# A Comparison Between QM/MM and QM/QM Based Fitting of Condensed-Phase Atomic Polarizabilities

| | |
|---|---|
| Journal: | *Physical Chemistry Chemical Physics* |
| Manuscript ID: | CP-ART-06-2014-002401.R1 |
| Article Type: | Paper |
| Date Submitted by the Author: | 24-Jun-2014 |
| Complete List of Authors: | Vosmeer, C.; VU University Amsterdam, AIMMS Divsion of Molecular Toxicology, Department of Chemistry and Pharmaceutical Sciences, Faculty of Sciences<br>Kiewisch, Karin; VU University Amsterdam, Amsterdam Center for Multiscale Modeling and Division of Theoretical Chemistry, Department of Chemistry and Pharmaceutical Sciences, Faculty of Sciences<br>Keijzer, Karlijn; VU University Amsterdam, AIMMS Divsion of Molecular Toxicology, Department of Chemistry and Pharmaceutical Sciences, Faculty of Sciences<br>Visscher, Lucas; VU University Amsterdam, Amsterdam Center for Multiscale Modeling and Division of Theoretical Chemistry, Department of Chemistry and Pharmaceutical Sciences, Faculty of Sciences<br>Geerke, Daan; VU University Amsterdam, AIMMS Divsion of Molecular Toxicology, Department of Chemistry and Pharmaceutical Sciences, Faculty of Sciences |
| | |

Faculty of Sciences / Department of Chemistry and Pharmaceutical Sciences
*Division of Molecular and Computational Toxicology*

Date

24th June 2014

| Telephone | Fax | E-mail |
|---|---|---|
| +31-20-59 87606 | +31-20-59 87610 | d.p.geerke@vu.nl |

Postal address

Dr. D.P. Geerke, De Boelelaan 1083, room P-262, NL-1081 HV Amsterdam

*vrije* Universiteit *amsterdam*

Dear Editor,

Thank you for the referee reports of our manuscript

“A Comparison Between QM/MM and QM/QM Based Fitting of Condensed-Phase Atomic Polarizabilities”

by

C.R. Vosmeer, K. Kiewisch, K. Keijzer, L. Visscher and D.P. Geerke.

We kindly thank the referees for their careful reading of our manuscript and for their recommendation for publication in *Physical Chemistry Chemical Physics*. We also thank the referees for their suggestions for further improvements. Please find below a detailed description of how we have incorporated the suggestions of the referees in the revised version of our manuscript. In the revised manuscript, we have highlighted added texts in blue (see the uploaded justification file “ann_paper.pdf”; for completeness we also uploaded a non-annotated version of the manuscript (“paper.pdf”)).

We have uploaded the revised version of our manuscript today via the *PCCP* submission platform, and we are looking forward to your decision on our manuscript.

Sincerely yours,

Daan P. Geerke
Assistant-Professor in Biomolecular Simulation and Modeling
AIMMS Division of Molecular and Computational Toxicology
Department of Chemistry and Pharmaceutical Sciences, VU University Amsterdam

Faculty of Sciences / Department of Chemistry and Pharmaceutical Sciences
*Division of Molecular and Computational Toxicology*

**Comments of the authors to the issues raised by the reviewers:**

*Reviewer 1*

1) ΔESP values were evaluated at Connolly grid points that were generated using the same routines and settings as described in reference 21. This is described in Section 2.2, page 3 ("grid-point coordinates […] grid settings.[21]").

2) We have added two references for the PW91k kinetic energy functional (references 55 and 56) on page 3.

3) The reason for the differences between our results and the findings of Cerutti *et al.* is that we determined atomic polarizabilities, based on induced dipole moments (fitted using QM/MM determined ΔESPs) and the solvent electric field at the solute atoms, whereas Cerutti determined solute partial charges by directly fitting based on QM/MM determined ESPs. We discussed the origin of these differences on page 3 of the manuscript ("The obtained overlap […] used in Equation 1."), and to further emphasize the difference, we have added the phrase "(based on solute ESPs)" on page 3 of the manuscript.

*Reviewer 2*

1) On page 1 of the revised manuscript, we have added references to work of, amongst others, J. Gao, C.E. Dykstra, G. Karlström, P. Söderhjelm, N. Gresh, S. Patel, R.A. Friesner and P. Cieplak (references 1-10, 13-19, and 25).

2) In Section 2.1, we have corrected the error in the time step, which is indeed 2 fs.

3) Both in the QM/MM and in the QM/FDE calculations, we have included only those solvent molecules for which the oxygen atom was within 1.4 nm of any of the solute atoms. We have clarified this by adding the phrase "for which […] solute atoms," on page 3 of the revised manuscript.

4) It would be quite expensive to obtain for all considered solvent configurations a self-consistent solution for the polarized solvent electron density, whereas performing such calculations for a few configurations only or for configurations obtained from simulations of much smaller system sizes would not be of additive value, due to poor statistics and finite-size effects, respectively. In order to address this issue in an alternative way, we have compared results from QM/FDE calculations using three very differently polarized solvent frozen densities (as obtained for a water molecule in vacuum, in implicit solvent, and in a small ice cluster). We have emphasized the rationale of using this approach on page 4 of the manuscript ("In order to […] hydrogen-bonding water molecules."), and our comparison showed that the choice of how to polarize the solvent molecules did not significantly affect the QM/FDE determined distributions of polarizabilities (see Section 3.2 and Figure 3).

5) We agree with the reviewer that we have previously drawn a too general conclusion on basis-set sensitivity by stating that “the fitted distributions are not significantly affected by the choice of basis set” (originally submitted manuscript, page 3). Therefore, we have removed this phrase from page 3 in the revised manuscript, and replaced it by “similar distributions […] when compared”.

6) On page 4 (Section 3.2) of the revised manuscript, we changed “significantly” into “substantially”.

7) By referring to the carbon polarizability as an “effective parameter” in the context of force-field development, we mean that its value can only be determined with large uncertainty (*i.e.*, from a broad range of fitted values, see also our reply to comment 8 of Reviewer 2, and Figure 3, panels B and D). This has been previously discussed by us in reference 21, and to clarify the meaning of “effective”, we have added the phrase “(*i.e.*, tunable)” on page 5 of the revised manuscript.

8) In reference 21, we discussed the origin of the wide spread of fitted values as typically obtained for $\alpha_C$ (*cf.* Figure 3, panels B and D of the current manuscript). As a result (and as mentioned in our reply to comment 7 of reviewer 2), condensed-phase fitting of atomic polarizabilities does not give as much direction to the parameterization of $\alpha_C$ when compared to $\alpha_O$ parameterization. For this reason, we have focused our discussion in Section 3.2 primarily on the $\alpha_O$ results. In order to emphasize this, we have added “In combination with […] fitted values for $\alpha_C$ (Figure 3),” on page 5 of the revised manuscript.

# A Comparison Between QM/MM and QM/QM Based Fitting of Condensed-Phase Atomic Polarizabilities

**C. Ruben Vosmeer,[a] Karin Kiewisch,[b] Karlijn Keijzer,[a] Lucas Visscher,[b] and Daan P. Geerke*[a]**

*Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX*
*First published on the web Xth XXXXXXXXXX 20XX*
**DOI: 10.1039/b000000x**

Recently we proposed a combined QM/MM approach to estimate condensed-phase values of atomic polarizabilities for use in (bio)molecular simulation. The setup relies on a MM treatment of the solvent when determining atomic polarizabilities to describe the response of a QM described solute to its external electric field. In this work, we study the effect of using alternative descriptions of the solvent molecules when evaluating atomic polarizabilities of a methanol solute. In a first step, we show that solute polarizabilities are not significantly affected upon substantially increasing the MM dipole moments towards values that are typically reported in literature for water solvent molecules. Subsequently, solute polarization is evaluated in the presence of a QM described solvent (using the frozen-density embedding method). In the latter case, lower oxygen polarizabilities were obtained than when using MM point charges to describe the solvent, due to introduction of Pauli-repulsion effects.

## 1 Introduction

Over the past years, the interest in polarizable force fields has been steadily increasing.[1–10] This is partly driven by the increase in available computational power, allowing for inclusion of additional degrees of freedom and levels of complexity in modeling nonbonded interactions in molecular simulations. The explicit inclusion of electronic polarization effects in biomolecular force fields is considered to be an answer to issues in the transferability of force field parameters between environments of different polarity.[11] Typically, non-polarizable force fields optimized to describe intermolecular interactions in polar environments may overestimate electrostatic interactions in less polar environments, and vice versa.[12] This in turn can effect the accuracy of predicting protein-ligand binding or other processes during which local electric fields vary.

Recent calibration of polarizable force fields has included parameter optimization for a large diversity of solvents and biomolecular building blocks.[8,13–19] As pointed out previously,[20,21] it is not trivial to determine appropriate values for the additional parameters that describe polarizabilities for use in non-additive force fields. Their determination commonly relies on gas-phase quantum-chemical calculations or estimates,[22–24] yielding values that are typically too high for use in the condensed phase. The need to scale down gas-phase estimates of atomic polarizabilities for use in condensed-phase simulation is typically explained to be the consequence of Pauli repulsion between electron distributions of nearby molecules,[20,25] but also of variations in the local electric field within molecular volumes[26] and of intramolecular polarization effects.[21]

Recently, we proposed an approach to directly obtain condensed-phase atomic polarizabilities for use in biomolecular simulation.[21] This method relies on molecular dynamics (MD) based sampling of solvent configurational space around a solute, using classical force fields. The generated configurations are subjected to combined QM/MM calculations, in which the external electrostatic potential (ESP) of the quantum-mechanically (QM) treated solute is determined in the presence of the solvent partial charges. Subsequently, atomic dipole moments are fitted to reproduce the difference in solute ESP when compared to the gas-phase electrostatic potential. From the fitted (induced) dipole moments and electric fields due to the solvent charges at the atomic sites, the polarizabilities are derived.[21]

Our approach assumes that partial charges (from a MM solvent model) appropriately describe the solvent charge distributions in the combined QM/MM Hamiltonian that is used for the determination of the solute's induced dipole moments. Replacing the full QM electron density of the solvent by MM point charges keeps computational costs of the (induced) ESP calculation tractable, and was shown to provide atomic polarizabilities for water and methanol solutes that are close to values used in previously reported polarizable models.[20,26–28]

---

[a] *AIMMS Division of Molecular Toxicology, Department of Chemistry and Pharmaceutical Sciences, Faculty of Sciences, VU University Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, the Netherlands; Tel: +31 (0)20 598 7606; Fax: +31 (0)20 598 7610; E-mail: d.p.geerke@vu.nl*
[b] *Amsterdam Center for Multiscale Modeling and Division of Theoretical Chemistry, Department of Chemistry and Pharmaceutical Sciences, Faculty of Sciences, VU University Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, the Netherlands*

In the current work, we study the effect of using alternative (improved) solvent descriptions on calculated values for the solute's atomic polarizabilities. In reference 21, we found that QM/MM fitted polarizabilities for methanol and water are not very sensitive to the choice of the solvent force field used in MD sampling and QM/MM calculations. For example, fitted values for carbon and oxygen polarizabilities of methanol in either SPC[29] or SPC/E water[30] typically differ by less than 5 %. However, molecular dipole moments of SPC and SPC/E vary between 2.23 and 2.35 D only, whereas experimentally or quantum-mechanically derived values for condensed-phase water dipole moments are reported within a range of 2.6 – 2.9 D or higher.[31–34] In order to determine polarized point charge distributions of hydrated solute molecules (for use in non-polarizable force fields) from QM/MM based ESP fitting, Cerutti *et al.*[35] recently stressed the need to increase the solvent charges that enter the solute's QM hamiltonian towards values representing fully polarized water molecules (with a dipole moment $> 2.6$ D). In the current work, we study the effect of increasing the MM point charges accordingly in our QM/MM determination of atomic polarizabilities of a methanol solute. For that purpose, we still use a classical solvent model (SPC[29]) to generate solvent configurations, because these models are known to appropriately describe thermodynamic properties of water, and hydrogen bonding and local structure in bulk water. Only in the QM/MM calculations of the solute's (induced) ESP, the charges are increased to reach a condensed-phase value of 3.0 D for the molecular dipole moment of water, in a similar approach to the one used by Cerutti *et al.*[35]

In the second part of this study, we evaluate the effect of using a QM description of the solvent on the (QM) determination of the solute's atomic polarizabilities. This is achieved by applying the Frozen-Density Embedding (FDE) method[36–38] to describe the solvent electron density. For this purpose, pre-optimized solvent densities are mapped onto the solvent configurations obtained from classical MD. The solute wave function and ESP are then optimized and determined in the presence of the frozen solvent electron densities. This approach allows to fit condensed-phase polarizabilities based on QM/QM (QM/FDE) derived induced ESPs to include a higher level of detail into the determination of the induced atomic dipole moments of the solute, while keeping computational costs within achievable limits. As such, we can compare evaluation of atomic polarizabilities when using either a MM or a QM description of the solvent.

## 2 Computational methods

For a methanol solute in water, we fitted induced atomic dipole moments (and derived atomic polarizabilities) to describe the changes in solute ESP in the presence of the solvent (when compared to the solute gas-phase ESP), using the approach introduced in our previous work.[21] For this purpose, solvent configurations around the solute are generated first (in classical MD simulations). Subsequently, the obtained configurations are used for the (QM) determination of the change in solute ESP ($\Delta$ESP) induced by the solvent. Using the least-squared fitting procedure described before,[21] induced dipole moments $\vec{\mu}$ are fitted on the (non-hydrogen) atomic centers $i$ that best describe $\Delta$ESP. In a next step, the external electric fields $\vec{E}_i$ at the positions of the solute atoms due to the solvent charge distribution are determined for use in deriving distributions of atomic polarizabilities $\alpha_i$ from the $\vec{\mu}_i$'s according to:

$$\alpha_{i,k} = \frac{\mu_{i,k}}{E_{i,k}} \qquad (1)$$

with $k$=$x$, $y$ or $z$, and $\alpha_i = (\alpha_{i,x} + \alpha_{i,y} + \alpha_{i,z})/3$. Finally, a Gaussian curve is fitted to each of the distributions of atomic polarizabilities obtained with the different configurations, of which the mean value is reported as the fitted (average) value for $\alpha_i$.

The following describes the modifications introduced in the current work in the protocol to generate solvent configurations (Section 2.1) and in the QM/MM (or QM/FDE) calculation of the solute's induced ESP (Section 2.2). The details of the Frozen Density Embedding (QM/FDE) calculations are discussed in Section 2.3.

### 2.1 Generation of solvent configurations

During MD, atomic positions of the methanol solute molecule were kept constrained to their (gas-phase) optimized geometry. The gas-phase optimized geometry was obtained using the ADF software package (version 2012),[39,40] with the B3LYP exchange correlation functional[41–43] and using the QZ4P basis set.[44] MD simulations were performed at constant volume, to maintain the optimized solute geometry during a 1 ns production simulation after 1 ns equilibration. During production, 500 solvent configurations were collected every 1000 time steps of 2 fs. Two independent simulations were performed using either the SPC[29] or TIP4P[45] water model to describe the 1408 solvent molecules. All other MD settings used to generate solvent configurations around the methanol solute were identical to those described in our previous work.[21]

### 2.2 Determination of the solute's ESP in the presence of solvent point charges

In order to use an established framework for FDE (Section 2.3), the optimization of the solute wave function and the solute's ESP determination (with the B3LYP exchange correlation functional[41–43]) were performed using the ADF software package (version 2012)[39,40] and the Slater-type QZ4P basis set[44] (instead of the previously used[21] Gaussian-type aug-cc-pVTZ basis set[46,47]). Because generation of a Connolly grid[48] for use in the ESP evaluation is currently not implemented in ADF, grid-point coordinates were generated with the GAMESS-US software[49,50] using previously described routine and grid settings.[21] The obtained grid coordinates were subsequently imported into ADF and the solute ESP was evaluated at the grid points. For the configurations obtained from MD, solute ESPs were calculated in absence and presence of all solvent molecules for which the oxygen atom was within 1.4 nm of any of the solute atoms. Solvent molecules were treated by using either the MM point charges of the water model used to generate the solvent configurations, or by using point charges that were increased by 25 %. In addition, solute ESPs and electric fields at the atoms were determined by describing the solvent using FDE (Section 2.3).

### 2.3 Frozen Density Embedding

In the QM/FDE determination of the solute ESP, the solvent electron density was generated by superposing the B3LYP/QZ4P optimized electron density calculated for a single water molecule onto all solvent molecules for which the oxygen atom was within 1.4 nm of any of the solute atoms, by translating and rotating the positions of the nuclei to the positions of the atoms of the solvent molecules using ADF and a geometry tolerance threshold of $10^{-2}$ for the mapping. Three different approaches to generate the optimized electron density of the single water ”template” molecule were followed: first, a gas-phase single-point density optimization was performed with the coordinates of the first water molecule selected from the first MD configuration. In the second case, instead of a gas-phase calculation, the conductor-like screening model (COSMO,[51] a continuum solvent model) was used during optimization. In a third setup, the density of the template water molecule was determined for a (central) water molecule surrounded by four other water molecules in a cluster with nearly $C_{2v}$-symmetric geometry, mimicking a small part of an ice cluster. Deviations from $C_{2v}$ symmetry only arose from the fact that during MD simulations, the structure of the single water molecules slightly differed from $C_{2v}$ symmetry within a certain threshold (defined by the geometric tolerance of $10^{-4}$ used in the SHAKE[52] algorithm to constrain interatomic distances during MD). The geometry of the central water molecule was imposed on all water molecules in the small ice-like cluster. The optimized ”ice water” density of the central molecule was subsequently used in generating the FDE densities of the MD solvent configurations.

All QM/FDE calculations were performed with the ADF program (version 2012),[39,40,53,54] employing the B3LYP exchange-correlation functional[41,43] and the kinetic-energy functional PW91k.[55,56] For all fragments, the QZ4P basis set[44] was used.

For the QM/FDE calculations, the electric fields at the oxygen and carbon atoms of the methanol solute were calculated from the derivative of the electrostatic potential due to the frozen solvent fragments (with respect to the $x$, $y$ and $z$ coordinates). For this purpose, cubic grids of 25x25x25 points centered at the carbon and oxygen atoms, and a grid spacing of 0.0001 Bohr were used. The electric field contributions were taken as the linear parameters from a polynomial fit to the 25 points along the $x$, $y$ and $z$ axes.

## 3 Results and Discussion

### 3.1 QM/MM determined polarizabilities

For every of the 500 solvent configuration frames obtained from MD with SPC as solvent model, induced atomic dipole moments at methanol’s carbon and oxygen were fitted that best describe the QM/MM determined change in the solute’s ESP. Gaussian-fitted distributions of the derived atomic polarizabilities are displayed in Figure 1. Using the SPC solvent model and the B3LYP/QZ4P level of DFT to describe the solute, we obtained distributions of oxygen polarizabilities with a mean value of $1.0 \times 10^{-3}$ nm$^3$, while for carbon the distribution of polarizabilities is broad and has a mean value of $0.9 \times 10^{-3}$ nm$^3$. Figure 1 shows that similar distributions are obtained using the QZ4P basis set when compared to our previous results, obtained with the aug-cc-pVTZ basis set to describe the solute.

Subsequently, we studied the effect of increasing the MM point charges in the QM/MM calculations by 25 %, such that the dipole moment of the water molecules was enhanced to 3.0 D, which is close to condensed-phase values reported in literature.[31–34] Both for methanol’s carbon and oxygen atoms, Figure 2, panel A shows direct overlap between the fitted distributions of polarizabilities when using either SPC charges or SPC charges increased by 25 % for the MM part of the QM/MM Hamiltonian. Thus, increasing the dipole moment of the solvent molecule to (average) reported condensed-phase values does not affect the outcomes of our QM/MM determination of condensed-phase polarizabilities. This is in contrast with recent results of Cerutti *et al.*,[35] who could improve QM/MM based condensed-phase fitting of solute *point charges* (based on solute ESPs) by increasing the solvent point charges in the QM/MM calculations. The obtained overlap in the distributions of fitted polarizabilities (Figure 2, panel A) indicates that the increase in the electric fields at the solute atomic centers (Figure 2, panel B) coincides with a corresponding increase in the induced atomic dipole moments, in line with the linear response assumption used in Equation 1.

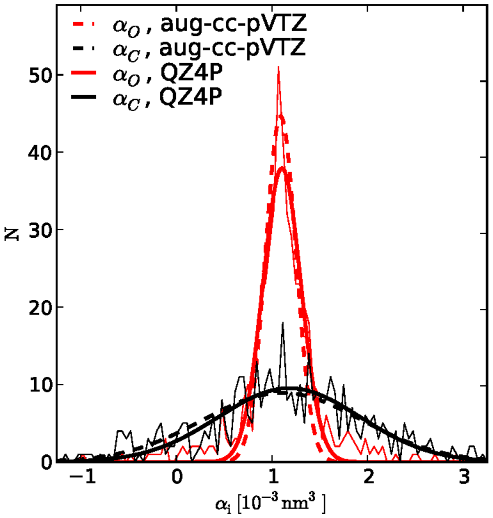


**Fig. 1** Gaussian-fitted counted distributions of QM/MM determined oxygen ($\alpha_O$, red) and carbon ($\alpha_C$, black) polarizabilities for a methanol solute in SPC water, obtained for 500 solvent configurations from MD, using the QZ4P basis set[44] for the methanol solute (solid lines). For reference, raw data are shown by thin lines, and the Gaussian-fitted counted distributions obtained using the aug-cc-pVTZ basis set[46,47] are shown by dashed lines.

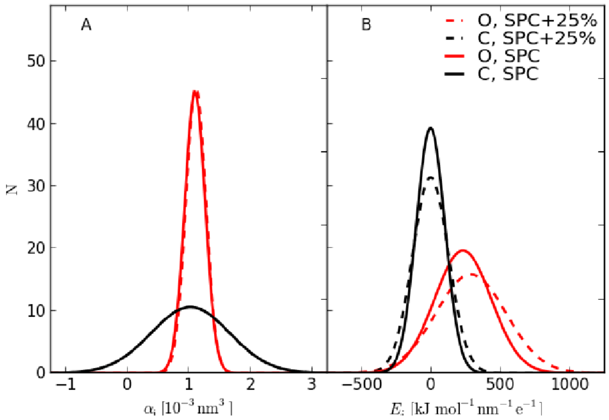


**Fig. 2** (A) Gaussian-fitted counted distributions of QM/MM determined oxygen ($\alpha_O$, red) and carbon ($\alpha_C$, black) polarizabilities for a methanol solute in SPC water, obtained for 500 solvent configurations from MD, using the QZ4P basis set[44] for the methanol solute (solid lines). In addition, Gaussian-fitted counted distributions for $\alpha_O$ and $\alpha_C$ are shown as obtained using increased solvent point charges (by 25 %) in the QM/MM calculations (dashed lines). (B) Gaussian-fitted counted distributions of the size of the electric field at methanol's oxygen (red) and carbon (black) atoms obtained for 500 solvent configurations from MD using the SPC model, and using SPC point charges (solid line) or SPC point charges increased by 25 % (dashed line) to describe the solvent charge distributions.

### 3.2 QM/FDE determined polarizabilities

Subsequently, we studied the effect of using a quantum-mechanical description of the solvent on the determination of the solute's atomic polarizabilities. To keep our calculations computationally feasible, the Frozen-density embedding (FDE) approach was used. In order to study the effect of following different strategies to determine the frozen electron density of the solvent molecules, results were compared for fitting atomic polarizabilities of a methanol solute that was embedded in water solvent for which the electron density of the single "template" molecules was determined either (i) in the gas phase, (ii) in an implicit (COSMO[51]) solvent model, or (iii) for the central (polarized) molecule of a $C_{2v}$-symmetrized cluster of five hydrogen-bonding water molecules. In addition, we compared use of the SPC water model in generating the solvent configurations with use of the TIP4P model, which has a geometry that is closer to the QM optimized geometry of an isolated water molecule.[45]

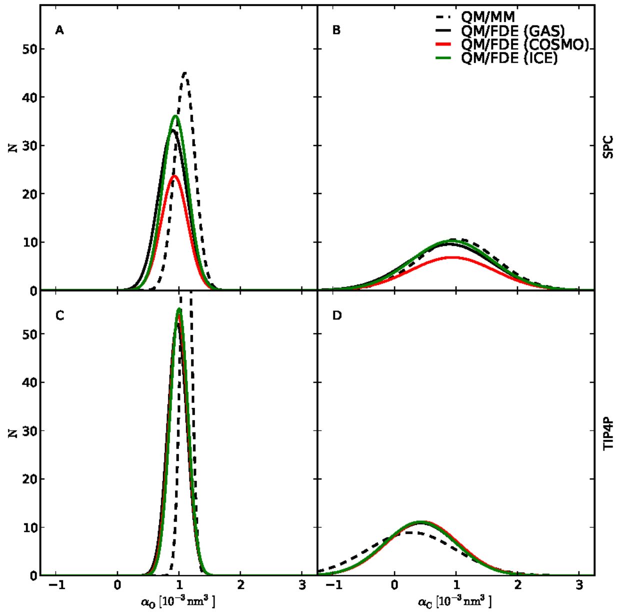


**Fig. 3** Gaussian-fitted counted distributions of QM/MM (dashed lines) and QM/FDE (solid lines) fitted polarizabilities for the oxygen (panels A and C) and carbon (panels B and D) atoms of a methanol solute in SPC (panels A and B) and TIP4P (panels C and D) water, obtained for 500 MD configurations. In the FDE calculations, frozen densities of each individual water molecule were either optimized in the gas phase (black), in COSMO implicit solvent (red), or for a central molecule in a $C_{2v}$-symmetrized cluster of five water molecules (green).

Figure 3 shows that both in SPC and TIP4P water, the obtained QM/FDE polarizabilities are not substantially affected by the approach chosen to generate the frozen electron density to describe the solvent. In addition, fitted average values for the oxygen polarizability ($\alpha_O$) do not significantly depend on the choice of the water model in the MD simulations to generate the configurations for use in the QM/FDE calculation: obtained $\alpha_O$ values range from 0.90 to $0.94 \times 10^{-3}$ nm$^3$ when using SPC, and from 0.97 to $1.00 \times 10^{-3}$ nm$^3$ when using the TIP4P model (Figure 3, panels A and C). In contrast, average values obtained for carbon's polarizability ($\alpha_C$) differ from $0.89 - 0.93 \times 10^{-3}$ nm$^3$ to $0.42 - 0.45 \times 10^{-3}$ nm$^3$, respectively (Figure 3, panels B and D). Although the QM/MM and QM/FDE obtained distributions of $\alpha_C$ values show strong overlap when using SPC to generate the solvent configurations (Figure 3, panel B), the mean value of the $\alpha_C$ distribution changes from 0.29 to $0.43 \times 10^{-3}$ nm$^3$ when going from QM/MM to QM/FDE fitting using the TIP4P configurations (Figure 3, panel D). In combination with the large spread of fitted values for $\alpha_C$ (Figure 3), the differences in the obtained mean values reemphasize that $\alpha_C$ can be considered as an effective (*i.e.*, tunable) parameter in the process of polarizable force field parametrization.[21]

Interestingly, average fitted values for $\alpha_O$ were systematically lower in the QM/FDE calculations than in QM/MM (by $0.15 - 0.19 \times 10^{-3}$ nm$^3$ using SPC and by $0.14 - 0.17 \times 10^{-3}$ nm$^3$ with TIP4P, see Figure 3, panels A and C, respectively). The lower polarizabilities for oxygen obtained with FDE, can be explained from the introduction of Pauli repulsion between solute and solvent electron densities. This was already suggested as an explanation for the need of using lower polarizabilities for condensed-phase systems when compared to gas-phase values.[20] The introduction of solute-solvent Pauli repulsion upon going from a MM to a FDE representation of the solvent is illustrated by differences between the solute electron density as determined either in the presence of MM point charges or FDE-described water molecules. Such difference densities are shown in Figure 4 for a hydrated methanol solute that donates or accepts a hydrogen bond. In both cases, lower electron densities are observed at the hydrogen bond-donor or accepting group when using a QM (FDE) solvent description instead of MM point charges, which confirms that inclusion of solute-solvent Pauli repulsion contributes to a decrease in solute polarizability in the condensed phase. From the fits in Figure 3, panels A and C, this effect accounts for a decrease in polarizability of 9-12 % for methanol's oxygen, whereas differences between gas-phase and condensed-phase fitted (molecular) polarizabilities of approximately 30 % are commonly reported for polarizable methanol models.[20,24] These findings confirm that other factors such as intramolecular polarization[21] and variations in electric fields[26] play a role as well in the need for lower polarizabilities in the condensed phase.

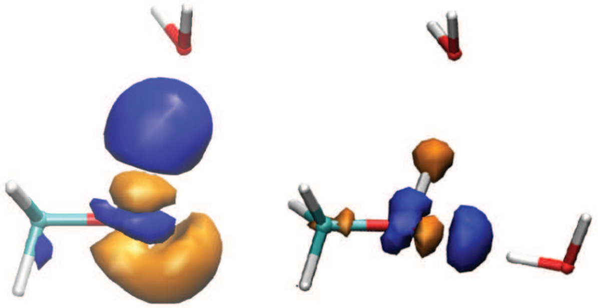


**Fig. 4** Difference density maps for a methanol solute as obtained in water solvent described by either FDE or MM point charges, and as evaluated for the first (left) and eleventh (right) configuration obtained from MD simulation using the TIP4P water model. Differences $\Delta\rho$ of the isosurface density value are shown for $\Delta\rho = \pm 0.001$ a.u. Blue denotes negative $\Delta\rho$ values ($\rho$(FDE) $<$ $\rho$(MM)) and orange denotes positive values for $\Delta\rho$ ($\rho$(FDE) $>$ $\rho$(MM)).

## 4 Conclusions

In this study we have evaluated the effect of the descriptions of the solvent molecules on the QM/MM and QM/QM fitted values of condensed-phase polarizabilities of a methanol solute in a water solvent. It was found that increasing the solvent's molecular dipole moment from its MM model value by 25 % to a value that is within the range reported in literature does not change the QM/MM values for the polarizabilities, because the increase in electric fields at the solute atoms due to the solvent is equivalent to the increase in fitted values for the induced atomic dipole moments.

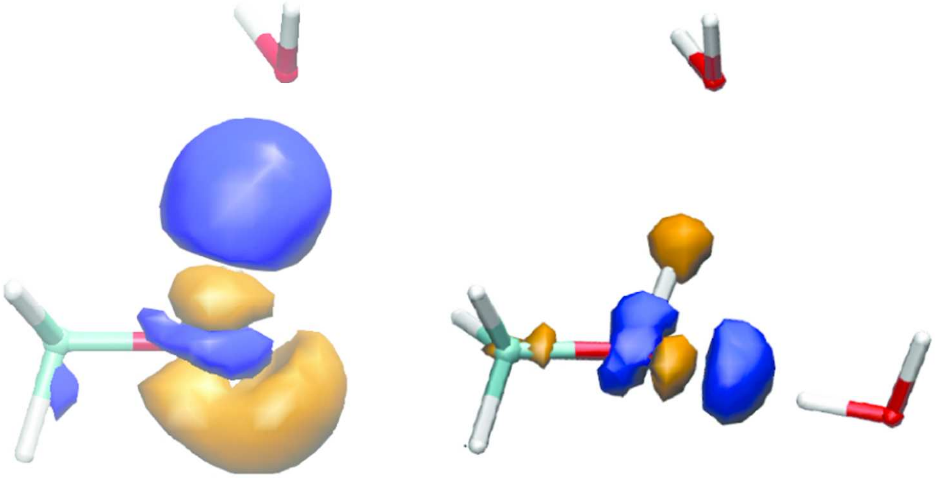
Introducing a QM description of the solvent by means of frozen density embedding was found to affect the fitted values for the oxygen polarizabilities of the methanol solute. Compared to the QM/MM results, QM/FDE values for the fitted oxygen polarizabilities were found to be approximately 10 % lower. This difference can be taken into account in the process of QM/MM based force field calibration of atomic polarizabilities for use in (bio)molecular simulation, and is due to the introduction of solute-solvent Pauli repulsion effects in the QM/FDE calculations, which can only partly explain the difference in condensed-phase and gas-phase estimates for the solute polarizabilities.

## Acknowledgements

We kindly thank dr. William C. Swope and dr. Julia E. Rice for stimulating discussions and for fitting tools. Financial support from the Netherlands Organisation for Scientific Research (NWO VENI grant 700.59.406) is gratefully acknowledged.

## References

1. J. M. Stout and C. E. Dykstra, *J. Am. Chem. Soc.*, 1995, **117**, 5127–5132.
2. S. Rick and S. Stuart, *Potentials and algorithms for incorporating polarizability in computer simulations*, John Wiley & Sons, Inc., New Jersey, USA, 2002, vol. 18.
3. H. B. Yu and W. F. van Gunsteren, *Comput. Phys. Commun.*, 2005, **172**, 69–85.
4. W. Xie and J. Gao, *J. Chem. Theory Comput.*, 2007, **3**, 1890–1900.
5. N. Gresh, G. A. Cisneros, T. A. Darden and J.-P. Piquemal, *J. Chem. Theory Comput.*, 2007, **3**, 1960–1986.
6. A. Warshel, M. Kato and A. V. Pisliakov, *J. Chem. Theory Comput.*, 2007, **3**, 2034–2045.
7. P. Cieplak, F.-Y. Dupradeau, Y. Duan and J. Wang, *J. Phys. Cond. Mat.*, 2009, **21**, 333102.
8. J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson and T. Head-Gordon, *J. Phys. Chem. B*, 2010, **114**, 2549–2564.
9. X. Zhu, P. E. M. Lopes and A. D. MacKerell, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2012, **2**, 167–185.
10. S. Ou, T. R. Lucas, Y. Zhong, B. A. Bauer, Y. Hu and S. Patel, *J. Phys. Chem. B*, 2013, **117**, 3578–3592.
11. D. P. Geerke and W. F. van Gunsteren, *J. Phys. Chem. B*, 2007, **111**, 6425–6436.
12. C. Oostenbrink, A. Villa, A. E. Mark and W. F. van Gunsteren, *J. Comput. Chem.*, 2004, **25**, 1656–1676.
13. G. A. Kaminski, H. A. Stern, B. J. Berne and R. A. Friesner, *J. Phys. Chem. A*, 2004, **108**, 621–627.
14. J. R. Maple, Y. Cao, W. Damm, T. A. Halgren, G. A. Kaminski, L. Y. Zhang and R. A. Friesner, *J. Chem. Theory Comput.*, 2005, **1**, 694–715.
15. Y. Zhong and S. Patel, *J. Phys. Chem. B*, 2010, **114**, 11076–11092.
16. E. Wernersson and P. Jungwirth, *J. Chem. Theory Comput.*, 2010, **6**, 3233–3240.
17. C. M. Baker, V. M. Anisimov and A. D. MacKerell, *J. Phys. Chem. B*, 2011, **115**, 580–596.
18. P. Ren, C. Wu and J. W. Ponder, *J. Chem. Theory Comput.*, 2011, **7**, 3143–3161.
19. Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder and P. Ren, *J. Chem. Theory Comput.*, 2013, **9**, 4046–4063.
20. G. Lamoureux, A. D. MacKerell and B. Roux, *J. Chem. Phys.*, 2003, **119**, 5185–5197.
21. C. R. Vosmeer, A. S. Rustenburg, J. E. Rice, H. W. Horn, W. C. Swope and D. P. Geerke, *J. Chem. Theory Comput.*, 2012, **8**, 3839–3853.
22. J. Applequist, J. R. Carl and K. K. Fung, *J. Am. Chem. Soc.*, 1972, **94**, 2952–2960.
23. K. Miller, *J. Am. Chem. Soc.*, 1990, **112**, 8543–8551.
24. V. M. Anisimov, G. Lamoureux, I. V. Vorobyov, N. Huang, B. Roux and A. D. MacKerell, *J. Chem. Theory Comput.*, 2005, **1**, 153–168.
25. P. Söderhjelm, A. Öhrn, U. Ryde and G. Karlström, *J. Chem. Phys.*, 2008, **128**, 014102.
26. B. Schropp and P. Tavan, *J. Phys. Chem. B*, 2008, **112**, 6233–6240.
27. H. B. Yu, D. P. Geerke, H. Y. Liu and W. F. van Gunsteren, *J. Comput. Chem.*, 2006, **27**, 1494–1504.
28. V. M. Anisimov, I. V. Vorobyov, B. Roux and A. D. MacKerell, *J. Chem. Theory Comput.*, 2007, **3**, 1927–1946.
29. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren and J. Hermans, *Intermolecular forces*, Reidel, Dordrecht, The Netherlands, 1981, pp. 331–338.
30. H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269–6271.
31. G. Maroulis, *Chem. Phys. Lett.*, 1998, **289**, 403–411.
32. B. Guillot, *J. Mol Liq.*, 2002, **101**, 219–260.
33. Y. S. Badyal, M. L. Saboungi, D. L. Price, S. D. Shastri, D. R. Haeffner and A. K. Soper, *J. Chem. Phys.*, **112**, 9206–9208.
34. P. L. Silvestrelli and M. Parrinello, *Phys. Rev. Lett.*, 1999, **82**, 3308–3311.
35. D. S. Cerutti, J. E. Rice, W. C. Swope and D. A. Case, *J. Phys. Chem. B*, 2013, **117**, 2328–2338.
36. T. A. Wesolowski and A. Warshel, *J. Phys. Chem.*, 1993, **97**, 8050–8053.
37. J. Neugebauer, C. R. Jacob, T. A. Wesolowski and E. J. Baerends, *J. Phys. Chem. A*, 2005, **109**, 7805–7814.
38. C. R. Jacob, J. Neugebauer, L. Jensen and L. Visscher, *Phys. Chem. Chem. Phys.*, 2006, **8**, 2349–2359.
39. C. Fonseca Guerra, J. G. Snijders, G. te Velde and E. J. Baerends, *Theor. Chem. Acc*, 1998, **99**, 391–403.
40. G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders and T. Ziegler, *J. Comput. Chem.*, 2001, **22**, 931–967.
41. R. H. Hertwig and W. Koch, *Chem. Phys. Lett.*, 1997, **268**, 345–351.
42. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
43. P. J. Stephens, F. J. Devlin, C. F. Cabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.
44. E. van Lenthe and E. J. Baerends, *J. Comput. Chem.*, 2003, **24**, 1142–1156.
45. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
46. R. A. Kendall, T. H. Dunning and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796–6806.
47. F. de Proft, F. Tielens and P. Geerlings, *J. Molec. Struc. (Theochem.)*, 2000, **506**, 1 – 8.
48. U. Singh and P. Kollman, *J. Comput. Chem.*, 1984, **5**, 129–145.
49. M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis and J. A. Montgomery, *J. Comput. Chem.*, 1993, **14**, 1347–1363.
50. M. Gordon and M. Schmidt, *Theory and applications of computational chemistry: the first forty years*, Elsevier, Amsterdam, 2005.
51. C. C. Pye, T. Ziegler, E. van Lenthe and J. N. Louwen, *Can. J. Chem*, 2009, **87**, 790–797.
52. J.-P. Ryckaert, G. Ciccotti and H. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.
53. C. R. Jacob, J. Neugebauer and L. Visscher, *J. Comput. Chem.*, 2008, **29**, 1011–1018.
54. C. R. Jacob, S. M. Beyhan, R. E. Bulo, A. S. P. Gomes, A. W. Götz, K. Kiewisch, J. Sikkema and L. Visscher, *J. Comput. Chem.*, 2011, **32**, 2328–2338.
55. A. Lembarki and H. Chermette, *Phys. Rev. A*, 1994, **50**, 5328–5331.
56. T. A. Wesolowski, H. Chermette and J. Weber, *J. Chem. Phys.*, 1996, **105**, 9182–9190.



Condensed-phase polarizabilities were evaluated for a solute in the presence of a QM treated solvent (using the frozen-density embedding method), and compared to results for a solute surrounded by MM solvent point charges. In this way, the effects of solute-solvent Pauli repulsion on solute polarizabilities could be estimated.
80x41mm (300 x 300 DPI)