 0.375 | 0.147 | 0.152 |

(Figure 2). Comparisons of the optimized parameters (Table 2) indicate that they also tend to be quite similar for matching molecules and atom types. For example, polarizabilities optimized for the direct and self-consistent models with RESP charges (models 3 and 5) agree well with each other (Figure 7). This agreement suggests that the optimized polarizabilities might be swapped between the two molecules with little loss of accuracy. In fact, using polarizabilities optimized for model 5 in model 3 increases the mean RMSE, $R$, by only 0.05 kcal/mol-$e$, whereas the reverse swap decreased the mean RMSE by 0.03 kcal/mol-$e$. Thus, the optimized values are effectively interchangeable. Similarly, the co-optimized charges of models 4 and 6 are quite similar to corresponding baseline RESP charges (Figure 8). Overall, the four inducible dipole models end up with similar parameters and yield similar levels of accuracy.

## ■ CONCLUSIONS

Incorporating conformation-dependent electronic polarization into force fields promises to advance the predictive power of molecular simulations. Although various polarization models exist, the self-consistent atomic polarizability model has emerged as a well-regarded standard. This approach has been implemented in its full form and approximated via the Drude oscillator method, as well as the first-order, or direct, approximation. Conformation-dependent polarization has also
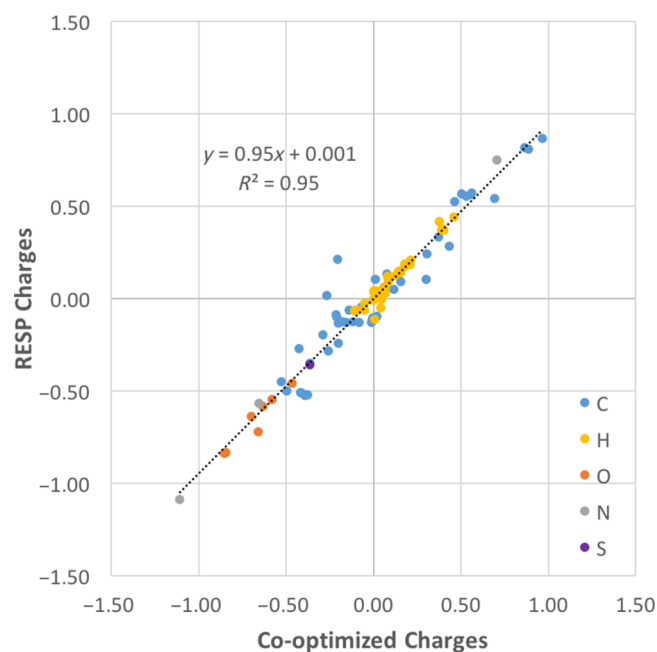
(a) Aspartate Polarizabilities          (b) Tyrosine Polarizabilities
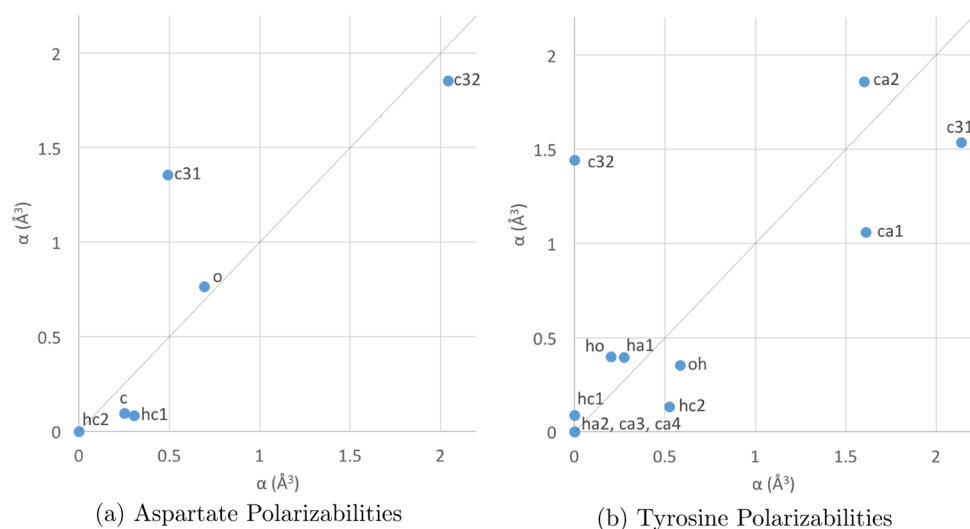
Figure 6. Comparison of model 6 atomic polarizabilities from different molecular conformations of (a) aspartate and (b) tyrosine analogues.
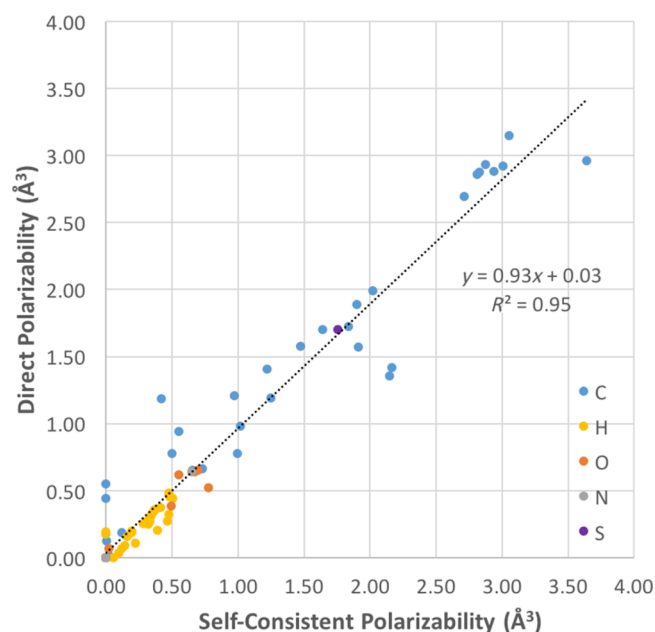


Figure 7. Comparison of self-consistent and direct polarizabilities for models 3 and 5 across all molecules.



Figure 8. Comparison of co-optimized and RESP charges for models 5 and 6 across all molecules.

been modeled via a redistribution of charge among atomic point charges,[30,32] instead of using added point dipoles. (A combined approach has also been reported.[52]) Charge redistribution approaches cannot capture polarization that is not directed along bonds, but they are simpler because they do not add point dipoles to the representation of the molecule.

In considering the relative merits of various polarization models, it is useful to distinguish between how a model represents polarization and how it computes the polarization within this representation. This study has considered two polarization representations, atom-centered point dipoles and redistribution of atom-centered charges, and it has examined two response models for the point-dipole representation, namely, direct and self-consistent inducible dipoles. (Although not studied here, there are also useful response models for the charge-based representation.[30,53]) It should be emphasized that, although the response models used in polarizable force fields are physically motivated, they are simplified representations of

the mechanisms controlling how electrons in molecules shift in response to inducing fields. This holds not only for the charge-redistribution models but also for the physically appealing picture of self-consistent, atom-centered inducible point dipoles. Thus, given that the parameters have been properly optimized, the accuracy of a polarization model can be limited by its representation of polarization, its response model, or both.

A central finding of this study is that it is the response model, rather than a polarization representation, that limits the accuracy of the self-consistent atomic polarization model: point dipoles optimized independent of any response model yield much more accurate ESPs than those optimized using a full polarization model. Indeed, the inducible dipole response model is so problematic that it yields ESPs less accurate than those achievable with a point-charge representation of polarization. Thus, although a key advantage of the dipole

representation should be its ability to capture out-of-plane polarization, the induced-dipole implementations studied here are no more accurate than point charges at capturing out-of-plane induction of the planar nitrobenzene molecule.

This means that a polarization model using a point-charge representation could outperform the standard inducible dipole model, if it were outfitted with a good response model. Further accuracy might be available at modest computational cost by adding out-of-plane charge centers above and below aromatic rings; these could improve the representation of both the baseline potential and the polarization normal to the ring. On the other hand, because the point-dipole representation can yield greater accuracy than the point-charge representation, it would also be of great interest to seek improved response models for the point-dipole representation. Although any specific directions for improvements are currently speculative, several possible approaches may be considered. One is to allow for anisotropic polarizabilities. Another would be to modify the current electrostatic treatment of short-ranged induced dipole−induced dipole interactions, which clearly does not capture the complex details of QM electronic reorganization. Finally, the success of models 1 and 2 suggests that one might develop empirical response models for chemical fragments.

We also observed only a slight improvement in accuracy on going from the first-order, or direct, approximation of the induced-dipole model to the fully self-consistent model, even though the self-consistent model is more physically complete. Perhaps greater improvement would be observed if a similar study could be done for molecules in the condensed phase, in which there would be many more dipole−dipole interactions. Nonetheless, our results support prior suggestions[17−20] that the direct approximation offers an advantageous combination of accuracy and computational efficiency. It is also interesting that polarizabilities optimized for the self-consistent model were essentially interchangeable with those optimized for the direct model.

Somewhat greater improvement was observed upon replacing regular RESP charges with partial charges co-optimized with the atomic polarizabilities, particularly for the ionized compounds. Thus, although good results can be obtained with RESP charges combined with the inducible polarization models, it makes sense to re-optimize charges in the context of the inducible dipoles for charged compounds. Indeed, this leads to improvement in both the polarized and unpolarized baseline ESPs.

In this study, the atomic polarizabilities were optimized to replicate the QM ESPs generated by molecules in the fields of external point charges. The agreement of molecular polarizabilities computed from these fitted atomic polarizabilities with those computed directly from the QM calculations supports the physical plausibility of the values assigned. On the other hand, some of the optimized polarizabilities differ significantly from previously published values. This often appears to result from compensating deviations of neighboring atoms. In addition, there is evidence that the solutions to the optimization problem can be degenerate, in the sense that equally good (or nearly so) fits can be obtained with different polarizabilities, much as observed in the optimization of point charges to match QM ESPs.[33] Procedures analogous to those used in RESP, such as the addition of weak restraints and/or the use of multiple conformations in fitting, could be used to generate more uniform polarizability assignments across chemically similar atoms.

When assessing the reliability of the present conclusions, it is reasonable to consider the degree to which the ability of a polarization model to fit the QM ESPs of polarized molecules is a useful metric of the model's quality. The central argument in support of this view is that this approach directly probes the relevant physics, and, indeed, other groups have used a similar approach.[23,52,54] In addition, RESP, one of the most successful approaches to assigning partial atomic charges, works by fitting charges to QM ESPs, so a similar approach should also be suitable for adjusting and evaluating polarization models. More particularly, the use of QM calculations at the HF/6-31G* level to calculate the reference QM ESPs is consistent with the standard RESP protocol. On the other hand, as this QM calculation leads to molecular polarizabilities that underestimate the experimental results by roughly 30%, the reliability of the parameters might benefit from use of a larger basis set, such as aug-cc-pVTZ. It is also worth noting that partial charges fitted to the HF/6-31G* results in vacuum yield dipole moments that somewhat overestimate gas-phase experimental results. These partial charges are regarded as suitable for simulations in the condensed phase, where some self-polarization occurs. However, an explicit treatment of polarizability should allow molecular dipole moments to adjust automatically to the environment, so it would be appropriate, while accounting explicitly for polarization, to set baseline gas-phase partial charges with a QM method that yields molecular dipole moments appropriate to the gas phase.

In summary, the accuracy of a polarization model is determined not only by how it represents polarization but also by the response model it uses to compute polarization. Although atom-centered point dipoles can do an excellent job of representing molecular polarization, the inducible dipole response models typically used with this representation fall well short of the theoretical maximum accuracy it could attain. Therefore, it should be possible to develop more accurate polarization models not through more detailed representations of polarization but instead through improved response models.

## ■ AUTHOR INFORMATION

**Corresponding Author**
*E-mail: mgilson@ucsd.edu. Phone: (858) 822-0622.

**Notes**

## ■ ACKNOWLEDGMENTS

## ■ 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) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.

(3) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollmann, P. A. Application of RESP Charges to Calculate Conformational Energies, Hydrogen Bond Energies, and Free Energies of Solvation. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.

(4) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force field. *J. Comput. Chem.* **2004**, *25*, 1157−74.

(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*, 1656−76.

(6) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−616.

(7) MacKerell, A. D.; Wiorkiewicz-Kuczera, J.; Karplus, M. An All-Atom Empirical Energy Function for the Simulation of Nucleic Acids. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.

(8) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of Alpha/Gamma Conformers. *Biophys. J.* **2007**, *92*, 3817−29.

(9) Henriksen, N. M.; Fenley, A. T.; Gilson, M. K. Computational Calorimetry: High-Precision Calculation of Host-Guest Binding Thermodynamics. *J. Chem. Theory Comput.* **2015**, *11*, 4377−94.

(10) Ponder, J. W.; Case, D. A. Force Fields for Protein Simulations. *Adv. Protein Chem.* **2003**, *66*, 27−85.

(11) Cieplak, P.; Dupradeau, F.-Y.; Duan, Y.; Wang, J. Polarization Effects in Molecular Mechanical Force Fields. *J. Phys.: Condens. Matter* **2009**, *21*, 333102.

(12) Palmo, K.; Mannfors, B.; Mirkin, N. G.; Krimm, S. Potential Energy Functions: From Consistent Force Fields to Spectroscopically Determined Polarizable Force Fields. *Biopolymers* **2003**, *68*, 383−94.

(13) Vanommeslaeghe, K.; MacKerell, A. D. CHARMM Additive and Polarizable Force Fields for Biophysics and Computer-Aided Drug design. *Biochim. Biophys. Acta* **2015**, *1850*, 861−71.

(14) Baker, C. M. Polarizable Force Fields for Molecular Dynamics Simulations of Biomolecules. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2015**, *5*, 241−254.

(15) Rick, S. W.; Stuart, S. J. Potentials and Algorithms for Incorporating Polarizability in Computer Simulations. *Rev. Comput. Chem.* **2002**, *18*, 89−146.

(16) Simmonett, A. C.; Pickard, F. C.; Shao, Y.; Cheatham, T. E.; Brooks, B. R. Efficient Treatment of Induced Dipoles. *J. Chem. Phys.* **2015**, *143*, 074115.

(17) Straatsma, T. P.; McCammon, J. A. Molecular Dynamics Simulations with Interaction Potentials Including Polarization Development of a Noniterative Method and Application to Water. *Mol. Simul.* **1990**, *5*, 181−192.

(18) Schyman, P.; Jorgensen, W. L. Exploring Adsorption of Water and Ions on Carbon Surfaces using a Polarizable Force Field. *J. Phys. Chem. Lett.* **2013**, *4*, 468−474.

(19) King, G.; Warshel, A. Investigation of the Free Energy Functions for Electron Transfer Reactions. *J. Chem. Phys.* **1990**, *93*, 8682.

(20) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *J. Phys. Chem. B* **2013**, *117*, 9956−72.

(21) Drude, P.; Riborg, C.; Millikan, R. A. *The Theory of Optics. Translated from the German by CR Mann and RA Millikan*; Longmans & Co.: London, 1902.

(22) Lamoureux, G.; Roux, B. Modeling Induced Polarization With Classical Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm. *J. Chem. Phys.* **2003**, *119*, 3025−3039.

(23) Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D. Determination of Electrostatic Parameters for a Polarizable Force Field Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2005**, *1*, 153−168.

(24) Yu, H.; Hansson, T.; Van Gunsteren, W. F. Development of a Simple, Self-Consistent Polarizable Model for Liquid Water. *J. Chem. Phys.* **2003**, *118*, 221−234.

(25) 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.

(26) Sharp, K.; Jean-Charles, A.; Honig, B. A Local Dielectric Constant Model for Solvation Free Energies which Accounts for Solute Polarizability. *J. Phys. Chem.* **1992**, *96*, 3822−3828.

(27) Tan, Y.-H.; Tan, C.; Wang, J.; Luo, R. Continuum Polarizable Force Field within the Poisson−Boltzmann Framework. *J. Phys. Chem. B* **2008**, *112*, 7675−7688.

(28) Truchon, J.-F.; Nicholls, A.; Roux, B.; Iftimie, R. I.; Bayly, C. I. Integrated Continuum Dielectric Approaches To Treat Molecular Polarizability and the Condensed Phase: Refractive Index and Implicit Solvation. *J. Chem. Theory Comput.* **2009**, *5*, 1785−1802.

(29) Rappé, A. K.; Goddard, W. A., III Charge Equilibration for Molecular Dynamics Simulations. *J. Phys. Chem.* **1991**, *95*, 3358−3363.

(30) Rick, S. W.; Stuart, S. J.; Berne, B. J. Dynamical Fluctuating Charge Force Fields: Application to Liquid Water. *J. Chem. Phys.* **1994**, *101*, 6141−6159.

(31) Mortier, W. J.; Van Genechten, K.; Gasteiger, J. Electronegativity Equalization: Application and Parametrization. *J. Am. Chem. Soc.* **1985**, *107*, 829−835.

(32) Ferenczy, G. G.; Reynolds, C. A. Modeling Polarization Through Induced Atomic Charges. *J. Phys. Chem. A* **2001**, *105*, 11470−11479.

(33) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; 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.

(34) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A. Development of an Accurate and Robust Polarizable Molecular Mechanics Force Field from Ab Initio Quantum Chemistry. *J. Phys. Chem. A* **2004**, *108*, 621−627.

(35) Berka, K.; Laskowski, R.; Riley, K. E.; Hobza, P.; Vondrášek, J. Representative Amino Acid Side Chain Interactions in Proteins. A Comparison of Highly Accurate Correlated ab Initio Quantum Chemical and Empirical Potential Procedures. *J. Chem. Theory Comput.* **2009**, *5*, 982−992.

(36) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *Antechamber: An Accessory Software Package for Molecular Mechanical Calculations.* http://ambermd.org.

(37) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, Revision D.01; Gaussian Inc.: Wallingford, CT, 2009.

(38) Ferrin, T. E.; Huang, C. C.; Jarvis, L. E.; Langridge, R. The MIDAS Display System. *J. Mol. Graphics* **1988**, *6*, 13−27.

(39) van der Walt, S.; Colbert, S. C.; Varoquaux, G. The NumPy Array: A Structure for Efficient Numerical Computation. *Comput. Sci. Eng.* **2011**, *13*, 22−30.

(40) Applequist, J.; Carl, J. R.; Fung, K.-K. Atom Dipole Interaction Model for Molecular Polarizability. Application to Polyatomic Molecules and Determination of Atom Polarizabilities. *J. Am. Chem. Soc.* **1972**, *94*, 2952−2960.

(41) Cieplak, P.; Caldwell, J.; Kollman, P. Molecular Mechanical Models for Organic and Biological Systems Going Beyond the Atom Centered Two Body Additive Approximation: Aqueous Solution Free Energies of Methanol and N-Methyl Acetamide, Nucleic Acid Base, and Amide Hydrogen Bonding and Chloroform. *J. Comput. Chem.* **2001**, *22*, 1048−1057.

(42) Wang, Z.-X.; Zhang, W.; Wu, C.; Lei, H.; Cieplak, P.; Duan, Y. Strike a Balance: Optimization of Backbone Torsion Parameters of

AMBER Polarizable Force Field for Simulations of Proteins and peptides. *J. Comput. Chem.* **2006**, *27*, 781−90.

(43) Thole, B. Molecular Polarizabilities Calculated with a Modified Dipole Interaction. *Chem. Phys.* **1981**, *59*, 341−350.

(44) Ren, P.; Ponder, J. W. Consistent Treatment of Inter- and Intramolecular Polarization in Molecular Mechanics Calculations. *J. Comput. Chem.* **2002**, *23*, 1497−506.

(45) Wang, J.; Cieplak, P.; Li, J.; Hou, T.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical Calculations I: Parameterization of Atomic Polarizability. *J. Phys. Chem. B* **2011**, *115*, 3091−3099.

(46) Zhu, C.; Byrd, R. H.; Lu, P.; Nocedal, J. Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization. *ACM Trans. Math. Software* **1997**, *23*, 550−560.

(47) Jones, E.; Oliphant, T.; Peterson, P.; et al. SciPy: Open Source Scientific Tools for Python. http://www.scipy.org.

(48) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796.

(49) Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. IV. Calculation of Static Electrical Response Properties. *J. Chem. Phys.* **1994**, *100*, 2975.

(50) *NIST Computational Chemistry Comparison and Benchmark Database*, 2015. http://cccbdb.nist.gov.

(51) *CRC Handbook of Chemistry and Physics*, 2015−2016. http://www.hbcpnetbase.com.

(52) Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B. J.; Friesner, R. A. Fluctuating Charge, Polarizable Dipole, and Combined Models: Parameterization from ab Initio Quantum Chemistry. *J. Phys. Chem. B* **1999**, *103*, 4730−4737.

(53) Winn, P. J.; Ferenczy, G. G.; Reynolds, C. A. Towards Improved Force Fields: III. Polarization Through Modified Atomic Charges. *J. Comput. Chem.* **1999**, *20*, 704−712.

(54) Harder, E.; Anisimov, V. M.; Vorobyov, I. V.; Lopes, P. E. M.; Noskov, S. Y.; MacKerell, A. D.; Roux, B. Atomic Level Anisotropy in the Electrostatic Modeling of Lone Pairs for a Polarizable Force Field Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2006**, *2*, 1587−1597.