

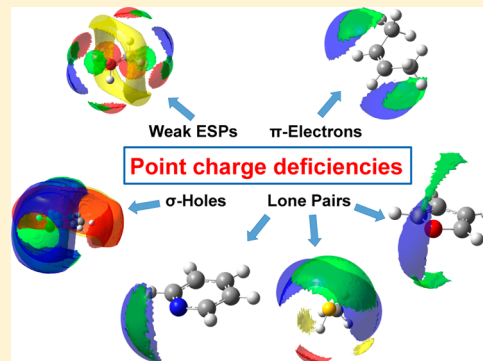
Charge Anisotropy: Where Atomic Multipoles Matter Most

Christian Kramer,* Alexander Spinn, and Klaus R. Liedl

Department of Theoretical Chemistry, Faculty for Chemistry and Pharmacy, Center for Molecular Biosciences Innsbruck (CMBI), Leopold-Franzens University Innsbruck, Innrain 80/82, A-6020 Innsbruck, Austria

S Supporting Information

ABSTRACT: Specific intermolecular interactions are largely guided by electrostatics. However, the most common model for electrostatic interactions—atomic point charges—fails to reproduce anisotropic charge distributions, such as lone pairs and sigma holes. Although this has long been known, point charges are still widely used in chemical modeling and reasoning. In this contribution, we analyze the deficiency of atomic point charges in reproducing the electrostatic potential (ESP) around molecules and find that multipole moments up to quadrupoles can, with a much lower error than point charges, reproduce the relevant ESP for all cases. Mapping the surface to the closest atom allows to compare ESP errors between atom types and to identify cases with the most urgent need for atomic multipoles. Our analysis shows that almost all heteroatoms require multipoles to correctly describe their charge distribution, with the most serious cases being nitrogen, sulfur, and halogens. Comparison with small molecule crystallography data studies supports our findings and emphasizes the need for incorporating anisotropic charge descriptions in chemical models. The scheme introduced here can be used to identify anisotropic binding preferences for atom types where there is too little coverage in crystal structure databases.



INTRODUCTION

Point charges are one of the basic chemical concepts for rationalizing, modeling, and predicting molecular interactions. This is reflected, for example, in the number of publications in the *Journal of the American Chemical Society* in the years 2012 and 2013 that use the terms RESP^{1,2} and ChelpG,³ which are two standard methods for fitting atomic point charges to the electrostatic potential around molecules (RESP, 4S; ChelpG, 12). Point charges are used to model the electrostatic interaction between atoms and molecules, which, when analyzed by energy decomposition schemes, often turns out to be the major driving force.^{4,5}

While point charges are (a) easy to picture and understand and (b) computationally inexpensive, there is no clear physical basis for assigning point charges to atoms. Therefore, many different methods for computing atomic point charges have been developed, which can yield very different results.⁶ Also, it has been known for a very long time that atomic point charges cannot model anisotropic features such as sigma holes⁷ and lone pairs.^{8,9} In contrast, all supposedly more-accurate quantum mechanical approaches inherently produce anisotropic charge distributions. Yet, point charges are still widely used for rationalizing molecular interactions, especially when it comes to modeling large biomolecules and polymers that cannot easily be treated by higher level methods. When doing so, one of the most pressing questions becomes “To what extent does the point charge approximation hold, and which features cannot be reproduced?” Whereas it is clear that there is an anisotropic charge distribution for some atom types, there is no comprehensive overview over the degree of anisotropy for all

commonly used atom types. This is a serious lack, since this would be very important information for interpreting modeled molecular interactions in large chemical systems. There is little value in approximations that are inexpensive and simple, but neglect major features of the chemical system.

In this work, we analyze the extent of atomic charge anisotropy by analyzing the error in reproducing the electrostatic potential around a molecule with point charges and mapping it back to the closest atoms. This complements the efforts of other groups that have previously analyzed the anisotropy of molecular interactions based on small molecule and protein crystal data.¹⁰ It also puts previous computational work on anisotropic features of individual elements in a wider picture.^{7,11} The advantage of the computational approach taken here is that we are independent of crystal data availability and we can obtain quantitative results that are comparable across all atom types. Thus, we are able to globally map the degree of charge anisotropy across the entire space of organic chemistry and sketch a wider picture for the need for more-refined approximations than point charges. The results are important all over chemistry, since they highlight the cases where we need to use multipoles in order to understand and predict chemical interactions.

METHODS

Reference Compounds. A set of 65 representative small molecules with the most common functional groups in organic

Received: June 30, 2014

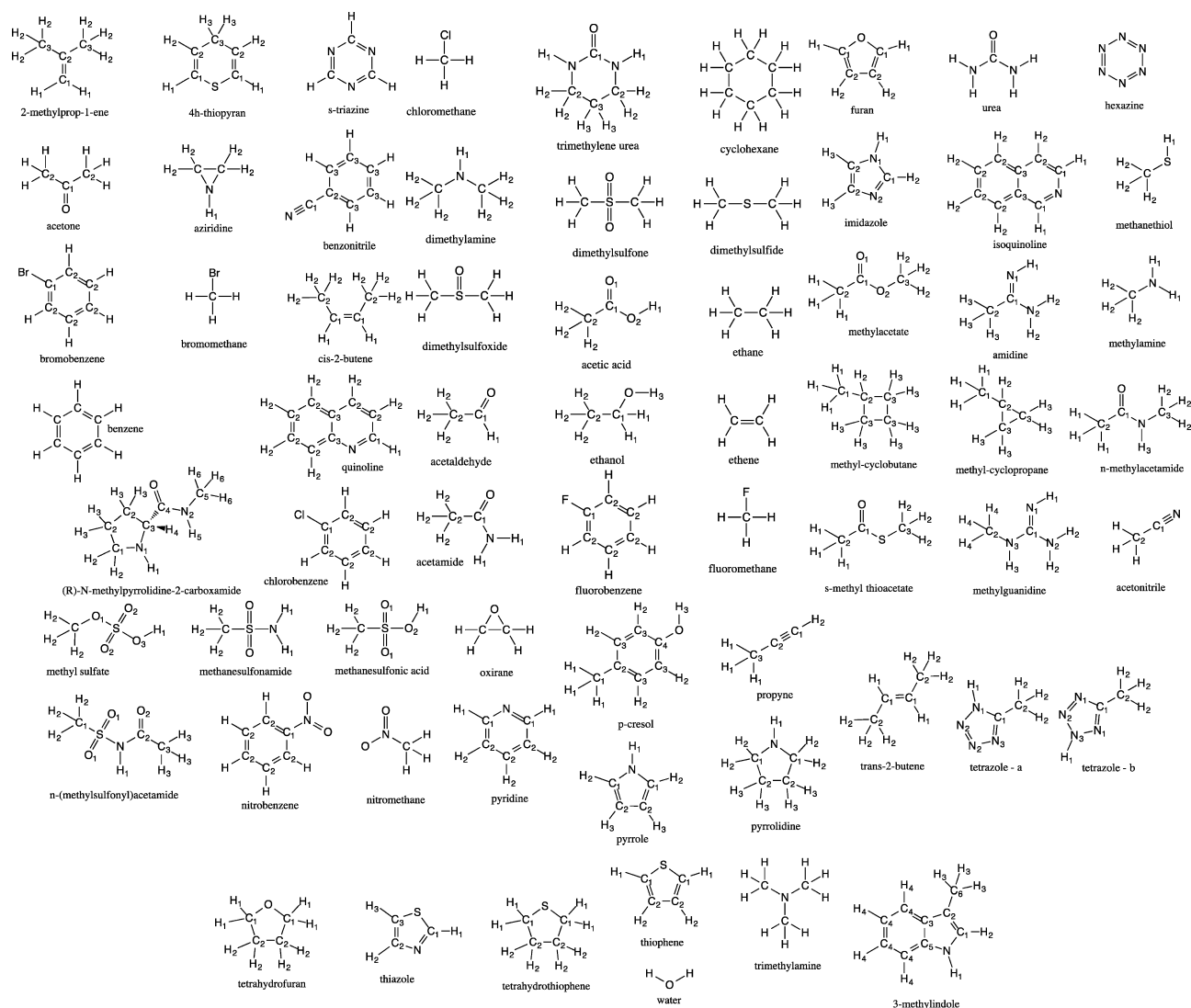


Figure 1. Reference molecules used for calculating charge anisotropy. The indices next to the atoms indicate the independent atom types, which are individual for each molecule.

molecules containing H, C, N, O, F, S, Cl, and Br were used for the reference calculations. Iodine was not included, since it has been shown to behave similarly to bromine and chlorine.¹² We did not consider charged compounds, since the charge overlays most of the effects we want to study here. The reference compounds are shown in Figure 1.

We used an atom typing scheme where every atom is assigned an atom type according to its element, the number of bonded neighbors and the element, and the number of neighbors of the neighbors, as previously published.¹³ This way, we make sure that chemically equivalent atoms have the same multipole coefficients. One could theoretically relax the atom-typing scheme even more and strictly assign different atom types to all atoms but the symmetrically equivalent ones, but it turns out that this is not a limiting factor for the accuracy of ESP reproduction and the largest errors are on atom types that only occur once.

Reference ESP. All structures were geometry-optimized in the gas phase at the M06-2X¹⁴ aug-cc-pVDZ¹⁵ level of theory, and the electrostatic potential was calculated on the same level on a rectangular grid with a grid spacing of 0.3 Å (0.1 Å for the images shown in this contribution) that extends 5 Å beyond the

molecule in the *x*-, *y*-, and *z*-directions, as previously published.¹⁶ We chose this level in order to keep the calculations at a computationally affordable cost. The quality of the reproduction of the electrostatic potential in the relevant region does not vary any more when going to higher basis sets or higher levels of theory (see reference calculations in the Supporting Information: PC and MTP parameters of acetaldehyde and dimethylsulfide have been fit to the ESP calculated at the M06-2X aug-cc-pVDZ and the CCSD(T) aug-cc-pVTZ level for comparison). Also, initial tests (details not shown) revealed that the fitted charge parameters converge at a spacing of ~0.3 Å: a smaller spacing will only increase the computing time and will not substantially change the parameters any more. Only grid points in a distance of 1.66–2.2 times the vdW radius to the nearest atom were kept in order to focus on the space that is most relevant for close intermolecular interactions.¹⁷ The ESP closer than 1.66 times the vdW radius is not relevant, since the interactions in this region are becoming repulsive and it is therefore unlikely that atoms from other molecules come that close. On the other hand, the ESP strongly decreases with distance from the

surface, and the error further than 2.2 times the vdW radius becomes rather small.

Point Charge and Multipole ESP Fitting. Atomic point charges, dipole moments, and quadrupole moments were fitted to the ESP, as previously published.^{13,16} Briefly, the contribution of each atom type's point charge and multipole parameters to the ESP at each grid point can be separated into a geometric term and a coefficient. The geometric terms can be precalculated and a linear equation system can be set up for the coefficients. This system can be solved directly and efficiently with classical linear algebra methods. It turns out that charge and multipole ESP fitting requires restraints to keep the parameters in a physically acceptable range, since atoms that are far away from the molecular surface only have a small influence on the outcome and can therefore be assigned unrealistically high parameters. Although we only use small molecules with a low risk for this effect, we nevertheless use restraints to keep the parameters in a physically acceptable range (for details, see ref 16). Note that every molecule is fitted individually, and atom types are not artificially forced to be similar between different molecules. If the multipoles were to be used in molecular dynamics simulations,^{18–20} the protocol for deriving them might require higher-level quantum mechanical calculations and conformational dependency would have to be considered. Also, other methods that avoid direct (and partially arbitrary) restraints in the parameter fitting could be used.²¹ However, since the purpose of this study is to map the degree of charge anisotropy of common chemical functionalities, we prefer to keep the calculations as simple as possible and direct interested readers to articles that touch issues of deriving multipoles for molecular dynamics simulations.^{22,23}

Statistical Analysis. Each grid point is assigned to the nearest atom for the analysis. The grid point to atom assignment for acetaldehyde is shown in Figure 2.

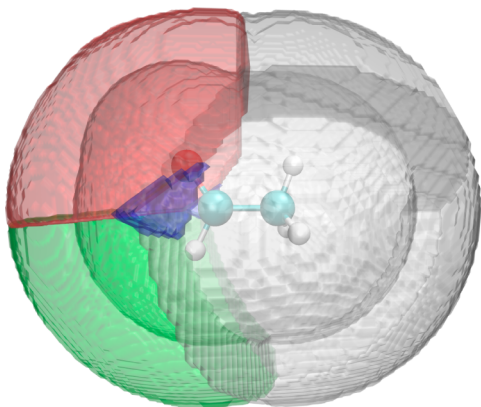


Figure 2. Grid point assignment for acetaldehyde. All points are located in the first interaction shell outside of the molecule between 1.66 and 2.2 times the vdW Radius. Colors indicate the closest atoms: red, oxygen; green, hydrogen of the CHO group; blue, carbon of the CHO group; white, hydrogens of the CH₃ group. There are no grid points assigned to the carbon of the CH₃ group.

Atom types with fewer than 20 grid points assigned were skipped. The reproduction of the quantum mechanical ESP around the individual atoms is analyzed using three different metrics: Pearson's correlation coefficient, the root mean squared error (RMSE), and the relative mean absolute error (RMAE).

Pearson's R^2 correlation coefficient is used in order to analyze whether the ESP is qualitatively correctly reproduced. A low R^2 indicates that the ESP produced by the point charges or multipoles has different maxima and minima compared to the QM ESP and therefore misses donor or acceptor features.

$$R^2 = \frac{[\sum_i (\text{ESP}_{i,\text{model}} - \overline{\text{ESP}}_{\text{model}})(\text{ESP}_{i,\text{QM}} - \overline{\text{ESP}}_{\text{QM}})]^2}{\sum_i (\text{ESP}_{i,\text{model}} - \overline{\text{ESP}}_{\text{model}})^2 \sum_i (\text{ESP}_{i,\text{QM}} - \overline{\text{ESP}}_{\text{QM}})^2}$$

The summation is over all n grid points, indexed by i , that belong to a specific atom, where $\text{ESP}_{i,\text{model}}$ indicates the ESP due to the point charge or multipole model, $\text{ESP}_{i,\text{QM}}$ indicates the ESP from the quantum mechanical calculation, $\overline{\text{ESP}}_{\text{model}}$ and $\overline{\text{ESP}}_{\text{QM}}$ indicate the average ESP of all grid points that belong to a certain atom type due to the model and the QM calculation, respectively.

The Root Mean Squared Error (RMSE) is used to quantitatively measure the absolute error in reproducing the ESP. Atom types with a high R^2 and a high RMSE have the correct features, but the directionality of the polar interactions is underestimated. The absolute values of the RMSE can best be interpreted when compared among different atom types. The RMSE is given in volts.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_i (\text{ESP}_{i,\text{model}} - \text{ESP}_{i,\text{QM}})^2}$$

The Relative Mean Absolute Error (RMAE) measures the fraction of the total variation in the ESP that is modeled correctly. A low RMSE and a high RMAE is obtained for nonpolar and weakly polar atom types whose ESP environment is reproduced badly. A high R^2 value and a high RMAE value are obtained for atom types where there is a correlation between the model ESP and the QM ESP, but the actual values of the model ESP are completely off. The RMAE value is given as a percentage.

$$\text{RMAE} = \frac{\sum_i |\text{ESP}_{i,\text{model}} - \text{ESP}_{i,\text{QM}}|}{\sum_i |\text{ESP}_{i,\text{QM}}|}$$

All quantum mechanical calculations were carried out using Gaussian09.²⁴ The ESP fitting and analysis was implemented in an own python program, which is available from the authors upon request.

RESULTS

Atomic multipole moments up to quadrupoles reproduce the relevant ESP around each individual atom with a much lower error than point charges. Figure 3 shows a plot of the RMSEs from the point charge model versus the RMSEs from the multipole model for all atoms of all molecules considered in this study.

Atomic Multipole Moments up to Quadrupoles. The largest overall errors are obtained for the sulfur of dimethylsulfide and dimethyl sulfoxide (RMSE = 9.3 and 9.1 mV). The largest RMAEs are obtained for the acid carbon (C₁) of acetic acid (19%) and the aromatic carbon of nitrobenzene (C₁) where the nitro group is attached (14%). However, there are only very few grid points assigned to both these atoms, and the ESP around them takes small absolute values due to a change in sign. For small absolute ESP values, small changes can result in large RMAEs. The lowest R^2 is obtained for the same aromatic carbon of nitrobenzene (0.97) and the sulfur of

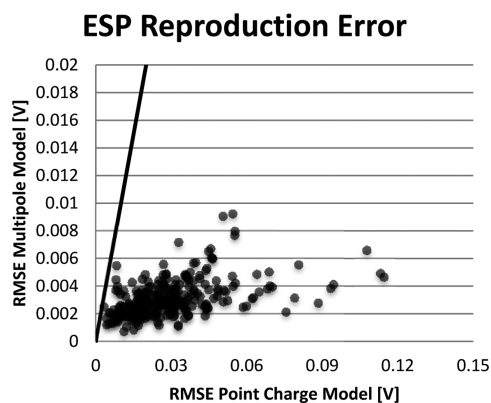


Figure 3. Comparison of the RMSEs obtained with point charges versus the RMSEs obtained with multipole moments for all atom types of all molecules analyzed. The black line indicates the line of equality. Note that the axes are scaled differently.

dimethylsulfide (0.98). All these values are much better than the corresponding values for the point-charge model. The RMSE from the MTP models is typically at least 10 times lower than the corresponding RMSE from the point-charge model.¹³ Therefore, we will not interpret the multipole models much deeper and just take as given that multipoles up to quadrupoles reproduce the relevant ESP for organic molecules very well.

Atomic Point-Charge Model. In the following, we present the results per element, starting with the halogens where the well-known sigma hole serves as a benchmark for comparing to other atom types.

Fluorine, Chlorine, and Bromine. R^2 , RMSE, and RMAE for F, Cl, and Br are shown in Table 1.

Table 1. Quality of the QM-ESP Reproduction with Point Charges around Halogens

element	parent molecule	R^2	RMSE [mV]	RMAE [%]
F	fluoromethane	0.996	19	4.5
	fluorobenzene	0.991	8.5	2.2
Cl	chloromethane	0.62	31	11
	chlorobenzene	0.35	38	16
Br	bromomethane	0.34	37	15
	bromobenzene	0.31	44	21

Table 1 shows that the ESP of fluorine within the reference compounds can be modeled well with point charges, whereas this approximation fails for chlorine and bromine. This is to be expected, since chlorine and bromine both have the sigma hole along the sigma axis. The sigma hole describes an electro-positive region along the sigma axis of halogens, which renders halogens as donors with a highly localized directionality. The sigma hole has previously been described both computationally⁷ and by analysis of interaction preferences within X-ray data.²⁵ The sigma hole of chloromethane is rather weak, as can be seen from $R^2 = 0.62$, which is twice as high as the R^2 for chlorobenzene and both bromo derivatives which have a stronger sigma hole. Typically, fluorine does not have a sigma hole, as in the reference compounds shown here. However, even fluorine can have a sigma hole, if there are strongly electron-withdrawing groups in its vicinity.^{12,26} The RMSE for the halogens with the stronger sigma hole is ~ 40 mV. This

number can serve as a reference for the error in ESP reproduction that is obtained when a feature such as the sigma hole is completely missed.

Sulfur. The ESP around sulfur can only be modeled very badly by point charges. Table 2 shows R^2 , RMSE, and RMAE for sulfur of the various sulfur-containing reference compounds.

Table 2. Quality of the QM-ESP Reproduction with Point Charges around Sulfur in Various Environments

parent molecule	R^2	RMSE [mV]	RMAE [%]
methanethiol	0.19	68	19
<i>s</i> -methyl thioacetate	0.21	44	39
tetrahydrothiophene	0.59	48	10
dimethylsulfide	0.81	54	14
4H-thiopyran	0.89	30	10
thiazole	0.89	26	32
thiophene	0.93	23	15
dimethyl sulfoxide	0.95	51	30

The ESP around the thiol sulfur in methanethiol is reproduced very badly with $R^2 = 0.19$ and RMSE = 68 mV. This indicates that some donor/acceptor features are completely missing. The same holds for *s*-methyl thioacetate and in parts for tetrahydrothiophene. The R^2 is higher for the other sulfurs, but the RMSE is higher than 50 mV for dimethylsulfide and dimethyl sulfoxide. This indicates that the directionality of the interactions is not correctly modeled. In particular, the description of the lone pairs is insufficient. Also, the RMAE for thiazole and dimethyl sulfoxide is $\sim 30\%$, indicating that the relative strength of the ESP cannot be reproduced. Figure 4 shows a comparison of the quantum mechanical ESP and the ESP reproduced with point charges for tetrahydrothiophene.

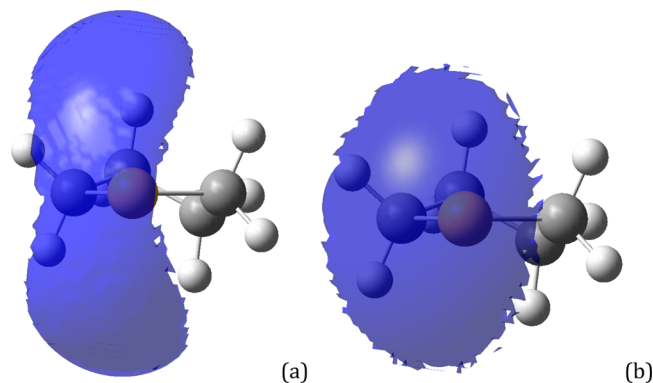


Figure 4. Volume around tetrahydrothiophene with ESP lower than -0.45 V (threshold chosen for illustration purposes), calculated with (a) QM and (b) point charges. The comparison shows that the point charge model yields different regions of low ESP close to sulfur along the C–S axis.

Figure 4 shows that the point-charge model reproduces the ESP around sulfur of tetrahydrothiophene rather badly. Specifically, it overestimates the acceptor feature in the plane and underestimates the donor feature above and below the plane. In the plane, sulfur can serve as donor, with two highly localized donor features along the C–S axis. This feature has previously been described by Politzer et al.,²⁷ and small

molecule crystal structure studies have also confirmed the propensity of sulfur to act as donor along this axis.^{28,29}

Oxygen. Our reference compounds contain 31 different types of oxygen, where the quality of the ESP reproduction varies greatly. In Table 3, we present selected cases where the point charge approximation fails. The full list for all oxygen atoms can be found in the Supporting Information.

Table 3. Quality of the QM-ESP Reproduction with Point Charges around Selected Oxygens

parent molecule (index)	R^2	RMSE [mV]	RMAE [%]
furan	0.57	44	15
methyl acetate (O2)	0.90	35	29
methylsulfate (O1)	0.95	46	26
methylsulfate (O3)	0.97	45	15
tetrahydrofuran	0.995	45	7.0
acetic acid (O2)	0.98	18	15
acetaldehyde	0.987	25	4.0

The R^2 value for furan is only 0.57. This indicates that one of the interaction features is missing in the point-charge model. Figure 5 shows a comparison of the QM ESP with the point charge ESP for furan.

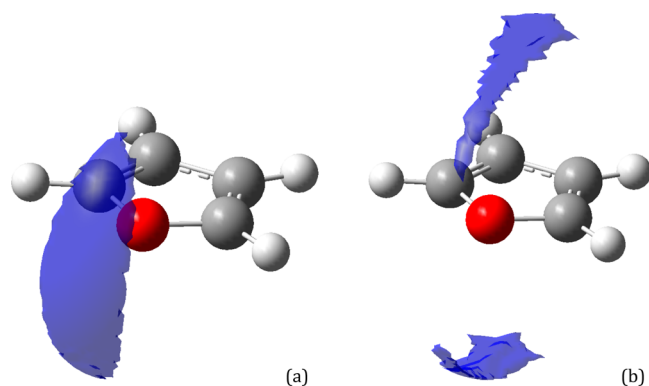


Figure 5. Volume around furan with ESP lower than -0.38 V (threshold chosen for illustration purposes) calculated with (a) QM and (b) point charges. As can be seen, the area comprising the minima of the ESP values is located in different areas, leading to a qualitative error when using point charges only.

The minima of the QM ESP of furan cannot be reproduced with point charges, leading to a qualitatively (and quantitatively) incorrect picture of donor and acceptor properties. For the other oxygen types, the correlation with the QM ESP is above 0.90. However, some of the oxygens have a large RMSE over 40 mV, as shown in Table 3. This error indicates that the strength of the directionality of the interactions (lone pairs) is not well-modeled with only point charges. Also, the RMAE is above 10% and approaches 30% for several oxygen types, indicating that the relative strength of the polarity cannot be well-modeled using point charges. Interestingly, the more problematic oxygens are all sp^3 oxygens within ether or hydroxy functionalities. Carboxy oxygens, for example, in acetaldehyde, are less problematic with point charges. Nevertheless, atomic multipole moments can still reduce the error in describing the relevant ESP for acetaldehyde by 1 order of magnitude.

Nitrogen. Nitrogens have the highest RMSE errors for reproducing the QM ESP. Table 4 shows a selection out of the

36 different nitrogen types in our reference compounds with the most problematic nitrogens.

Table 4. Quality of the QM-ESP Reproduction with Point Charges around Selected Nitrogens

parent molecule (index)	R^2	RMSE [mV]	RMAE [%]
pyrrolidine	0.88	114	17
dimethylamine	0.80	113	17
tetrazole_b (N3)	0.72	88	164
aziridine	0.89	79	11
pyridine	0.91	63	11
imidazole (N2)	0.93	60	7.6

The RMSE around nitrogens is the highest among all atom types in this study. This is because the nitrogens' lone pair has a high directionality that cannot be captured with atomic point charges. Table 4 shows that this is the case, both for nitrogen species that would be charged at pH 7.4, such as dimethylamine, and for species with a lower pK_a , such as pyridine. Tetrazole is a quite dramatic case, since the RMAE here is even above 100%, indicating that the relative strength of the ESP is totally off around the N–H nitrogen. This clearly shows that point charges alone cannot appropriately reproduce the binding features of nitrogens with lone pairs. The deficiency of point charge models has been noted previously,³⁰ yet the most common protein force fields still model, for example, the histidine side chain with atom-centered point charges.

An academic yet interesting case is hexazine, the nitrogen analogue of benzene. Hexazine consists of six nitrogens in a planar ring arrangement, similar to the carbons of benzene (see Figure 6). It has never been synthesized, so the analysis is

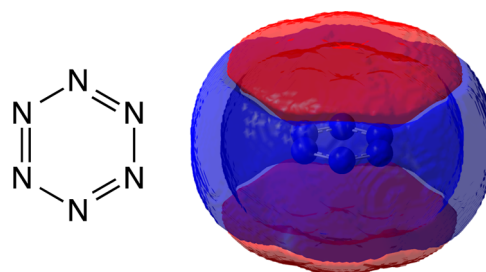


Figure 6. Hexazine and the sign of its QM ESP in the first interaction belt (red, positive ESP; blue, negative ESP).

rather of theoretical interest. Because of symmetry, the point charge on each nitrogen atom must be zero for uncharged hexazine. A QM calculation indicates that the π -electrons of hexazine are highly delocalized and anisotropically distributed around the ring. The RMSE for reproducing the ESP of hexazine with point charges (which all have to be zero) is 198 mV (the maximum possible error), whereas the RMSE for reproducing the ESP with multipoles up to quadrupoles is 2.1 mV, roughly 2 orders of magnitude lower.

Carbon. Sixty one (61) different carbon atom types occur in the reference compounds. The errors in reproducing the ESP vary widely, in terms of RMSE from 3.2 mV for the carboxy carbon (C_1) of acetaldehyde to 44 mV for the C_1 carbon in imidazole. Table 5 shows a selection of the carbon atom types with the most severe errors.

The errors around sp^2 hybridized carbons are generally larger than the errors around sp^3 hybridized carbons. This

Table 5. Quality of the QM-ESP Reproduction with Point Charges around Selected Carbons

ID/parent molecule	R^2	RMSE [mV]	RMAE [%]
tetrazole_b (C ₁)	0.77	37	77
benzonitrile (C ₃)	0.93	20	42
imidazole (C ₁)	0.97	44	19
cis-2-butene (C ₁)	0.95	43	19
amidine (C ₁)	0.996	42	13

makes sense, since sp² hybridized carbons generally have a more quadrupolar charge distribution. The largest errors in terms of RMSE are obtained for the carbon between the two nitrogens in imidazole and the sp² carbons in cis-2-butene and amidine, all having an RMSE slightly above 40 mV. The largest relative errors are obtained for the aromatic carbon in tetrazole and the aromatic carbon where the nitrile group is attached in benzonitrile.

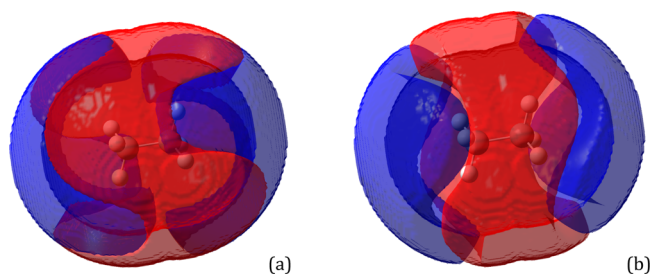
Hydrogen. Hydrogens are of crucial importance for accurately modeling the ESP, since they are very close to the surface. There are 128 different hydrogen types in the reference compounds. The RMSE varies between 5.1 mV and 74 mV. In Table 6, we present a selection of the hydrogens with the most serious deficiencies in the modeled ESP.

Table 6. Quality of the QM-ESP Reproduction with Point Charges around Selected Hydrogens

ID/parent molecule	R^2	RMSE [mV]	RMAE [%]
water	0.95	74	20
dimethylamine (H ₁)	0.99	69	27
methanethiol (H ₁)	0.91	65	38
methanethiol (H ₂)	0.84	58	28
methylamine (H ₂)	0.89	55	47
cyclohexane	0.08	21	95
ethane	0.002	24	99
methylcyclobutane (H ₃)	0.34	29	75
tetrahydrofuran (H ₂)	0.57	33	20

The largest errors in terms of RMSE (>65 mV) are observed around polar hydrogens attached to heteroatoms and hydrogens on carbons next to heteroatoms (~55 mV). The ESP around these hydrogens and thus the directionality of the hydrogen bonds is modeled particularly badly with atomic point charges only. However, the ESP around aliphatic hydrogens is also problematic. While the RMSE for aliphatic hydrogens is small since they are rather apolar, the R^2 value can become very small and the RMAE can be very large, indicating that the relative positioning of maxima and minima is qualitatively wrong. This becomes apparent from comparing the QM ESP around ethane and the ESP modeled with point charges, as shown in Figure 7.

The maxima and minima of the ESP around ethane are in a completely incorrect position when modeled using point charges. The overall variation in the ESP around ethane and other alkanes is low, but not zero. However, since the correlation between modeled and QM ESP is so low, it appears that there is no point in using point charges for modeling ethane. The influence of the bad performance of point charges on modeled alkane properties has been studied previously.³¹

**Figure 7.** (a) QM ESP and (b) ESP modeled with point charges around ethane (red, positive ESP; blue, negative ESP). Ethane molecules are shown in the same orientation in both cases.

The individual errors in terms of RMSE per element are summarized in Figure 8. The distributions of R^2 per element can be found in the Supporting Information.

DISCUSSION

We have analyzed the error in reproducing the local ESP around a variety of atom types in reference molecules that represent the most common functional groups in organic chemistry. The results show that, for every element, there are atom types whose surrounding ESP is modeled particularly badly with atomic point charges. For chlorine and bromine, atomic point charges fail because they are not capable of modeling the sigma hole. For sulfur, point charges fail, for example, in thioles and thioesters and thioethers, because they are not able to model the donor features along the C–S axis. For nitrogen and sp³ oxygens, point charges fail because they are not able to model the high directionality of the lone pairs. For sp²-hybridized carbons, point charges fail to model the anisotropic distribution of the π -electrons. Lastly, point charges fail to model the high directionality of the ESP around hydrogens attached to heteroatoms, which can act as H-bond donors, and they completely fail to reproduce the weak polarity around alkanes. Therefore, caution must be exercised when modeling and rationalizing intermolecular interactions based on point charges. There are many interaction preferences that cannot be explained with point charges. The above analysis should stir up everybody who likes to think in terms of point or partial charges, since they operate at a very low level of resolution. Atomic multipoles up to quadrupoles can alleviate most of the problems of point charges.

Technically, a variety of models are able to capture anisotropic charge distributions. All quantum mechanical approaches naturally model anisotropic charge distributions. Then, several computationally less expensive force field extensions have been developed that are all capable of representing anisotropic charge distributions: For this analysis, we used atomic multipole moments up to quadrupoles which have sparked a wider interest in the context of next-generation force fields.^{32–37} Alternatives to atomic multipole moments for modeling anisotropic charge distributions include off-site point charges,^{30,38,39} and Gaussian multipoles.⁴⁰

Mathematically, multipole moments and off-site charges are not just arbitrarily extending the function space—they have well-established interpretations in chemistry (lone pairs, π -electrons, sigma holes). If sensible constraints are used in the fit, the multipole parameters do not overfit the ESP but result in interpretable atomic features. It also needs to be noted that anisotropy is not limited to the charge model—the polar flattening of the electron density that leads to effects such as

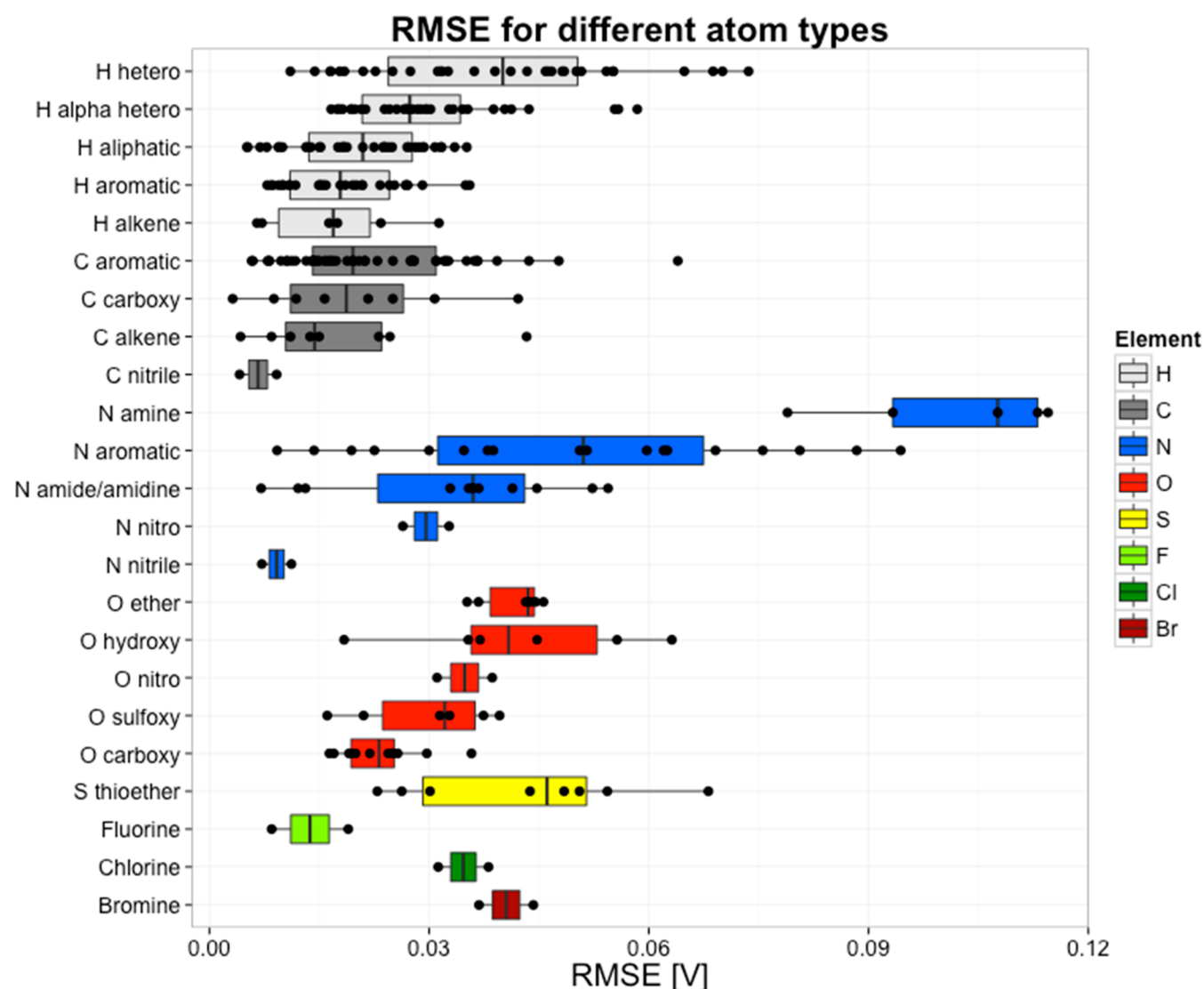


Figure 8. Distribution of RMSE per atom type shown as a boxplot. The left and right of the box indicate the 25th and 75th percentiles. The whiskers indicate 1.5 times the range between 25th and 75th percentile (values outside this range are often considered as outliers). The black dots indicate the values calculated for all atom types of the individual molecules.

sigma holes also causes some degree of anisotropy of the van der-Waals interactions.⁴¹

Since most anisotropic interactions are driven by electrostatics, the analysis scheme presented here can be used as a general tool for identifying the propensity to undergo anisotropic interactions. The combination of our computational tool with small molecule crystal data analysis can generate strong evidence for anisotropic binding preferences. While we have only calculated the anisotropic charge distribution for a set of the most common functional groups, the analysis can easily be extended to more-specific fragments for which there is low or zero crystal structure coverage.

CONCLUSIONS

In this contribution, we have quantified the failure of point charges in accurately reproducing the electrostatic potential (ESP) around specific atom types. The introduced analysis scheme allows mapping the ESP to the closest atoms and thereby comparing atom types among each other.

The analysis shows that, for every common organic chemistry element, there are cases where point charges fail

badly. They either fail completely in reproducing the relevant interaction features, or they underestimate the directionality of the interactions. Atomic multipole moments up to quadrupoles, in contrast, can represent the ESP of all molecules studied very accurately. Therefore, we want to stress that, for modeling and understanding intermolecular interactions, it is important to use multipoles. Often enough, atom-centered point charges are qualitatively and quantitatively insufficient and cannot explain a wide variety of interaction types.

While we have concentrated on the most common functional groups in organic chemistry, the scheme introduced here can directly be used to screen for potential anisotropic preferences in rare functional groups and provide evidence for anisotropic binding preferences in situations where there is too little coverage in crystal structure databases.

ASSOCIATED CONTENT

Supporting Information

A Microsoft Excel sheet containing the individual atomic errors for all reference compounds calculated in this study and results from reference calculations to justify the basis set and the level

of theory used is provided as Supporting Information. A summary plot showing the distribution of the correlations between point charge and ab initio ESP per element is also provide in the Supporting Information. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Tel.: 0043 512 507 57103. E-mail: Christian.Kramer@uibk.ac.at.

Notes

The authors declare no competing financial interest.

REFERENCES

- (1) Singh, U. C.; Kollman, P. A. An Approach to Computing Electrostatic Charges for Molecules. *J. Comput. Chem.* **1984**, *5*, 129–145.
- (2) Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11*, 431–439.
- (3) 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.
- (4) *Intermolecular Forces and Clusters*; Wales, D. J., Ed.; Structure and Bonding; Springer: Berlin, New York, 2005.
- (5) Stone, A. J. *The Theory of Intermolecular Forces*, Second Edition; Oxford University Press: Oxford, U.K., 2013.
- (6) Wiberg, K. B.; Rablen, P. R. Comparison of Atomic Charges Derived via Different Procedures. *J. Comput. Chem.* **1993**, *14*, 1504–1518.
- (7) Clark, T.; Hennemann, M.; Murray, J. S.; Politzer, P. Halogen Bonding: The Σ -hole. *J. Mol. Model.* **2007**, *13*, 291–296.
- (8) Stone, A. J.; Alderton, M. Distributed Multipole Analysis. *Mol. Phys.* **1985**, *56*, 1047–1064.
- (9) Rein, R. On Physical Properties and Interactions of Polyatomic Molecules: with Application to Molecular Recognition in Biology. In *Advances in Quantum Chemistry*; Löwdin, Per-Olov, Ed.; Academic Press: New York, 1973; Vol. 7, pp 335–396.
- (10) Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.* **2010**, *53*, S061–S084.
- (11) Murray, J. S.; Lane, P.; Politzer, P. Expansion of the Sigma-hole Concept. *J. Mol. Model.* **2009**, *15*, 723–729.
- (12) Politzer, P.; Murray, J. S.; Clark, T. Halogen Bonding and Other Σ -hole Interactions: A Perspective. *Phys. Chem. Chem. Phys.* **2013**, *15*, 11178–11189.
- (13) 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.
- (14) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Non-covalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (15) 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.
- (16) Kramer, C.; Bereau, T.; Spinn, A.; Liedl, K. R.; Gedeck, P.; Meuwly, M. Deriving Static Atomic Multipoles from the Electrostatic Potential. *J. Chem. Inf. Model.* **2013**, *53*, 3410–3417.
- (17) Rowland, R. S.; Taylor, R. Intermolecular Nonbonded Contact Distances in Organic Crystal Structures: Comparison with Distances Expected from van Der Waals Radii. *J. Phys. Chem.* **1996**, *100*, 7384–7391.
- (18) Bereau, T.; Kramer, C.; Meuwly, M. Leveraging Symmetries of Static Atomic Multipole Electrostatics in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, *9*, 5450–5459.
- (19) 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.
- (20) Kramer, C.; Gedeck, P.; Meuwly, M. Multipole-Based Force Fields from Ab Initio Interaction Energies and the Need for Jointly Refitting All Intermolecular Parameters. *J. Chem. Theory Comput.* **2013**, *9*, 1499–1511.
- (21) Tan, J. S.; Boerrigter, S. X. M.; Scaringe, R. P.; Morris, K. R. Application of Error-ranked Singular Value Decomposition for the Determination of Potential-derived Atomic-centered Point Charges. *J. Comput. Chem.* **2009**, *30*, 733–742.
- (22) 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*, 9956–9972.
- (23) Ren, P.; Wu, C.; Ponder, J. W. Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules. *J. Chem. Theory Comput.* **2011**, *7*, 3143–3161.
- (24) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, USA, 2009.
- (25) Politzer, P.; Murray, J. S.; Clark, T. Halogen Bonding: An Electrostatically-driven Highly Directional Noncovalent Interaction. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7748–7757.
- (26) Metrangola, P.; Murray, J. S.; Pilati, T.; Politzer, P.; Resnati, G.; Terraneo, G. Fluorine-Centered Halogen Bonding: A Factor in Recognition Phenomena and Reactivity. *Cryst. Growth Des.* **2011**, *11*, 4238–4246.
- (27) Murray, J. S.; Lane, P.; Politzer, P. Simultaneous Σ -hole and Hydrogen Bonding by Sulfur- and Selenium-containing Heterocycles. *Int. J. Quantum Chem.* **2008**, *108*, 2770–2781.
- (28) Zhou, P.; Tian, F.; Lv, F.; Shang, Z. Geometric Characteristics of Hydrogen Bonds Involving Sulfur Atoms in Proteins. *Proteins Struct. Funct. Bioinform.* **2009**, *76*, 151–163.
- (29) Allen, F. H.; Bird, C. M.; Rowland, R. S.; Raithby, P. R. Hydrogen-Bond Acceptor and Donor Properties of Divalent Sulfur (Y-S-Z and R-S-H). *Acta Crystallogr., Sect. B: Struct. Sci.* **1997**, *53*, 696–701.
- (30) Dixon, R. W.; Kollman, P. A. Advancing Beyond the Atom-centered Model in Additive and Nonadditive Molecular Mechanics. *J. Comput. Chem.* **1997**, *18*, 1632–1646.
- (31) Van de Streek, J.; Verwer, P.; Bennema, P.; Vlieg, E. On the Irrelevance of Electrostatics for the Crystal Structures and Polymorphism of Long Even N-alkanes. *J. Comput. Chem.* **2002**, *23*, 365–370.
- (32) Williams, D. E. Representation of the Molecular Electrostatic Potential by Atomic Multipole and Bond Dipole Models. *J. Comput. Chem.* **1988**, *9*, 745–763.
- (33) Sokalski, W. A.; Hariharan, P. C.; Kaufman, J. J. Library of Cumulative Atomic Multipole Moments. I. Nucleic Acid Bases. *Int. J. Quantum Chem.* **1987**, *32*, 111–126.
- (34) Stone, A. J. Distributed Multipole Analysis: Stability for Large Basis Sets. *J. Chem. Theory Comput.* **2005**, *1*, 1128–1132.

(35) Piquemal, J.-P.; Gresh, N.; Giessner-Prettre, C. Improved Formulas for the Calculation of the Electrostatic Contribution to the Intermolecular Interaction Energy from Multipolar Expansion of the Electronic Distribution. *J. Phys. Chem. A* **2003**, *107*, 10353–10359.

(36) 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.

(37) Joubert, L.; Popelier, P. L. A. Improved Convergence of the “Atoms in Molecules” Multipole Expansion of Electrostatic Interaction. *Mol. Phys.* **2002**, *100*, 3357–3365.

(38) Jorgensen, W. L.; Schyman, P. Treatment of Halogen Bonding in the OPLS-AA Force Field; Application to Potent Anti-HIV Agents. *J. Chem. Theory Comput.* **2012**, *8*, 3895–3801.

(39) Chessari, G.; Hunter, C. A.; Low, C. M. R.; Packer, M. J.; Vinter, J. G.; Zonta, C. An Evaluation of Force-Field Treatments of Aromatic Interactions. *Chem.—Eur. J.* **2002**, *8*, 2860–2867.

(40) 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.

(41) El Kerdawy, A.; Murray, J. S.; Politzer, P.; Bleiziffer, P.; Heßelmann, A.; Görling, A.; Clark, T. Directional Noncovalent Interactions: Repulsion and Dispersion. *J. Chem. Theory Comput.* **2013**, *9*, 2264–2275.