

Charge Anisotropy of Nitrogen: Where Chemical Intuition Fails

Alexander Spinn,[†] Philip H. Handle,[†] Johannes Kraml, Thomas S. Hofer, and Klaus R. Liedl*



Cite This: *J. Chem. Theory Comput.* 2020, 16, 4443–4453



Read Online

ACCESS |



Metrics & More

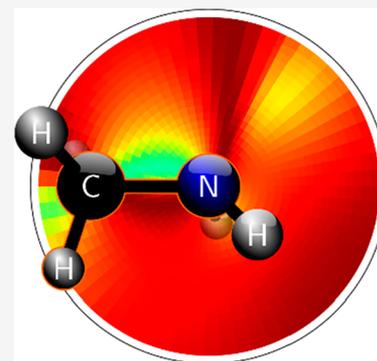


Article Recommendations



Supporting Information

ABSTRACT: For more than half a century computer simulations were developed and employed to study ensemble properties of a wide variety of atomic and molecular systems with tremendous success. Nowadays, a selection of force-fields is available that describe the interactions in such systems. A key feature of force-fields is an adequate description of the electrostatic potential (ESP). Several force-fields model the ESP via point charges positioned at the atom centers. A major shortcoming of this approach, its inability to model anisotropies in the ESP, can be mitigated using additional charge sites. It has been shown that nitrogen is the most problematic element abundant in many polymers as well as large molecules of biological origin. To tackle this issue, small organic molecules containing a single nitrogen atom were studied. In performing rigorous scans of the surroundings of these nitrogen atoms, positions where a single extra charge can enhance the ESP description the most were identified. Significant improvements are found for ammonia, amines, and amides. Interestingly, the optimal location for the extra charge does not correlate with the chemically intuitive position of the nitrogen lone pair. In fact, the placement of an extra charge in the lone-pair location does not lead to significant improvements in most cases.



INTRODUCTION

Atomistic and molecular simulations are one of the most common tools employed in numerical statistical mechanics. They can be utilized to study very diverse systems spanning atoms, small molecules, peptides, colloids, polymers, and biomolecules.^{1–3} A prerequisite for such simulations is the physical representation of the inter- and intramolecular interactions present in systems of interest. Nowadays, a host of ready-made force-fields is available, including AMBER,^{4–6} CHARMM,^{7–9} GROMOS,^{10–12} and OPLS^{13–15} sparing most users the tedious task of force-field parametrization. A cornerstone of force-field parametrization is the description of the electrostatic potential (ESP) of the molecules. Most commonly, the ESP is approximated by a set of point charges located at the centers of all atoms.^{4–17} Despite the success of thus modeled systems, point charges have obvious shortcomings in representing anisotropies in the ESP.^{18–26} One way to remedy this issue is the use of atom centered multipoles.^{22,27,28} The use of multipoles in simulations is, however, computationally more demanding than the use of point charges.

Another, computationally less expensive approach is the use of off-center charges (OCCs) and/or extra charges (ECs) to model anisotropies. In the OCC case, the number of charges still equals the number of atoms in the system, but the locations of the charges do not necessarily coincide with the atoms' centers. In the EC case the set of atom-centered point charges is augmented by additional off-center charge sites. Good illustrations of this approach are molecular models of water. An OCC is present in the TIP4P-type^{29–31} and OPC³²

water models, where the point charge of the oxygen is moved toward the hydrogens along the bisector of the HOH angle. A combination of an OCC and an EC is used in the early ST2 model³³ and the more recent TIPSP-type models.^{34,35} Here, the two lone-pairs of the oxygen are represented by point charges, while again no charge is located at the center of the oxygen atom. Finally, the very recent TIPSP/2018 model³⁶ makes use of five point charges. Three of them are located at the atoms' centers, while two ECs represent the lone-pairs.

In studies of other molecules ECs were introduced to enhance the description of anisotropies around nitrogen,^{37–46} oxygen,^{40,41,47–50} sulfur,^{41,51,52} and halogen atoms.^{46,51,53,54} Also, the ethyl anion was modeled with an EC at the lone-pair site,^{55,56} and recently methods were devised to systematically place ECs around atoms exceeding a given anisotropy threshold.^{57,58} Another study proposed an approach where all charges are allowed to be placed off-center, which can enhance the description of the ESP even if the number of charges is smaller than the number of atoms.⁵⁹ Finally, ECs have also been introduced in the drude polarizable force-field^{60,61} to describe lone-pairs. However, the discussion of polarizabilities, despite their importance,^{26,62–65} is beyond the scope of this article.

Received: March 2, 2020

Published: May 19, 2020



The majority of the above-mentioned studies rely either on chemical intuition or fitting procedures to place the ECs. In this work a different approach is employed. By systematically scanning the surroundings of nitrogen atoms, areas where a single EC significantly enhances the approximation of the molecular ESP with point charges can be located. In this way it is possible to suggest regions for EC placement that might be missed if one relies too heavily on chemical intuition. Specifically, a set of 13 small organic molecules each containing a single nitrogen is investigated. This choice is motivated by a previous study of some of the authors,²⁸ in which nitrogens were identified as those atoms for which a plain atom centered point charge model exhibits the largest deficiencies. Furthermore, the molecules studied contain functional groups that are relevant for amino acids, peptides, and hence biomolecular systems as well as polymer materials.

METHODS

Figure 1 shows the 13 nitrogen containing molecules considered in this work. All molecules were studied in their neutral form, i.e., protonated or deprotonated variants were not considered. For each molecule, the impact of a single EC on the ESP of the point charge model is investigated. In particular it is investigated how the point charge ESP performs with respect to the ESP obtained from quantum mechanical (QM) calculations. The applied workflow is as follows. (i) Similar to previous work,^{28,66} the molecular structure was optimized in the gas-phase using QM methods and the corresponding QM-ESP was extracted. The latter was calculated on a 3D grid within a cuboid that extends 5 Å from the molecule in each direction. A 0.3 Å grid spacing was used, and the ESP calculation was performed at the same QM level as the structure optimizations. Please note that the thus extracted ESP is not an electrostatic surface potential as for instance obtained through the Merz–Kollman framework, but an electrostatic grid potential. To be consistent with previous work,^{28,66} the abbreviation ESP will be used throughout this article. Polarization effects that would occur in the condensed phase were not considered. (ii) Depending on the chemical nature of the nitrogen a position for the EC relative to the nitrogen's coordinates was selected (further details on the EC placement will be given below). (iii) All atomic point charges and, if present, the EC were fitted to minimize the error in the ESP using the least-squares method following previously described procedures.^{28,66,67} Since the most relevant electrostatic interactions take place in the so-called first interaction belt (1.66–2.20 r_{vdW}),^{67,68} the fit was restricted to this region. Steps ii and iii were repeated for several EC positions to scan the surroundings of the nitrogen atoms, thus charting regions where the placement of an EC is most beneficial. Additionally, the plain atomic point charge (APC) model, i.e., the system without an EC, was studied as reference.

The position of the EC was calculated relative to the nitrogen's coordinates in all cases. Here, two different routines were used to generate the location of the EC. For ammonia and amines with three identical substituents, a 1D scan was performed, while a 2D scan was performed for amines with two identical substituents and for amides. Amines with three different substituents were not considered in this study.

For the 1D scan, the EC was placed on the (pseudo)- C_3 axis of the R_3N molecule as illustrated in Figure 2A. Note that R can also represent a hydrogen atom and that the addition of “pseudo” indicates that for molecules other than NH_3 this axis

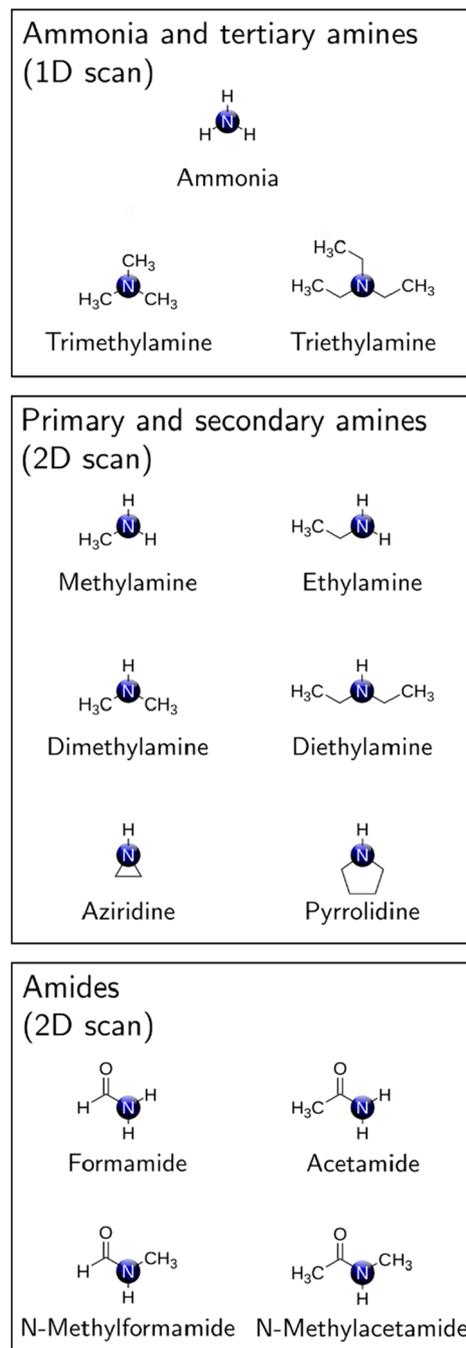


Figure 1. Set of 13 molecules investigated in this study. For ammonia and the tertiary amines, a 1D scan for optimal EC placement was performed. For the primary and secondary amines, as well as the amides, a 2D scan for optimal EC placement was performed.

is strictly speaking not a C_3 axis. The distance r from the nitrogen atom was varied in increments of 0.1 Å up to a maximum of 3 Å in both directions. Positive values for r indicate the direction opposing the substituents (i.e., up in Figure 2A), while negative values for r indicate the direction toward the substituents (i.e., down in Figure 2A).

In case the nitrogen atom is part of an amide moiety, a 2D scan was performed. Here, the EC was placed in the plane of the amide group as is illustrated in Figure 2B. A similar approach was used for the primary and secondary amines, where the nitrogen is bound to two identical and one differing

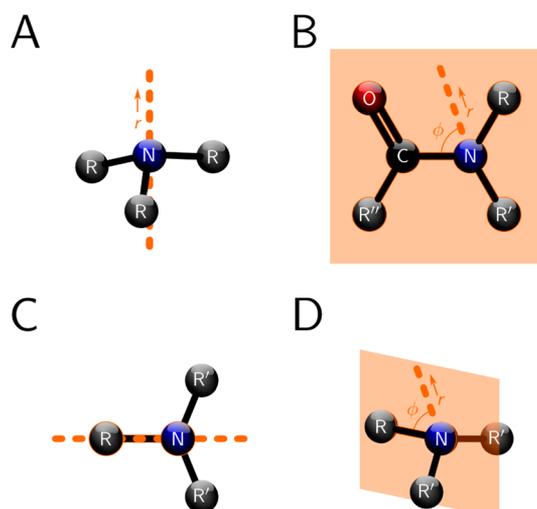


Figure 2. Illustration of the EC placement procedure. For R_3N molecules, the EC was placed on the (pseudo)- C_3 axis, varying the distance r from the nitrogen atom (A). For amides, the EC was placed in the plane of the amide moiety varying the distance r from the nitrogen atom and the angle ϕ (B). For RR'_2N molecules, the EC was placed in the mirror plane with respect to the first neighboring atoms of the nitrogen atom – top view (C) and 3D view (D). Also, here, the distance r from the nitrogen atom and the angle ϕ were varied. Note that R , R' , and R'' can also represent hydrogen atoms.

substituents (i.e., RR'_2N). That is, the EC was placed in a plane that is oriented in such a way that it contains the first atom of R , the nitrogen atom, and the bisector of the $R'NR'$ angle. This corresponds to a mirror plane, if only the nitrogen and its immediate neighbors are considered. This construction is illustrated in Figure 2C,D. In both two-dimensional scans the distance r from the nitrogen atom was varied in steps of 0.1 Å to a maximum of 2.1 Å, while the angle ϕ was varied in steps of 0.1 rad ($\sim 5.7^\circ$).

The plane used for the 2D scans was chosen in such a way that a rather simple scan (2D instead of 3D, one EC instead of multiple ECs) can capture the most promising locations for the placement of the EC. For amides, the plane of the amide

moiety was the obvious choice. The mirror plane constructed for the amines is also reasonable, since a placement of ECs outside this plane would require the use of two ECs due to the local symmetry. Moreover, this plane also includes the intuitive lone-pair location where ECs were often placed in other studies,^{37,38,41–46} thus enabling a direct comparison.

For the selection of the QM level of theory necessary for step i, combinations of three QM methods, being HF, B3LYP, and CCSD, and four different basis sets, namely 6-31G*, 6-311++G(3df,3pd),^{72–74} aug-cc-pVDZ,^{75,76} and aug-cc-pVTZ,^{75,76} were compared for ammonia. For the remaining molecules, the QM calculations were performed at the B3LYP/6-311++G(3df,3pd) level. All QM calculations were performed using Gaussian 09.⁷⁷

RESULTS

QM-ESP. As a first step the influence of the level of theory on the QM-ESP of a single ammonia molecule was tested. In particular, the accuracy of the QM calculation was benchmarked until an increase in the QM level did not significantly alter the QM-ESP. Note that the structure optimization and the QM-ESP calculations were always performed at the same level. Figure 3A reports the root-mean-square error (RMSE) within the first interaction belt between the different QM-ESPs. It can be seen that all calculations employing a large basis-set, i.e., not 6-31G*, yield highly similar results. That is, they have a small RMSE when compared to each other. This is consistent with previous work where only small differences in the QM-ESP were found between M06-2X/aug-cc-pVDZ and CCSD(T)/aug-cc-pVTZ calculations.²⁸

In addition to the RMSE between the different QM-ESPs, differences between the optimized geometry of ammonia were evaluated. For this purpose, the root-mean-square deviation (RMSD) of the different geometries was calculated (see Figure 3B). Here, slightly more variation in the data is found. Most notably the HF optimized geometries deviate markedly from those obtained using other levels of theory. Please note the very small difference between B3LYP/aug-cc-pVTZ and B3LYP/6-311++G(3df,3pd), implying that the latter provides an adequate compromise between accuracy and computational

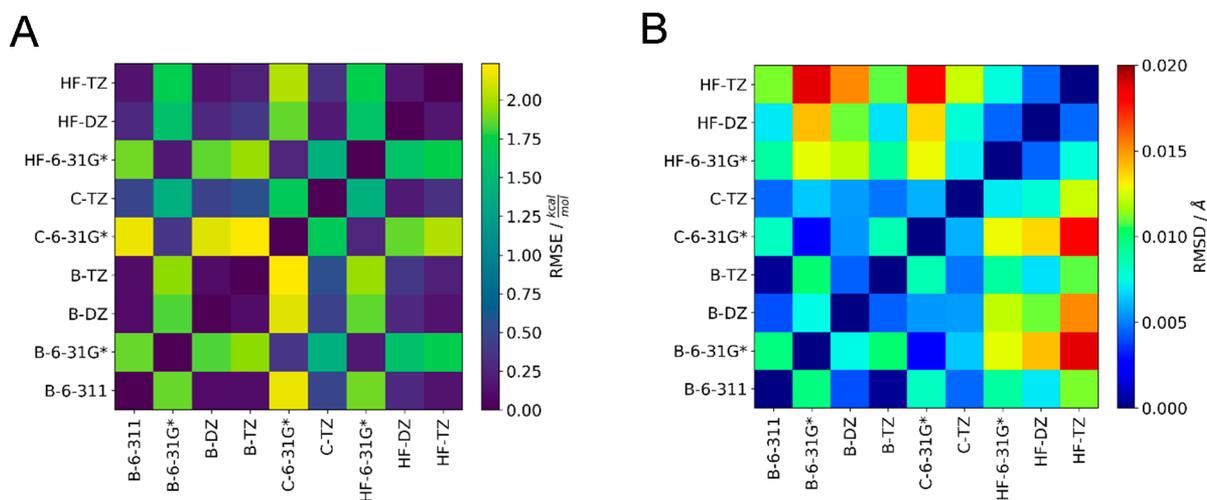


Figure 3. RMSE between QM-ESPs calculated at different levels of theory for a single ammonia molecule (A) and the RMSD between the corresponding minimum structures (B). For the labeling of the data, the following abbreviations were used: C, CCSD; B, B3LYP; DZ, aug-cc-pVDZ; TZ, aug-cc-pVTZ; and 6-311, 6-311++G(3df,3pd).

effort. Hence, B3LYP/6-311++G(3df,3pd) was used for the remainder of this work.

Ammonia and Tertiary Amines (1D Scans). For ammonia, trimethylamine, and triethylamine, the (pseudo)- C_3 axis for optimal EC placement was scanned. The results obtained for ammonia are depicted in Figure 4. Here, the

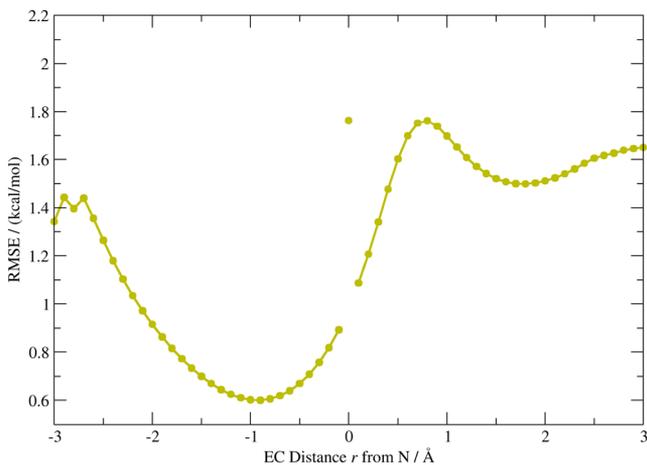


Figure 4. RMSE of the 1D scan for ammonia. At $r = 0$ the RMSE corresponds to the plain atomic point charge (APC) reference. The results are based on B3LYP/6-311++G(3df,3pd) calculations.

RMSE of the point charge ESP with respect to the QM reference (QM-ESP) is shown as a function of N-EC distance r . The RMSE value at $r = 0$ corresponds to the APC reference. Note that the EC cases ($r \neq 0$) have one point charge more than the APC reference ($r = 0$). Because of this additional parameter, all EC cases yield lower RMSE values than the APC reference. However, the location of the EC has a strong influence on how much the RMSE is decreased. As r is varied, two distinct minima are found at -0.9 and $+1.8$ Å, respectively. The minimum at negative distances is lower, reducing the RMSE by 66% when compared to the APC reference. The corresponding charge of the EC q_{EC} is -0.188 e.

In Figure 5 the results for trimethylamine are shown. Again two minima appear, a broad one at -1.4 Å and a more

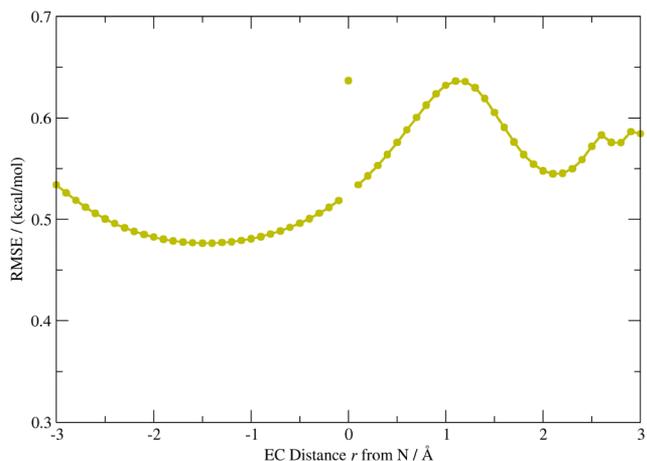


Figure 5. RMSE of the 1D scan for trimethylamine. At $r = 0$ the RMSE corresponds to the plain atomic point charge (APC) reference. The results are based on B3LYP/6-311++G(3df,3pd) calculations.

localized one at 2.1 Å. Here, also the minimum at negative distances is the lower one. The corresponding RMSE reduction with respect to the APC reference is 25% and the corresponding q_{EC} is -0.048 e. This reduction in RMSE is small when compared to ammonia. However, it has to be noted that the RMSE of the APC reference is already <0.7 kcal/mol, a value that was obtained for ammonia only after the introduction of an EC (cf. Figure 4).

The last molecule for which a 1D scan was performed is triethylamine. Again two minima in the RMSE are observed (see Figure 6). One minimum is located at -0.8 Å and one at

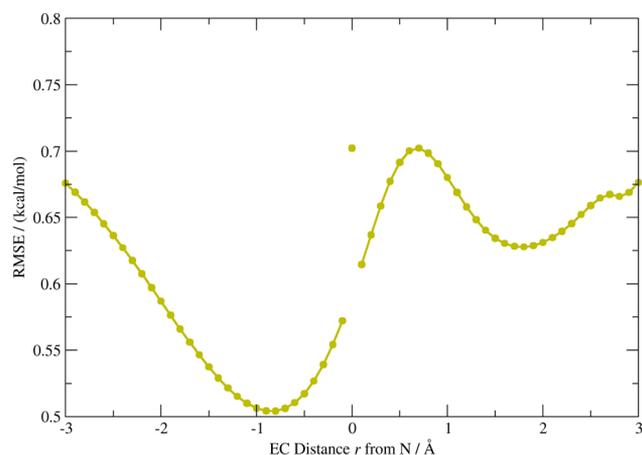


Figure 6. RMSE of the 1D scan for triethylamine. At $r = 0$ the RMSE corresponds to the plain atomic point charge (APC) reference. The results are based on B3LYP/6-311++G(3df,3pd) calculations.

$+1.8$ Å. Similar to ammonia and trimethylamine the minimum at negative r yields a lower RMSE value. The corresponding reduction in the RMSE is 28% compared to the already low APC RMSE (≈ 0.7 kcal/mol), and the corresponding q_{EC} is -0.433 e.

For all three molecules, for which the EC was placed along the (pseudo)- C_3 axis, the charge fitting procedure based on the B3LYP/6-311++G(3df,3pd) QM-ESP suggests the placement of an EC carrying a negative partial charge at negative distances, i.e., opposite of the lone-pair location and toward the substituents. Despite the fact that the nitrogen still carries a charge (positive for ammonia, negative for trimethyl- and triethylamine), this finding is similar to the placement of the OCC for the TIP4P-type water models,^{29–31} among which arguably the best rigid water models are found.⁷⁸ For those models, the charge of the oxygen is moved toward the hydrogen atoms along the HOH bisector. Even at distances comparable to the O-OCC distance of TIP4P-type models (0.1 – 0.2 Å),^{29–31} the EC placement at negative distances is favored over the EC placement at positive distances for the three molecules studied here.

Concluding this section, we bring the Supporting Information (SI) to the reader's attention, where we discuss the influence of different QM calculations on the optimal EC location (section S-I) as well as the change in all point charges as the EC location is varied (section S-II).

Primary and Secondary Amines (2D Scans). In the case of methyl- and ethylamine, 2D scans for the EC placement were performed in the plane indicated in Figure 2C,D. The results of the two-dimensional scans are shown as polar heat maps in Figure 7, where the radial distance r and the angle ϕ

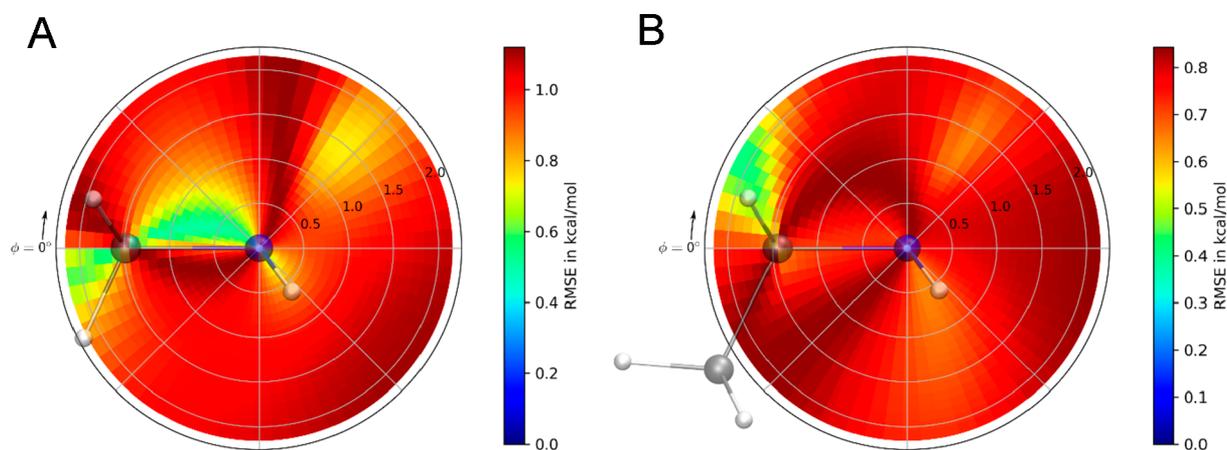


Figure 7. RMSE of the 2D scans for methylamine (A) and ethylamine (B). Here, the C–N bond is located in the scanned plane (cf. Figure 2C,D). The coloring indicates the RMSE and the maximum of the color scale is the RMSE of the plain atomic point charge (APC) reference. The results are based on B3LYP/6-311++G(3df,3pd) calculations.

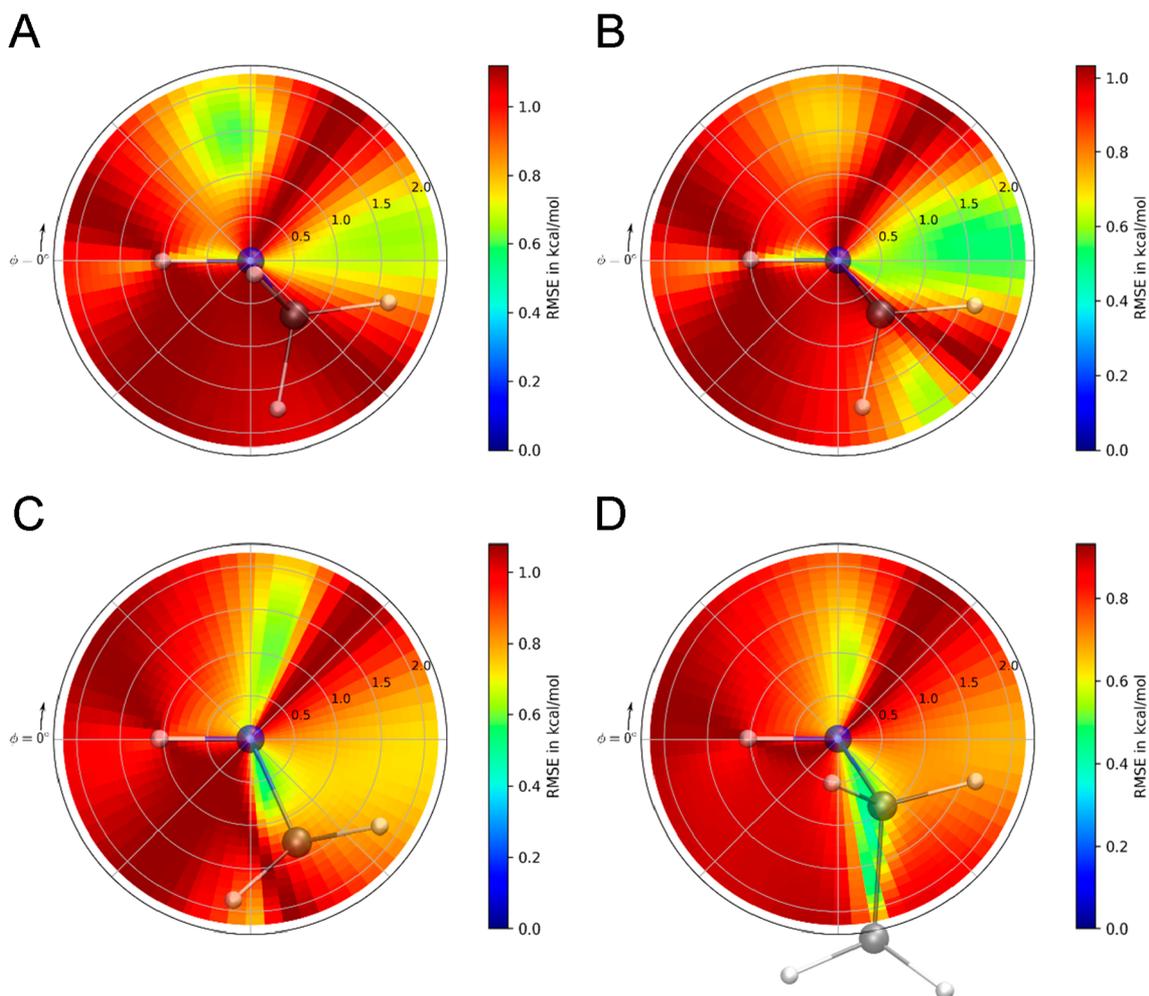


Figure 8. RMSE of the 2D scans for dimethylamine (A), diethylamine (B), aziridine (C), and pyrrolidine (D). Here, the H–N bond is located in the scanned plane (cf. Figure 2C,D). Note that for clarity the CH_3 group is not shown in the case of diethylamine (B). The coloring indicates the RMSE, and the maximum of the color scale is the RMSE of the plain atomic point charge (APC) reference. The results are based on B3LYP/6-311++G(3df,3pd) calculations.

specify the scanned EC position while the coloring indicates the RMSE of the respective point charge ESP when compared to the QM-ESP reference. In these plots the in-plane bond (corresponding to $\phi = 0$) is always oriented to the left of the

nitrogen atom. The maximum of the color scale corresponds to the RMSE of the APC reference. Due to the additional fitting parameter the EC case yields better results than the APC reference for any EC location. While there are some similarities

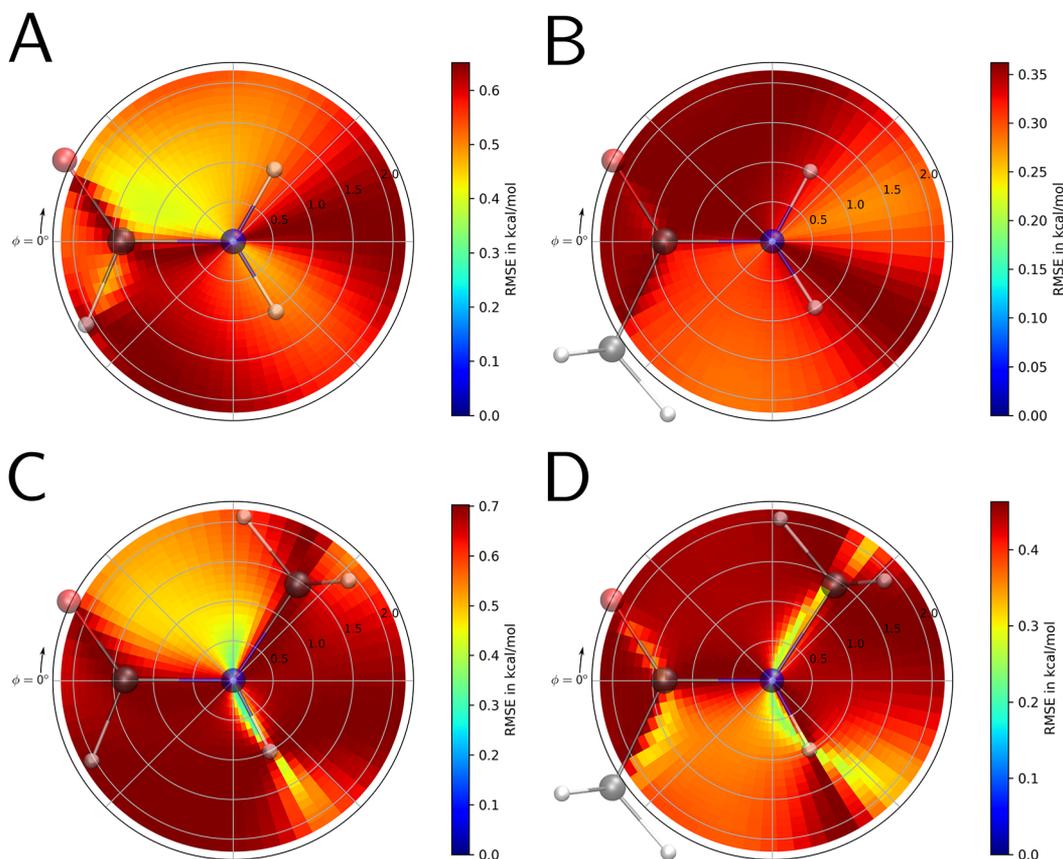


Figure 9. RMSE of the 2D scans for formamide (A), acetamide (B), *N*-methylformamide (C), and *N*-methylacetamide (D). Here, the whole amide moiety is located in the scanned plane (cf. Figure 2B). The coloring indicates the RMSE, and the maximum of the color scale is the RMSE of the plain atomic point charge (APC) reference. The results are based on B3LYP/6-311++G(3df,3pd) calculations.

between the two molecules, also remarkable differences can be noted. The most prominent minimum for EC placement in case of methylamine is above the C–N bond (see Figure 7A). In addition, more shallow minima are found in the direction of the two hydrogen atoms bound to the nitrogen and beyond 1 Å in the approximate lone-pair direction. However, this latter minimum is located at $\approx 135^\circ$ with respect to the C–N bond and not at the ideal tetrahedral angle (109.5°). In addition, a minimum beyond the CH_3 group is found. As is shown in the SI (Figure S6), the EC in the minimum above the bond carries a positive charge, while for all other minima negative q_{EC} are found. In Figure S6 it is also visible that the magnitude of the EC increases if the centers of the atoms are approached. This is the expected behavior, since the reproduction of identical multipole moments with smaller charge separation demands larger charges.

For ethylamine, the minima near the C–N bond and in the lone-pair direction are still present (see Figure 7B). Anyhow, these minima are markedly shallower when compared to the results for methylamine. The minimum in the approximate lone-pair direction is shifted to slightly lower angles, but it is still located at angles $>109.5^\circ$. It appears that for ethylamine the most favorable placement of a single EC is close to the edge of the scanned area. This minimum is located at distances beyond 1.5 Å from the nitrogen, where the q_{EC} is negative (see Figure S6). ECs around the C–N bond are positive, and areas to the right of the nitrogen are best fit by negative ECs.

For both primary amines, the optimal EC location reduces the RMSE by $\approx 50\%$ when compared to the APC reference.

For methylamine, this optimal location is very close to the nitrogen, while in the case of ethylamine it is found beyond a distance of 1.5 Å and close to the first carbon atom. EC placement closer to the nitrogen atom of ethylamine can reduce the RMSE by $\approx 20\%$, although the respective absolute RMSE change is comparable in size to the differences between B3LYP/6-311++G(3df,3pd) and higher level QM calculations for ammonia (cf. Figure 3A).

In Figure 8 results for dimethylamine, diethylamine, aziridine, and pyrrolidine are presented. These four secondary amines are similar, since each nitrogen has two identical substituents. The substituents are aliphatic chains containing either one or two carbon atoms. In the case of aziridine and pyrrolidine the two substituents form a ring. For dimethylamine, three pronounced minima appear (see Figure 8A). One for the placement of an EC close to the H–N bond, one in its extension beyond the nitrogen, and one above the nitrogen, slightly tilted toward the hydrogen. Remarkably, it is of little use to place an EC in the lone-pair direction, what chemical intuition would suggest. For the minima around and above the H–N bond, negative values for q_{EC} are obtained, while in the other areas q_{EC} is positive (see Figure S7). In the lone-pair direction q_{EC} is almost zero indicating again that this direction is not efficient for EC placement.

If the substituent length is increased by one carbon atom (diethylamine) similar results are found (see Figure 8B). The minima around the H–N bond and its extension are still present and both minima are more pronounced when compared to the results for methylamine. In addition, the

minimum above the nitrogen is far less developed, but a new minimum appears in the direction of the CH_2 groups. This feature is similar to what is found for ethylamine and it again is found at the edge of the scanned area. In all cases no minimum in the direction of the lone-pair can be observed. Figure S7 shows the corresponding q_{EC} data. The EC is negative above and below the H–N bond, and it is positive in all other areas, a very similar result as in the case of dimethylamine.

If instead of adding a carbon atom the two substituents are linked by a bond, i.e., when moving from dimethylamine (Figure 8A) to aziridine (Figure 8C), different results are obtained. Rather than three minima, as found for methylamine, only two distinct minima appear. The first is pointing to the ring center, while the second is located above the nitrogen atom. Interestingly, the latter minimum is at an angle of $\leq 109.5^\circ$ to the H–N bond. That is, it is located in an area where typically a lone-pair would be expected. Again, negative values for q_{EC} are found in the area around and above the H–N bond, while positive values for q_{EC} are observed in the minimum located within the ring moiety (see Figure S7).

The last molecule in this series is pyrrolidine. Here, the nitrogen atom is part of a five-membered heterocycle. One can also think of pyrrolidine as a derivative of diethylamine, in which the two terminal carbon atoms are linked by a bond. In Figure 8D it is visible that the results for pyrrolidine are very similar to aziridine. Two minima are present: One above the nitrogen and one in the ring center. In both cases, the latter minimum corresponds to lower RMSE values. The smaller minimum is located almost perpendicular to the H–N bond. That is, this minimum is not located close to the intuitive lone-pair site, as it was found for aziridine (cf. Figure 8C), but rotated toward the H–N bond. For pyrrolidine, the q_{EC} data is quite similar to the aziridine case. In the area around and above the H–N bond q_{EC} is negative, while it is positive below the H–N bond (see Figure S4D).

For all four secondary amines studied, the RMSE can be reduced by 50% with respect to the APC reference if the EC is placed in the optimal spot. While this minimum is located at $r > 1.0 \text{ \AA}$ for dimethylamine, it is located right next to the nitrogen for diethylamine, aziridine, and pyrrolidine.

Amides (2D Scans). In Figure 9 the data for the four studied amides are presented. Here, 2D scans for EC placement were performed in the plane indicated in Figure 2B. The results for formamide shown in Figure 9A indicate shallow minima around the downward pointing H–N bond and in the direction of the H–C bond. The third and lowest minimum is located above the C–N bond. It seems this minimum bridges the gap between the nitrogen and the oxygen, which apparently is the location where the anisotropy of the entire moiety is approximated best by the addition of a single EC. In this latter minimum q_{EC} is positive, while in all regions below the C–N bond q_{EC} is found to be negative (see Figure S8).

The next two amides investigated are acetamide and *N*-methylformamide. In each case a CH_3 group is added to formamide but on different sides of the molecule. It is visible in Figure 9B that the ESP of acetamide cannot be enhanced significantly through the introduction of an EC. Two areas for EC placement are visible that enable some enhancement. One such area is located between the nitrogen atom and the CH_3 group, but this minimum is very broad and shallow. Similarly, a broad and shallow minimum is found on the far side of the nitrogen atom, in the direction of the hydrogen atoms. This

area is tilted slightly upward. If instead a CH_3 group is added to the nitrogen atom (*N*-methylformamide), considerable improvement is possible (see Figure 9C). Distinct minima are found if the EC is introduced above the nitrogen or close to the N–H bond. For *N*-methylformamide, also two rotamers were considered, which have a different orientation of the methyl group. It is found that the area around the N–H bond is hardly affected by the rotation of the methyl group, while the area above the C(carbonyl)–N bond changes (see section S-II in the SI for a detailed discussion).

For both acetamide and *N*-methylformamide, the corresponding q_{EC} data are shown in Figures S8 and S9, respectively. In each case the EC is positive above the C(carbonyl)–N bond and negative below. For *N*-methylformamide, q_{EC} is also negative to the right of the N–C(methyl) bond. Additionally, the magnitude of q_{EC} is significantly larger for *N*-methylformamide, especially near the N–H bond.

The last amide studied is *N*-methylacetamide, representing the proxy for a peptide chain in this survey. Similarly to *N*-methylformamide, pronounced minima in the RMSE are present (see Figure 9D). A shallow minimum extends downward from the carbonyl carbon and clear minima are located along the C(methyl)–N–H chain. These connected minima also extend beyond the methyl carbon and the hydrogen. The respective q_{EC} are shown in Figure S10. Positive q_{EC} are found above the C(carbonyl)–N bond and to the right of the N–H bond. Negative q_{EC} are found below the C(carbonyl)–N bond and beyond the CH_3 group bound to the nitrogen.

The introduction of an EC in *N*-methylformamide or *N*-methylacetamide can improve the RMSE by $\approx 50\%$. While for formamide a reduction of $\approx 30\%$ is possible, the introduction of an EC in acetamide yields an RMSE improvement $< 20\%$ with respect to the APC reference. In any case, when compared to the amines (cf. Figures 7 and 8) the RMSE of the APC reference is already quite low for all four amides studied ($< 1 \text{ kcal/mol}$). Especially in case of acetamide, the RMSE of the APC reference is below 0.4 kcal/mol (cf. Figure 9B) and can be considered acceptable as is.

DISCUSSION

Based on the QM-ESP calculated at B3LYP/6-311++G-(3df,3pd) level, a set of 13 small organic molecules, each containing one nitrogen atom, was studied. The surroundings of the nitrogen atoms were systematically scanned for locations, where the placement of an extra-charge (EC) significantly enhances the point charge approximation of the reference QM-ESP. Depending on the local symmetry of the nitrogen atoms, scans along a single axis or within a plane were performed to identify such locations. Note that the location of the EC has almost no impact on the dipole moment of the studied molecules, but it significantly influences the quadrupole moment.

For ammonia, the introduction of a single EC can reduce the RMSE by $\approx 60\%$ compared to the plain atom centered point charge (APC) reference. For the studied amines, the introduction of an EC leads to a reduction of the RMSE by $\approx 50\%$. Exceptions are trimethyl- and triethylamine, where the RMSE is reduced by $\approx 35\%$. Mixed results are obtained for the four amides. While the RMSE of acetamide cannot be improved appreciably by the addition of an EC ($< 20\%$), some improvement is possible for formamide ($\approx 30\%$) and

significant improvements are possible for both *N*-methylformamide and *N*-methylacetamide ($\approx 50\%$). It was also found that the RMSE of the APC reference of all studied amides is already far below the RMSE of the APC of the amines indicating no immediate need for the introduction of an extra charge in these systems. Moreover, the changes in RMSE are comparable to the differences between B3LYP/6-311++G-(3df,3pd) and higher level QM calculations for ammonia. Nevertheless, it is striking that the addition of a single EC in *N*-methylacetamide (the peptide proxy), increasing the number of charges from 12 to 13, reduces the RMSE by 50%. This is a very efficient improvement that could be exploited to enhance the description of protein backbones.

Based on 1D scans of ammonia, trimethyl-, and triethylamine, the optimal spot for the EC is located opposite of the lone-pair direction. That is, the optimal EC placement is in the direction of the substituent along the (pseudo)- C_3 axis. In addition, the EC in this region is always negative. The placement of a negative charge opposite of the lone-pairs is an intriguing similarity to well performing molecular models of water (TIP4P-type^{29–31} and OPC³²) that all include an off-center charge (OCC) for the oxygen atom located at the HOH bisector close to the oxygen atom. Interestingly, it was shown⁷⁸ that these water models perform significantly better than the TIP5P water model,³⁴ in which two charges represent the lone-pairs of the oxygen.

2D scans were performed for primary and secondary amines as well as for amides. The data for the primary amines show that for methyl- and ethylamine a minimum is found somewhat close to the intuitive lone-pair position. However, a closer look reveals that these minima are located between 120 and 135° , which is significantly larger than the ideal tetrahedral angle (109.5°). Moreover, these minima are not the lowest minima, which are located above the C–N bond for methylamine and beyond the CH_2 group for ethylamine. The secondary amines dimethyl- and diethylamine show no minima close to the lone-pair position. For the corresponding cyclic molecules, aziridine and pyrrolidine, the lone-pair spot is very close to the minimum for aziridine, but not for pyrrolidine. The most pronounced minimum, however, is found inside the ring moiety for both molecules.

In the case of formamide the introduction of an EC leads to an improvement of the point-charge approximation of the QM-ESP if the EC is placed between the N and the O of the amide moiety. For acetamide, no clear minimum is observed, while for *N*-methylformamide and *N*-methylacetamide distinct minima are located close to the N–H bond and above the nitrogen.

Already in past studies, lone-pairs were explicitly described through point charges to enhance the description of nitrogen atoms. A particular focus was laid on heterocyclic aromatic molecules.^{37,38,40–46,60} Here, the lone-pair position was fixed in the aromatic plane pointing outward from the nitrogen, along the bisector of the CNC angle. The lone-pair was either described as an OCC^{43–46,60} or as an EC.^{37,38,40–42} Dixon and Kollmann⁴¹ studied also nonaromatic molecules containing nitrogen. In particular, they studied ammonia, methylamine, dimethylamine, and trimethylamine, in which lone-pair charges were added to the nitrogen atom. In their work the lone-pair EC is always located at the intuitive position (ideal tetrahedral angle of 109.5° and a distance of 0.35 \AA). While the scans performed in the present work do show minima in that direction for ammonia and trimethylamine, the optimal EC

location is found at the opposite side of the nitrogen, pointing toward the substituents. For methylamine, the lowest minimum is found close to the C–N bond and for dimethylamine the intuitive lone-pair location is actually a suboptimal spot to enhance the approximation of the QM-ESP with a single EC. Better locations are found along the H–N bond and its extension as well as above the nitrogen. Cole et al.⁵⁷ used a method based on dipole- and quadrupole moments to identify badly reproduced ESPs. On the thus identified atoms up to three ECs were introduced based on a minimization routine. In this study also amines were considered and two ECs have been introduced to the nitrogen atoms for both methylamine and dimethylamine. In a similar effort Horton et al.⁵⁸ introduced ECs if the error in the ESP exceeded a certain threshold. Here, the placement of the ECs is based on symmetry arguments. For ammonia, trimethylamine, and triethylamine, Horton et al.⁵⁸ actually use the (pseudo)- C_3 axis, as was done in the present work. However, for the primary and secondary amines ECs were placed on an axis that is at equal angles to all three bonds originating from the nitrogen atom. Instead, in the present study a 2D scan in the mirror plane containing the nitrogen and the bond to the unique substituent was performed. To compare the two approaches, the vectors as described by Horton et al.⁵⁸ were constructed. For the primary amines, this vector is at an angle of $\approx 109^\circ$ to all bonds, i.e., very close to the tetrahedral angle. For the noncyclic secondary amines, $\approx 108^\circ$ is found, for aziridine $\approx 120^\circ$ and for pyrrolidine $\approx 111^\circ$. Most of these angles are close to the ideal tetrahedral angle, and the corresponding vectors do not intersect the minima identified in the present survey (see Figures S11 and S12). Only for ethylamine ($\approx 109^\circ$) and pyrrolidine ($\approx 111^\circ$) this vector approaches an observed minimum. The method introduced by Horton et al.⁵⁸ also enables the addition of multiple ECs in case a single EC does not reproduce the ESP satisfactorily. In light of our results we therefore surmise that a slight adjustment in the way the first EC is placed might suffice to obtain an acceptable approximation of the QM-ESP rendering the addition of further ECs obsolete.

In summary, it was found that considerable improvement of the ESP approximation with point charges is possible if a single EC is added in the proximity of the nitrogen atom. It was also revealed that small chemical changes can alter the optimal spot for EC placement quite significantly. Nevertheless, also some trends were found. For ammonia and the tertiary amines, a negative EC at negative distances yields the best fit in all cases. For the primary amines, a negative EC along the C–N bond enhances the ESP, and for the secondary amines, a negative EC along the H–N bond improves the ESP. In these cases, however, this does not always correspond to the optimal placement, and also the amount of improvement varies considerably for these specific spots. For the two cyclic amines, we find that a positive EC located close to the ring center along the CNC bisector is the best EC placement. For the amides, finally, we find no consistent pattern. Only *N*-methylformamide and *N*-methylacetamide share similarities, as in both cases the placement of a negative EC close to the N–H bond is beneficial. Anyhow, our data do not suggest a one-size-fits-all rule, not even for similar molecules. On the contrary, a beneficial spot for EC placement in one molecule, e.g., an EC on the N–H bond of *N*-methylformamide, can actually be only marginally beneficial in a similar molecule (*N*-methylacetamide). For *N*-methylacetamide, the EC has to be placed at

$\approx 10^\circ$ with respect to the N–H bond to yield a comparable enhancement. Therefore, we refrain from suggesting generally applicable spots for beneficial EC placement.

CONCLUSION

In this work, areas around nitrogen atoms were systematically scanned for spots where EC placement is beneficial to reproduce a QM-based reference ESP. For most studied molecules, a single EC, usually with a magnitude <1.0 e, can significantly improve the description of the molecular ESP. Intriguingly, the optimal spots for EC placement do not follow chemical intuition. In fact, even some cases are revealed, where the placement of an EC is of little use if it is placed at the intuitive lone-pair spot. Despite some similarities between different molecules, these results do not allow the deduction of a general approach that would predict the optimal EC location for any nitrogen containing compound, as small chemical changes can alter the results significantly.

Obviously this survey of small organic molecules containing a single nitrogen just scratches the surface of possible efforts aimed at optimizing the point charge approximation of the molecular ESP. Future studies need to be directed at compounds with multiple and less symmetrically substituted nitrogen atoms. Within the scan approach presented here, the addition of a second nitrogen atom would require an additional 2D scan for each scanned point around the first nitrogen, effectively squaring the computational demand. Nevertheless, it is surely rewarding to also look into these cases, as they are relevant for a proper description of the amino acids histidine and arginine as well as the nucleobases. Ultimately, full force fields need to be built on top of a point charge model derived using the here presented scanning approach. Only in this way, it is possible to perform numerical simulations that enable a comparison with experimental data.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.0c00204>.

Additional data pertaining to the scans for optimal EC placement, including discussions on the influence of the QM level used and on the dependence of all point charges on the EC location and a comparison of our results to the results from literature (PDF)

AUTHOR INFORMATION

Corresponding Author

Klaus R. Liedl – Institute of General, Inorganic and Theoretical Chemistry, University of Innsbruck, A-6020 Innsbruck, Austria; orcid.org/0000-0002-0985-2299; Email: klaus.liedl@uibk.ac.at

Authors

Alexander Spinn – Institute of General, Inorganic and Theoretical Chemistry, University of Innsbruck, A-6020 Innsbruck, Austria; orcid.org/0000-0003-2527-5513

Philip H. Handle – Institute of General, Inorganic and Theoretical Chemistry, University of Innsbruck, A-6020 Innsbruck, Austria

Johannes Kraml – Institute of General, Inorganic and Theoretical Chemistry, University of Innsbruck, A-6020 Innsbruck, Austria

Thomas S. Hofer – Institute of General, Inorganic and Theoretical Chemistry, University of Innsbruck, A-6020 Innsbruck, Austria

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jctc.0c00204>

Author Contributions

[†]A.S. and P.H.H. contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Austrian Science Fund (FWF) via Grants P30565 and P30737.

REFERENCES

- (1) Frenkel, D.; Smit, B. *Understanding molecular simulation: from algorithms to applications*; Elsevier: Amsterdam, 2001; Vol. 1.
- (2) Leach, A. R. *Molecular modelling: principles and applications*; Pearson education: London, 2001.
- (3) Rovigatti, L.; Russo, J.; Romano, F. How to simulate patchy particles. *Eur. Phys. J. E: Soft Matter Biol. Phys.* **2018**, *41* (5), 59.
- (4) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713.
- (5) Debic, K. T.; Cerutti, D. S.; Baker, L. R.; Gronenborn, A. M.; Case, D. A.; Chong, L. T. Further along the road less traveled: AMBER ff15ipq, an original protein force field built on a self-consistent physical model. *J. Chem. Theory Comput.* **2016**, *12* (8), 3926–3947.
- (6) Wang, L.-P.; McKiernan, K. A.; Gomes, J.; Beauchamp, K. A.; Head-Gordon, T.; Rice, J. E.; Swope, W. C.; Martínez, T. J.; Pande, V. S. Building a more predictive protein force field: a systematic and reproducible route to AMBER-FB15. *J. Phys. Chem. B* **2017**, *121* (16), 4023–4039.
- (7) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evansck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102* (18), 3586–3616.
- (8) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D., Jr Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* **2012**, *52* (12), 3155–3168.
- (9) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14* (1), 71–73.
- (10) Poger, D.; Van Gunsteren, W. F.; Mark, A. E. A new force field for simulating phosphatidylcholine bilayers. *J. Comput. Chem.* **2010**, *31* (6), 1117–1125.
- (11) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and testing of the GROMOS force-field versions S4A7 and S4B7. *Eur. Biophys. J.* **2011**, *40* (7), 843.
- (12) Reif, M. M.; Hünenberger, P. H.; Oostenbrink, C. New interaction parameters for charged amino acid side chains in the GROMOS force field. *J. Chem. Theory Comput.* **2012**, *8* (10), 3705–3723.
- (13) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118* (45), 11225–11236.

- (14) Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. Improved peptide and protein torsional energetics with the OPLS-AA force field. *J. Chem. Theory Comput.* **2015**, *11* (7), 3499–3509.
- (15) Robertson, M. J.; Qian, Y.; Robinson, M. C.; Tirado-Rives, J.; Jorgensen, W. L. Development and Testing of the OPLS-AA/M Force Field for RNA. *J. Chem. Theory Comput.* **2019**, *15* (4), 2734–2742.
- (16) Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, *17* (5–6), 490–519.
- (17) Horta, B. A.; Merz, P. T.; Fuchs, P. F.; Dolenc, J.; Riniker, S.; Hünenberger, P. H. A GROMOS-compatible force field for small organic molecules in the condensed phase: The 2016H66 parameter set. *J. Chem. Theory Comput.* **2016**, *12* (8), 3825–3850.
- (18) Rein, R. On Physical Properties and Interactions of Polyatomic Molecules: With Application to Molecular Recognition in Biology. In *Advances in Quantum Chemistry*; Löwdin, P.-O., Ed.; Academic Press: New York, 1973; Vol. 7, pp 335–396.
- (19) Taylor, R.; Kennard, O.; Versichel, W. Geometry of the imino-carbonyl (N-H...O:C) hydrogen bond. 1. Lone-pair directionality. *J. Am. Chem. Soc.* **1983**, *105* (18), 5761–5766.
- (20) Taylor, R.; Kennard, O. Hydrogen-bond geometry in organic crystals. *Acc. Chem. Res.* **1984**, *17* (9), 320–326.
- (21) Platts, J. A.; Howard, S. T.; Bracke, B. R. F. Directionality of Hydrogen Bonds to Sulfur and Oxygen. *J. Am. Chem. Soc.* **1996**, *118* (11), 2726–2733.
- (22) Stone, A. J.; Price, S. L. Some new ideas in the theory of intermolecular forces: anisotropic atom-atom potentials. *J. Phys. Chem.* **1988**, *92* (12), 3325–3335.
- (23) Clark, T.; Hennemann, M.; Murray, J. S.; Politzer, P. Halogen bonding: the σ -hole. *J. Mol. Model.* **2007**, *13* (2), 291–296.
- (24) Metrangola, P.; Meyer, F.; Pilati, T.; Resnati, G.; Terraneo, G. Halogen bonding in supramolecular chemistry. *Angew. Chem., Int. Ed.* **2008**, *47* (33), 6114–6127.
- (25) Ponder, J. W.; Case, D. A. Force Fields for Protein Simulations. In *Advances in Protein Chemistry*; Academic Press: New York, 2003; Vol. 66, pp 27–85.
- (26) Mackerell, A. D. Empirical force fields for biological macromolecules: Overview and issues. *J. Comput. Chem.* **2004**, *25* (13), 1584–1604.
- (27) Stone, A. J. Distributed Multipole Analysis: Stability for Large Basis Sets. *J. Chem. Theory Comput.* **2005**, *1* (6), 1128–1132.
- (28) Kramer, C.; Spinn, A.; Liedl, K. R. Charge Anisotropy: Where Atomic Multipoles Matter Most. *J. Chem. Theory Comput.* **2014**, *10* (10), 4488–4496.
- (29) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.
- (30) Horn, H. W.; Swope, W. C.; Pitner, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120* (20), 9665–9678.
- (31) Abascal, J. L.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123* (23), 234505.
- (32) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **2014**, *5* (21), 3863–3871.
- (33) Stillinger, F. H.; Rahman, A. Improved simulation of liquid water by molecular dynamics. *J. Chem. Phys.* **1974**, *60* (4), 1545–1557.
- (34) Mahoney, M. W.; Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **2000**, *112* (20), 8910–8922.
- (35) Rick, S. W. A reoptimization of the five-site water potential (TIPSP) for use with Ewald sums. *J. Chem. Phys.* **2004**, *120* (13), 6085–6093.
- (36) Khalak, Y.; Baumeier, B.; Karttunen, M. Improved general-purpose five-point model for water: TIPSP/2018. *J. Chem. Phys.* **2018**, *149* (22), 224507.
- (37) Williams, D. E.; Weller, R. R. Lone-pair electronic effects on the calculated ab initio SCF-MO electric potential and the crystal structures of azabenzene. *J. Am. Chem. Soc.* **1983**, *105* (13), 4143–4148.
- (38) Williams, D. E.; Cox, S. R. Nonbonded potentials for azahydrocarbons: the importance of the Coulombic interaction. *Acta Crystallogr., Sect. B: Struct. Sci.* **1984**, *40* (4), 404–417.
- (39) Profeta, S., Jr.; Allinger, N. Molecular mechanics calculations on aliphatic amines. *J. Am. Chem. Soc.* **1985**, *107* (7), 1907–1918.
- (40) Kumar, A.; Mohan, C.; Mishra, P. Hybridization displacement charge in molecules and its effects on electrostatic potentials and fields. *Int. J. Quantum Chem.* **1995**, *55* (1), 53–60.
- (41) Dixon, R. W.; Kollman, P. A. Advancing beyond the atom-centered model in additive and nonadditive molecular mechanics. *J. Comput. Chem.* **1997**, *18* (13), 1632–1646.
- (42) Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Luyuan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A. others, OPLS3: a force field providing broad coverage of drug-like small molecules and proteins. *J. Chem. Theory Comput.* **2016**, *12* (1), 281–296.
- (43) Macchiagodena, M.; Mancini, G.; Pagliai, M.; Del Frate, G.; Barone, V. Fine-tuning of atomic point charges: Classical simulations of pyridine in different environments. *Chem. Phys. Lett.* **2017**, *677*, 120–126.
- (44) Pagliai, M.; Mancini, G.; Carnimeo, I.; De Mitri, N.; Barone, V. Electronic absorption spectra of pyridine and nicotine in aqueous solution with a combined molecular dynamics and polarizable QM/MM approach. *J. Comput. Chem.* **2017**, *38* (6), 319–335.
- (45) Pagliai, M.; Funghi, G.; Vassetti, D.; Procacci, P.; Chelli, R.; Cardini, G. Imidazole in Aqueous Solution: Hydrogen Bond Interactions and Structural Reorganization with Concentration. *J. Phys. Chem. B* **2019**, *123* (18), 4055–4064.
- (46) Roos, K.; Wu, C.; Damm, W.; Reboul, M.; Stevenson, J. M.; Lu, C.; Dahlgren, M. K.; Mondal, S.; Chen, W.; Wang, L.; Abel, R.; Friesner, R. A.; Harder, E. D. OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *J. Chem. Theory Comput.* **2019**, *15* (3), 1863–1874.
- (47) Allinger, N.; Chung, D. Y. Conformational analysis. 118. Application of the molecular-mechanics method to alcohols and ethers. *J. Am. Chem. Soc.* **1976**, *98* (22), 6798–6803.
- (48) Allinger, N. L.; Chang, S. H.-M.; Glaser, D. H.; Hönl, H. An improved molecular mechanics force field for alcohols and ethers. *Isr. J. Chem.* **1980**, *20* (1–2), 51–56.
- (49) Singh, U. C.; Kollman, P. A. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **1984**, *5* (2), 129–145.
- (50) Macchiagodena, M.; Mancini, G.; Pagliai, M.; Barone, V. Accurate prediction of bulk properties in hydrogen bonded liquids: amides as case studies. *Phys. Chem. Chem. Phys.* **2016**, *18* (36), 25342–25354.
- (51) Singh, U. C.; Kollman, P. A. Ab initio calculations on the structure and nature of the hydrogen bonded complex H₂S... HF. *J. Chem. Phys.* **1984**, *80* (1), 353–355.
- (52) Yan, X. C.; Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. Improved Description of Sulfur Charge Anisotropy in OPLS Force Fields: Model Development and Parameterization. *J. Phys. Chem. B* **2017**, *121* (27), 6626–6636.
- (53) 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* (10), 3895–3901.
- (54) Soteras Gutierrez, I.; Lin, F.-Y.; Vanommeslaeghe, K.; Lemkul, J. A.; Armacost, K. A.; Brooks, C. L.; MacKerell, A. D. Parametrization of halogen bonds in the CHARMM general force field: improved treatment of ligand–protein interactions. *Bioorg. Med. Chem.* **2016**, *24* (20), 4812–4825.

- (55) Jorgensen, W. L.; Briggs, J. M.; Gao, J. A priori calculations of pKa's for organic compounds in water. The pKa of ethane. *J. Am. Chem. Soc.* **1987**, *109* (22), 6857–6858.
- (56) Jorgensen, W. L.; Briggs, J. M. A priori pKa calculations and the hydration of organic anions. *J. Am. Chem. Soc.* **1989**, *111* (12), 4190–4197.
- (57) Cole, D. J.; Vilseck, J. Z.; Tirado-Rives, J.; Payne, M. C.; Jorgensen, W. L. Biomolecular force field parameterization via atom-in-molecule electron density partitioning. *J. Chem. Theory Comput.* **2016**, *12* (5), 2312–2323.
- (58) Horton, J. T.; Allen, A. E.; Dodda, L. S.; Cole, D. J. QUBEKit: automating the derivation of force field parameters from quantum mechanics. *J. Chem. Inf. Model.* **2019**, *59* (4), 1366–1381.
- (59) Unke, O. T.; Devereux, M.; Meuwly, M. Minimal distributed charges: Multipolar quality at the cost of point charge electrostatics. *J. Chem. Phys.* **2017**, *147* (16), 161712.
- (60) 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* (6), 1587–1597.
- (61) Lopes, P. E.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, A. D., Jr Polarizable force field for peptides and proteins based on the classical drude oscillator. *J. Chem. Theory Comput.* **2013**, *9* (12), 5430–5449.
- (62) Jorgensen, W. L.; Tirado-Rives, J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (19), 6665–6670.
- (63) 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* (8), 2549–2564.
- (64) Fanourgakis, G. S.; Xantheas, S. S. The Flexible, Polarizable, Thole-Type Interaction Potential for Water (TTM2-F) Revisited. *J. Phys. Chem. A* **2006**, *110* (11), 4100–4106.
- (65) Lopes, P. E.; Guvench, O.; MacKerell, A. D. Current status of protein force fields for molecular dynamics simulations. In *Molecular modeling of proteins*; Springer: Berlin, 2015; pp 47–71.
- (66) 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* (12), 3410–3417.
- (67) Kramer, C.; Gedeck, P.; Meuwly, M. Atomic multipoles: electrostatic potential fit, local reference axis systems, and conformational dependence. *J. Comput. Chem.* **2012**, *33* (20), 1673–1688.
- (68) Lee, B.; Richards, F. M. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.* **1971**, *55* (3), 379–IN4.
- (69) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.* **1971**, *54* (2), 724–728.
- (70) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys.* **1972**, *56* (5), 2257–2261.
- (71) Hariharan, P. C.; Pople, J. A. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* **1973**, *28* (3), 213–222.
- (72) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72* (1), 650–654.
- (73) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+ G basis set for first-row elements, Li–F. *J. Comput. Chem.* **1983**, *4* (3), 294–301.
- (74) Frisch, M. J.; Pople, J. A.; Binkley, J. S. Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets. *J. Chem. Phys.* **1984**, *80* (7), 3265–3269.
- (75) 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* (2), 1007–1023.
- (76) Kendall, R. A.; Dunning, T. H., Jr; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96* (9), 6796–6806.
- (77) Frisch, M. J.; et al. *Gaussian 09*, revision D.01; Gaussian, Inc.: Wallingford, CT, 2016.
- (78) Vega, C.; Abascal, J. L. F. Simulating water with rigid non-polarizable models: a general perspective. *Phys. Chem. Chem. Phys.* **2011**, *13* (44), 19663–19688.