Journal of Molecular Liquids 356 (2022) 119033

Contents lists available at ScienceDirect

Journal of Molecular Liquids

journal homepage: www.elsevier.com/locate/molliq

The dipole moment of alcohols in the liquid phase and in solution

Miguel Jorge [a,*], José R.B. Gomes [b], Maria Cecilia Barrera [c]

[a] Department of Chemical and Process Engineering, University of Strathclyde, 75 Montrose Street, Glasgow G1 1XJ, United Kingdom
[b] CICECO – Aveiro Institute of Materials, Department of Chemistry, University of Aveiro, Campus Universitário de Santiago, Aveiro, Portugal
[c] Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, 161 Cathedral St, Glasgow G4 0RE, United Kingdom

ARTICLE INFO

Article history:
Received 4 February 2022
Revised 24 March 2022
Accepted 26 March 2022
Available online 31 March 2022

Keywords:
Force fields
Polarization
QM/MM
Solvation
Molecular dynamics

ABSTRACT

Understanding polarization effects in condensed phases, like liquids and solutions, requires computational methods that can accurately predict dipole moments and energy of polarized molecules. In this paper, we report an improvement and extension of our recently developed Self-Consistent Electrostatic Embedding (SCEE) method, and apply it to determine the dipole moment of pure liquid alcohols, as well as of methanol dissolved in a variety of solvents (namely, other alcohols, water and hexadecane). We observe that the dipole moments of pure liquid alcohols are enhanced by ~0.9 D over their gas phase values, which is similar to the dipole enhancement previously observed for water, and much higher than what is predicted by dielectric continuum models. Our results demonstrate the importance of accounting for local solvation effects, namely the formation of hydrogen bonds, when calculating the extent of liquid phase polarization. In fact, we argue that the dipole enhancement upon solvation can be explained as a superposition of two effects: bulk screening described by the solvent dielectric constant and local solvation that requires a discrete molecular-level description of the system. SCEE is able to account for both effects simultaneously, and is thus a powerful tool to estimate polarization effects in liquids and solutions.

1. Introduction

Alcohols are extremely important and useful molecules, with wide ranging applications, including as solvents, fuels and chemical feedstocks. For example, methanol and ethanol produced from crops (so-called biomethanol and bioethanol, respectively) are some of the main sustainable alternatives to petrol for use in internal combustion engines [1,2]. Methanol is a major chemical feedstock in the production of many intermediate compounds [3], particularly of formaldehyde, which is then used in the manufacture of paints, plastics, textiles and cosmetics [4]. It is also considered a potential hydrogen carrier for bulk storage and transport of hydrogen [5]. Alcohol solvents play a crucial role in pharmaceutical development and manufacture, with the octanol/water partition coefficient being used, for example, to describe a drug's ability to diffuse through lipids [6]. Many alcohols also possess antiseptic properties, and have thus been on the frontline of our battle against the recent Covid-19 pandemic [7]. Apart from their practical importance, alcohols are also interesting from a fundamental point of view [8–10], as they are the simplest molecules that combine a hydrophobic moiety with a hydrogen-bonding functional group. In fact, long-chain alcohols are often considered to be the simplest amphiphiles, which is another reason why the octanol/water partition coefficient is widely used as a benchmark for the balance between lipophilicity and hydrophilicity [11].

Because alcohols are most often used in the liquid state or as solutions/mixtures, increasing our molecular-level knowledge of these systems assumes great importance [12]. This includes understanding how alcohol molecules are polarized when transferred from the gas phase to a liquid or solution, since these polarization effects have wide ranging implications for thermodynamic, structural and electronic properties of the medium, and are key to the efficiency of alcohols in the above-mentioned applications. Polarization of a molecule by the electric field induced by the surrounding molecules in a condensed phase causes the geometry and the electronic cloud of that molecule to distort, which carries an energy penalty – the so-called distortion energy. However, the distorted molecule interacts more favorably with the surrounding environment, leading to an overall decrease in the free energy of the system. Due to the molecular distortion caused by polarization, the dipole moment of polar molecules is significantly higher in condensed phases than in the gas phase [13–17]. Although, qualitatively, these effects have been known for decades (see, e.g. [18]

* Corresponding author.
E-mail address: miguel.jorge@strath.ac.uk (M. Jorge).

https://doi.org/10.1016/j.molliq.2022.119033
0167-7322/

and references therein), actually quantifying the degree of dipole enhancement and the energetic contributions to polarization is extremely challenging, mainly due to the difficulties in decoupling the electronic properties of individual molecules from those of the surrounding molecules in the liquid/solution.

To our knowledge, the only experimental estimate of the real liquid dipole moment has been carried out for water [19], which means that no experimental benchmark value is available for alcohols. However, we were able to find a few computational estimates for small alcohols (mainly methanol), based on either *Ab Initio* Molecular Dynamics (AIMD) [20] or Quantum Mechanics/Molecular Mechanics (QM/MM) approaches [21]. An early study by Gao and Xia [22] used QM/MM with the semi-empirical AM1 approach for the QM region and the TIP3P force field for the MM region, and computed a dipole moment of 2.06 D for methanol. Ten years later, Martín et al. [23] carried out more detailed QM/MM calculations using the Hartree-Fock (HF) method and the aug-cc-pVDZ basis set [24,25] with a mean-field description of the surrounding solvent obtained from Molecular Dynamics (MD) simulations. They obtained a dipole moment of 2.46 D for methanol, which was significantly larger than the value of Gao and Xia, with the authors attributing this to the shortcomings of the simplified AM1 approach [23]. Martín et al. also reported QM/MM dipole moments of 2.37 D for ethanol and 2.27 D for 1-propanol. In 2003, Pagliai et al. [26] reported a value of 2.64 D for methanol from AIMD simulations with the BLYP exchange–correlation functional (based on the Generalized-Gradient Approximation, GGA) on a box of 26 molecules. The same functional and method were used by Handgraaf et al. [27] with a larger box of 64 methanol molecules, yielding a value of 2.59 D. When studying the methanol vapor/liquid interface by AIMD, using both GGA-based BLYP and PBE functionals and a box containing 120 molecules, Kuo et al. [28] reported average bulk liquid methanol dipole moments of ~2.7 D. A more recent AIMD study by Sieffert et al. [29] considering a box of 64 molecules and GGA (BLYP and BP86) or hybrid (B97) functionals, some of which included Grimme's dispersion corrections [30,31], reported methanol dipole moment values ranging from 2.58 to 2.84 D.

Overall, the above studies, particularly more recent ones that use higher levels of theory, confirm that the methanol dipole moment is significantly enhanced in the liquid phase, relative to the gas phase value of 1.70 D [32–34]. However, as we discussed at length in our previous publication [17], the computational approaches used previously suffer from inherent limitations – e.g., difficulty in obtaining a reliable description of the structural and thermodynamic properties of the liquid/solution (particularly when relatively low QM levels of theory need to be employed); high computational cost; intrinsic ambiguity when decoupling the electronic properties of the central molecule from those of adjacent molecules; and/or challenges in fully accounting for the mutual polarization of solute and solvent molecules. As such, there is still a lack of a detailed and systematic study of the effects of polarization on alcohol molecules in the liquid state and in solution. In this paper, we attempt to fill this gap by applying a recently developed computational approach, the Self-Consistent Electrostatic Embedding (SCEE) method [17], to alcohol molecules of different polarity and chemical structure. In Section 2, we describe the basis of the SCEE approach and report some improvements and generalizations of the method that allow it to be applied to larger molecules as well as heterogeneous solutions – i.e. where the solute and solvent are different species. In Section 3, we present a detailed analysis of the dipole moments and polarization energies of several aliphatic alcohols, taking into account the increase in alkyl chain length as well as the location of the hydroxyl group. We end the paper with a summary of the main conclusions and suggestions for future work.

2. Methodology

2.1. Self-Consistent Electrostatic Embedding

The SCEE method, developed in our previous publication [17], is a QM/MM approach designed specifically to compute dipole moments and polarization energies in condensed phases, that aims to strike a balance between traditional electrostatic embedding (EE) – where the electrostatic MM environment surrounding the central QM molecule of interest is fixed, hence underestimating the degree of polarization – and polarized embedding (PE) – where the MM environment is able to adapt on-the-fly, hence capturing the correct degree of polarization but at a much greater computational cost. Like other QM/MM methods [21], SCEE calculations consider a QM region, in this case composed only of a central "solute" molecule, surrounded by an approximately spherical MM cluster of surrounding solvent molecules represented by point charges. The unique idea behind SCEE, shown schematically in Fig. 1, is to carry out a small number of EE calculations at different magnitudes of surrounding charges, then find the liquid-phase dipole moment by self-consistently matching the dipole moments of the QM and MM molecules.

We found that carrying out 3 EE calculations, followed by a quadratic fit of the QM dipole moment as a function of the MM dipole moment (see Figure S1 for an example of such a fit), was sufficient to achieve convergence in a computationally expedient way [17]. Moreover, we also identified the optimal conditions for calculating the dipole moment of liquid water [17]: QM calculations using the B3LYP functional [35] and the aug-cc-pVTZ basis set [24,25], cut-off radius of 1.5 nm for building the solvation clusters,

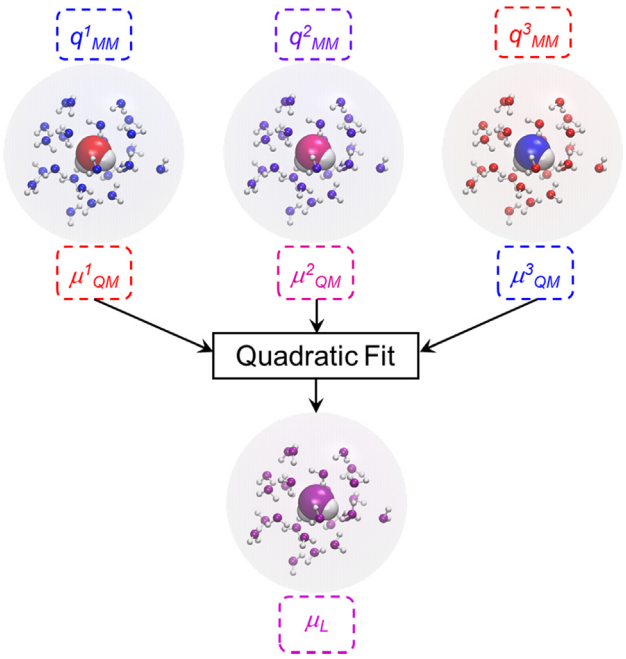

Fig. 1. Diagram showing the key idea behind the SCEE method. Each QM/MM calculation (top) calculates the dipole moment of the central QM molecule (μ_{QM}) when surrounded by a spherical cluster of MM molecules with fixed point charges (q_{MM}) of different magnitude (superscripts 1 to 3). In each calculation, the resulting QM dipole moment will be different from the surrounding MM dipole moments, represented here by using different colors for the QM and MM molecules. However, a relationship can be derived between those quantities to yield the self-consistent value of the liquid dipole moment (bottom), which is the same for QM and MM molecules. The process is repeated for a large number of configurations to yield an average liquid phase dipole moment.

and averaging over at least 150 independent molecular configurations, spaced by 20 ps, harvested from classical MD simulations (although we used 200 configurations as a conservative approach). We also showed that SCEE yielded liquid dipole moments in agreement with results from PE and AIMD calculations, while previous EE estimates could be made consistent with those of SCEE by approximately correcting for the shortcomings of the EE method [17].

Once the liquid phase dipole moments have been calculated, the different energetic components of polarization can be estimated using a simple sequence of single-point QM calculations starting from the fully polarized wave function of the central solute molecule. Specifically, we calculated the distortion contribution to the polarization energy (E_{Dist}) [36–38] from:

$$E_{Dist} = E_{QM} - E_{Vac} \quad (1)$$

where E_{QM} was obtained from the first self-consistent field (SCF) step of a single-point calculation of the polarized QM wave function in vacuum, and E_{Vac} corresponds to the energy of a reference solute molecule in vacuum. The distortion term quantifies the energy penalty paid by the molecule when moving from an unpolarized configuration in the gas phase to a fully polarized (and hence distorted) configuration in the liquid or solution phase.

The purely electronic contribution to the polarization energy (E_{Elec}) [16] is given by:

$$E_{Elec} = E_{\varepsilon_\infty} - E_{QM} \quad (2)$$

where $E_{\varepsilon\infty}$ was obtained from the first step of a single-point self-consistent reaction field (SCRF) calculation starting from the polarized QM wave function of the solute surrounded by a continuum with dielectric constant equal to the high-frequency dielectric constant of the solvent. For this single-point energy calculation, we used the SCIPCM model [39], as implemented in Gaussian 09 [40], with values of ε_∞ determined from the square of the experimental index of refraction at the sodium D-line frequency for each solvent (the values used in this work for the different alcohol molecules are shown in Table S1). This favorable (i.e. negative) energy term describes in an approximate way the interaction between the polarized solute wave function and the purely electronic degrees of freedom of the surrounding liquid [16]. The SCIPCM was preferred over other SCRF implementations [41] for this particular purpose because the solute cavity is determined self-consistently without the need to specify any parameters, whereas in most other approaches the energy would strongly depend on the cavity size and shape.

Finally, we can define a total polarization correction for non-polarizable force fields (E_{Tot}), introduced in the MDEC framework of Leontyev and Stuchebrukhov [16], as simply the sum of E_{Elec} and E_{Dist}. It is important to note that E_{Tot} is different from the total polarization energy, as defined by Orozco and co-workers [42,43] and normally denoted as E_{Pol}. E_{Tot} only considers the favorable (i.e. negative) energy contribution due to the electronic degrees of freedom of the liquid/solvent, whereas E_{Pol} considers *all* the favorable contributions of polarization, including the nuclear components. As a consequence, while E_{Pol} is always negative – in fact, it has been shown by ourselves [17] and others [42,43] to be approximately equal, within a first-order approximation, to $-E_{Dist} - E_{Tot}$ can be either positive or negative, but has been shown to be close to zero for water, due to a near complete cancellation of the distortion and electronic components [13,14,17]. For more details about the different energetic contributions to the polarization energy, the reader is referred to our earlier publication [17].

In this work, we applied the SCEE approach [17] to predict the liquid dipole moments of pure alcohols, as well as of alcohols solvated in different solvents. This required some adaptations of the method to deal with larger solute molecules (as compared to water used in our previous work [17]) and an extension to handle heterogeneous environments, i.e. where the solute and the solvent are different chemical species. These changes to the original method are described in detail below, and a step-by-step description of the full SCEE procedure (for both liquids and solutions) is provided in Supplementary Information.

In our previous work [17], we found that the results obtained for water were statistically independent of the choice of the MM model, provided it yielded a realistic representation of the thermodynamic and structural properties of the liquid. As a basis for the pure alcohol classical MD simulations, we used our recently developed PolCA model [44], which was designed to provide accurate predictions of both pure alcohol bulk properties and their solvation free energies in polar and non-polar solvents, by implicitly accounting for polarization effects through *post facto* analytical corrections [14,15,45,46]. A drawback of using a united-atom (UA) model like PolCA for this purpose, however, is that the aliphatic hydrogens are not explicitly represented. This is a potential problem because those atoms need to be in place in the central molecule for the QM optimization. To resolve this, we added "dummy" hydrogen atoms to the central "solute" molecule, which were bonded to the corresponding aliphatic carbon atoms by harmonic bond stretching, harmonic angle bending, and Ryckaert-Bellemans dihedral torsion terms with parameters taken from the OPLS all-atom (AA) model [47]. In other words, we ran an MD simulation of a pseudo-AA alcohol solute in a box of UA alcohol solvent molecules. We note that apart from those bonded potentials, the dummy atoms did not interact with any other atoms in the system, so the behavior of the original PolCA model for the bulk fluid was unchanged. However, this allowed us to obtain a series of input configurations for the QM part of the method where the aliphatic hydrogen atoms were placed at reasonable starting positions. As these were later optimized at the QM level in the subsequent QM/MM calculations, the precise nature of the bonded parameters is not critical.

When applying our original SCEE procedure to methanol, we observed that several configurations would not converge during the QM/MM optimization step. The problem was even worse for ethanol, since for that molecule more than 50% of the configurations were found not to converge. Upon further inspection, these convergence problems were deemed to be caused by the absence of a repulsive interaction to "shield" the bare solvent point charges in the original EE approach. In certain configurations, this caused the atoms of the central molecule to be unphysically distorted from their equilibrium positions, preventing convergence. To mitigate this issue, we decided to implement a new version of SCEE making use of the ONIOM methodology [48,49]. Instead of surrounding the central molecule with a cluster of bare point charges, as in our original method, we surrounded it by a cluster of MM molecules that included both Coulomb and Lennard-Jones interactions. The former were set up exactly as in the original SCEE method [17], while the latter were based on the same parameters as used in the MD simulations from which the configurations were harvested (i.e. the PolCA parameters in the case of alcohols). The only exception was that we included non-zero LJ parameters [44] for the hydroxyl hydrogen atom, which were absent in the classical PolCA force-field (and, indeed, in the vast majority of classical force fields that describe hydroxyl groups). After preliminary tests, we found that values of $\sigma = 0.027$ nm and $\varepsilon = 0.04$ kJ/mol created a sufficient repulsive "shield" to prevent atoms of the central molecule from approaching the surrounding (opposite sign) point charges at unphysical distances. Using this procedure, all configurations tested for all pure alcohol molecules were seen to converge. Moreover, the resulting liquid dipole moment of methanol obtained with ONIOM was in statistical agreement with the value obtained from the original procedure with bare point charges [17], but with-

Table 1
Dipole moments for methanol obtained from SCEE calculations using the original procedure with bare embedding charges and the improved procedure with ONIOM for the surrounding MM molecules. In all cases, the results are averages over 200 MD configurations, with the exception of those that did not converge or were identified as outliers by the IQR procedure (second and third rows, respectively). All QM calculations were done with the B3LYP functional and the aug-cc-pVTZ basis set. Dipoles are in Debye.

	Charges (Original)	ONIOM (Modified)
μ_L **(all data)**	2.66 ± 0.03	2.62 ± 0.03
Not converged	15	0
Number of outliers	7	5
μ_L **(outliers removed)**	2.61 ± 0.03	2.59 ± 0.03

out the convergence problems (see Table 1). Similar agreement was observed for SPC/E water (see Supplementary Information).

Despite the huge improvement in stability brought about by implementing the ONIOM procedure, a small number of configurations still led to rather large equilibrium dipole moments. This can be seen as a tail extending to large values in the distribution of dipole moments for methanol, shown in purple in Fig. 2. To prevent those configurations from biasing the calculation of average properties in the liquid state, we implemented a systematic outlier removal process based on the Interquartile Range (IQR) method [50]. More specifically, we identified all the configurations where the calculated dipole moment was outside the range between Q1 − 1.5 × IQR and Q3 + 1.5 × IQR, where Q1 is the first quartile and Q3 is the third quartile of the distribution (we note that there were very few, if any, points below the lower limit in the IQR procedure). For the methanol distribution, 5 configurations out of 200 (i.e. 2.5%) were identified using this criterion (see Table 1), and their dipole moment values are listed in Table S2. Table S2 shows that those configurations, highlighted in bold red font, led to very high distortion energies, often in excess of 100 kJ/mol. Most of those configurations also ranked at the top of the list of largest C-O or O-H distances, further indicating that the central molecule was undergoing unphysical distortion.

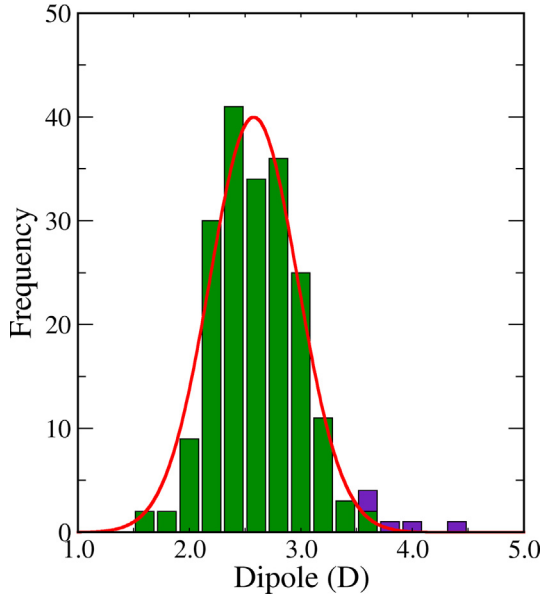

Fig. 2. Dipole moment distribution for liquid methanol. Results were sampled over 200 configurations obtained with the PolCA model, using SCEE at the B3LYP/aug-cc-pVTZ QM level of theory. The purple bars denote outliers removed through the IQR procedure, the green bars show the distribution after outlier removal, while the red line shows the best fit of the resulting distribution to a Gaussian expression.

Removing those few outliers led to a dipole moment distribution with the expected symmetric shape (green bars in Fig. 2), which conformed quite closely to a Gaussian function (red line in Fig. 2). We note that in our previous work, we had already identified a small tail extending to large values in the dipole moment distributions for water (see Fig. 7a of ref. [17]), although the effect was not as significant as observed here for methanol. In fact, after applying the same outlier removal procedure to the water dipole distributions from our previous paper, we observed only a very small decrease in the average value of the liquid dipole moment that was within the statistical uncertainty of the SCEE calculations (see Table S3). In Figure S2, we show the effects of outlier removal on the liquid phase dipole distributions for ethanol, 1-propanol, 2-propanol and t-butanol, which are qualitatively similar to the results shown in Fig. 2 for methanol.

One drawback of the ONIOM framework used in the new version of SCEE is that the wave function of the central QM molecule cannot be isolated from the surrounding MM region, hence precluding the calculation of polarization energies as described above (see Eqs. (1) and (2)). For that reason, we had to carry out an additional single-point calculation in a cluster of bare point charges, after optimization with ONIOM, from which the dipole moment, the energy and the wave function of the central molecule were extracted. Having to carry out this additional step does have a silver lining – because we are now separating the optimization of the central molecule (carried out with ONIOM) from the computation of the molecular dipole moment and energy (using standard EE), it was possible to reduce the level of theory of the optimization step, hence speeding up the calculations while maintaining the accuracy of the dipole and energy values. To test this, we tried optimizing with progressively decreasing basis set sizes, and found that the cc-pVTZ//aug-cc-pVTZ protocol (i.e., optimization run with ONIOM//single-point calculation with charges) was an optimal compromise between speed and accuracy (Table 2). While optimizing with the cc-pVDZ basis set led to reasonably accurate dipole moments and electronic energies, the distortion energy was significantly overestimated, leading to a total polarization correction that was too high by 1.1 kJ/mol, above the statistical uncertainty. In contrast, optimization with the cc-pVTZ basis set led to accurate dipoles and energies at a much lower computational cost than performing the optimization of the solute's atomic positions with the aug-cc-pVTZ basis set. Moreover, the improved ONIOM procedure was substantially faster than the original method based on bare charges, despite the need for an additional single-point calculation (Table 2). This is because the geometry of the central molecule was found generally to converge much faster with ONIOM.

Decoupling the optimization from the dipole calculation in the new SCEE procedure also opens up another possibility – using a higher level of theory in the latter step. In our previous work, it was found that diffuse functions needed to be used to obtain reliable dipole moments, in line with previous large-scale benchmark studies [51,52], and that the aug-cc-pVTZ basis set yielded a good compromise between accuracy and computational speed. However, because the full geometry optimization is now carried out with ONIOM at a lower level of theory (namely with the cc-pVTZ basis set, as discussed above), it is possible to use larger basis sets in the single-point step with bare point charges. We tested this possibility for water (configurations obtained from the SPC/E model [53], see [17] for details), using Dunning's augmented correlation-consistent basis sets [24,25] from double- up to quintuple-zeta. The resulting dipole moments and energies are shown in Table S4 and plotted as a function of basis set size in Fig. 3 and S3–S5.

As can be seen in Fig. 3, the dipole moment converges with basis set size, such that results obtained with triple-zeta basis sets and

Table 2
Dipole moments and polarization energies for methanol obtained from SCEE calculations with different optimization protocols. In all cases, the results are averages over ~200 MD configurations with the B3LYP functional. Dipoles are in Debye and energies in kJ/mol. Also shown is the average computational time, relative to the original calculation with bare point charges, for a single configuration undergoing all steps of the SCEE procedure with each protocol.

Method	QM Level	μ_L	E_{Elec}	E_{Dist}	E_{Tot}	Relative Time
Charges	aug-cc-pVTZ	2.61 ± 0.03	−24.1 ± 0.2	29.4 ± 0.7	5.3 ± 0.7	1.000
Oniom	aug-cc-pVTZ//aug-cc-pVTZ	2.59 ± 0.03	−24.3 ± 0.2	28.6 ± 0.6	4.3 ± 0.6	0.195
Oniom	cc-pVTZ//aug-cc-pVTZ	2.59 ± 0.03	−23.4 ± 0.2	27.6 ± 0.6	4.2 ± 0.6	0.138
Oniom	cc-pVDZ//aug-cc-pVTZ	2.61 ± 0.03	−23.9 ± 0.2	29.3 ± 0.6	5.4 ± 0.6	0.087

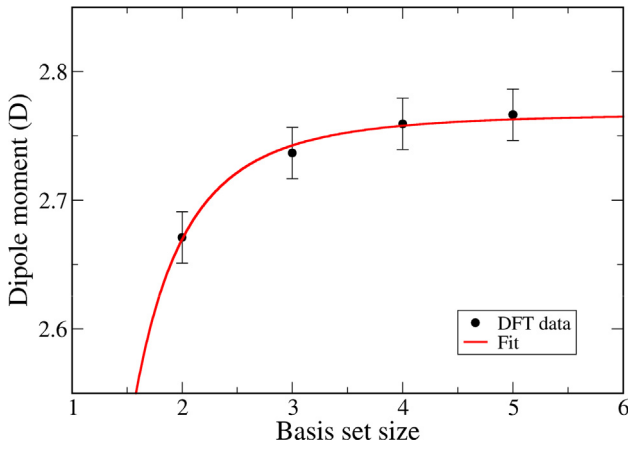

Fig. 3. Dipole moment of liquid water obtained from SCEE with different basis set sizes (i.e. aug-cc-pVXZ, with X from 2 to 5). Results were sampled over 200 configurations obtained with the SPC/E model, using the B3LYP functional. The red line shows a fit to Eq. (3).

beyond are statistically equivalent. It should be said that the computational effort increases significantly with basis set size, such that the calculations with aug-cc-pV5Z are at the limit of feasibility with current standard desktop computers, even for a small molecule like water. For methanol, we were able to run calculations using aug-cc-pV4Z (see discussion below) with an acceptable computational effort, while for ethanol and beyond, those calculations were found to be prohibitive; recall that one SCEE dipole determination requires 200 separate QM/MM calculations. As such, it would be advantageous to be able to extrapolate the dipole moments and energies to infinite basis set size based on calculations at lower levels of theory (namely at double-zeta and triple-zeta levels). We started by fitting the data to the following expression for extrapolation of correlated electronic structure calculations proposed by Truhlar et al. [54]:

$$X_n = X_\infty + A \times n^{-\alpha} \quad (3)$$

where X_n is the observable of interest at a given basis set size (n) and X_∞ is the value of that observable at the infinite basis set limit. We have fixed the exponent α to the recommended value of 3.4 [54], and used A and X_∞ as fitting parameters. Eq. (3) is able to accurately describe the evolution of the dipole moments, as shown by the red line in Fig. 3, and yield an estimate for the dipole moment of water at infinite basis set size. The fit to the different polarization energy components (Figures S3-S5) is also satisfactory.

This means that we can now use Eq. (3), together with results obtained at the triple- and double-zeta levels, to estimate the dipole moments and energies at infinite basis set for any molecule of interest. Because we are using only two points, an analytical expression can be derived as follows:

$$X_\infty = \frac{X_2 \times 3^{-\alpha} - X_3 \times 2^{-\alpha}}{3^{-\alpha} - 2^{-\alpha}} \quad (4)$$

In this equation, X represents the property of interest (dipole moment, energies, etc.). To test the accuracy of this approach, we estimated the dipole moment for methanol at the quadruple-zeta level from Eq. (3) after fitting only to the values obtained with triple- and double-zeta basis sets, and compared that estimate to the actual SCEE result obtained with the aug-cc-pV4Z basis set. The agreement was excellent: 2.601 ± 0.03 D from Eq. (3) compared with 2.603 ± 0.03 D from aug-cc-pV4Z calculations. It is important to reiterate that the dipoles and energies obtained at the triple-zeta level are already quite close to the infinite basis set limit, as shown in Fig. 3, S3-S5. However, the extrapolation procedure enables more accurate values to be obtained at a low additional computational cost, since double-zeta single-point calculations are computationally inexpensive. The accuracy of Truhlar's scheme for extrapolating results from DFT calculations was also demonstrated in previous studies [55,56], with a comparison against more complex power-law-type equations being provided in the Supporting Information of Ref. [56].

Before applying the SCEE procedure to alcohol molecules, we checked again the convergence of the resulting dipole moment with the size of the surrounding MM cluster. For methanol, the calculation converged for cluster radii above ~1.4 nm (see Figure S6), and therefore we retained the radius of 1.5 nm, as used previously for water [17], for subsequent analysis. However, for ethanol, the calculation only converged for radii above 1.6 nm (see Figure S7). This is due to the larger size of the central molecule, which requires a concomitantly larger solvation cluster. In fact, taking into account that convergence was observed for water beyond 1.2 nm [17], one would need to increase the radius of the cluster by ~0.2 nm for each carbon atom added to the solute chain – this is a useful "rule of thumb" if even larger molecules need to be analyzed. Based on this analysis, we used cluster radii for each solvent that were conservative enough to ensure complete convergence (see Table S5). We recall that the efficiency of the QM part of the calculation (by far the most time-consuming step of SCEE) is practically unaffected by the size of the MM cluster; the only additional cost is in the MD part, by simulating a larger box of liquid.

Another issue that arose during SCEE calculations for alcohol molecules has to do with conformational sampling in the gas phase. In our previous work [17], we used a single water molecule optimized in vacuum as a reference system for obtaining the gas-phase dipole moment (μ_G) and energy (E_{Vac}). For small molecules without torsional degrees of freedom involving chains of heavy atoms, this provides a realistic representation of the molecules' conformation state in the gas phase. Indeed, optimization of 200 methanol configurations extracted from molecular dynamics simulations at 25 °C led to the same energy minimum, with negligible variations in geometry, dipole moment and total energy using default optimization criteria. However, for larger molecules with at least three heavy atoms connected along a chain, i.e. ethanol and beyond, this is not the case. Analysis of the gas phase conformations of ethanol, for example, reveals the presence of two populations, corresponding to trans and gauche orientations of the C-C-O-H dihedral (Figure S8). Furthermore, at 25 °C, even conformations at the top of the energy barrier for rotation of this dihedral

have non-negligible probabilities of occurrence. This means that any estimate of gas phase properties (namely, dipole moments and energies, as required here) needs to sample over the entire conformational phase space of the molecule. This would allow for the gas phase properties to be calculated on an analogous basis to liquid phase properties, which are themselves based on statistical sampling over multiple configurations obtained from MD simulations at the temperature of interest.

To achieve this, we collected several evenly-spaced configurations from an MD simulation of a single alcohol molecule in the gas phase, under the same conditions as the pure liquid simulations. Each of those configurations was then optimized using our standard QM protocol, i.e. B3LYP/cc-pVTZ//aug-cc-pVTZ, but keeping all "backbone" dihedrals fixed (e.g., C-C-O-H for ethanol, C-C-O-H and C-C-C-O for 1-propanol, and so on). 200 configurations were sufficient to ensure convergence of both energy and dipole moment for ethanol (Fig. 4); however, for longer alcohols we used 400 configurations, since those properties were still fluctuating at ~200 configurations in the case of 1-propanol (Figure S9). Although these gas-phase optimizations are quite fast, and account for a relatively small fraction of the total computational effort of the SCEE method, we expect that more configurations may be needed for longer chain molecules, effectively placing an upper limit on the size of solute molecule that can be handled using this method.

In Table 3, we show the average gas phase dipole moments for all alcohol molecules tested here as "solutes", compared with experimental measurements. It is reassuring to note that all theoretical values are in reasonable agreement with experiment, and the accuracy of our chosen computational level of theory is supported by recent benchmark studies [51,52]. In particular, the results for the conformationally sampled dipole moments of the larger alcohols are much closer to experimental values than the dipole moments of the corresponding global energy minima (e.g., 1.58 D for ethanol and 1.48 D for 1-propanol). This supports our sampling procedure and confirms that it provides a realistic description of the environment experienced by those molecules in the gas phase.

As designed, the SCEE method self-consistently calculates the degree of polarization of the central (QM) molecule induced by the presence of the surrounding solvent (MM) molecules. Because the polarizing effect of the solvent is described through a set of point charges at the MM level, it only takes into account the nuclear contribution to the overall polarization. In contrast, if the full solute/solvent systems were described at the QM level, there would also be a contribution to polarization coming from purely

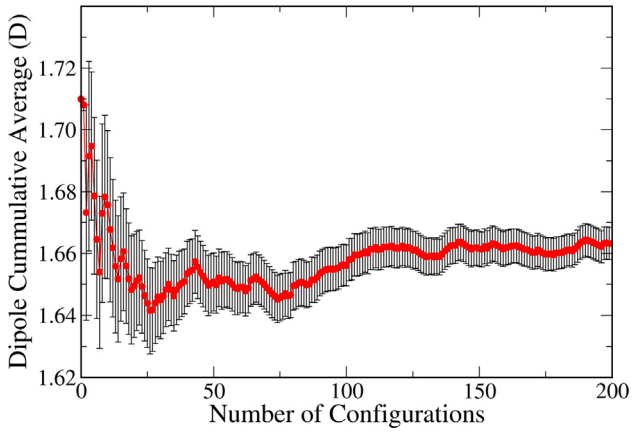

Fig. 4. Convergence of the gas phase dipole moment of ethanol as a function of the number of MD configurations used in the calculation of the average.

Table 3
Dipole moments (in Debye) for alcohols in the gas phase, obtained from experimental data and from our computational procedure. The computational results are an average of 400 independent configurations (200 for ethanol and 1 for methanol) obtained from an MD simulation of a single molecule in vacuum at 25 °C. Calculations were performed at the B3LYP/cc-pVTZ//aug-cc-pVXZ level of theory, where "X" denotes "D" or "T", and extrapolated to the infinite basis set limit.

Molecule	μ_{QM}	μ_{Exp}	Reference for experimental data
Methanol	1.65	1.70	[32–34]
Ethanol	1.66 ± 0.01	1.69 ± 0.03	[32–34]
1-Propanol	1.63 ± 0.01	1.68 ± 0.03	[32,33]
		1.58 ± 0.03	[34]
2-Propanol	1.69 ± 0.01	1.66 ± 0.03	[32,33]
		1.58 ± 0.03	[34]
t-Butanol	1.63 ± 0.01	1.67 ± 0.03	[33]
		1.66 ± 0.03	[34]

electronic degrees of freedom – i.e. due to fluctuations of the electronic clouds of the molecules in the absence of nuclear motion. This separation between nuclear and electronic contributions to polarization, based on the Born-Oppenheimer approximation, has been discussed in detail elsewhere [13,14,16,45]. Neglecting the electronic contribution is responsible, among other things, for systematic errors in predicting phase-change and electronic properties of fluids using classical non-polarizable models [13–15,45,46].

In the absence of a full QM description of the entire system, it has been shown that simple correction schemes can approximately account for the missing electronic effects in calculations that rely on classical interaction potentials [13–16,45,46]. In this spirit, we propose here a simple correction to the dipole moment obtained from SCEE calculations as described above. We define a correction factor based on the ratio of the dipole moments obtained from QM calculations of a single solute molecule in an electrostatic continuum model with two different dielectric constants: $\mu(\varepsilon)$ is obtained with the experimental dielectric constant of the solvent, which includes both nuclear and electronic contributions, while $\mu(\varepsilon - \varepsilon_\infty + 1)$ is obtained with a reduced dielectric constant that excludes the electronic contribution described by $\varepsilon_\infty - 1$ [15,45,46]. More precisely, the first calculation is carried out with an IEFPCM dielectric continuum model [57] with default settings for the solvent in question, while the second calculation uses the same model, but with a dielectric constant set manually to $\varepsilon - \varepsilon_\infty + 1$ (see Table S1 for a list of values used in these calculations). The corrected liquid phase dipole moment (μ_L) is then obtained from the SCEE dipole moment (μ_{SCEE}) by:

$$\mu_L = \mu_{SCEE} \times \frac{\mu(\varepsilon)}{\mu(\varepsilon - \varepsilon_\infty + 1)} \quad (5)$$

Two remarks should be made regarding Eq. (5). First of all, it has been shown by several authors [17,58–61] that uniform dielectric continuum models are not able to accurately describe the full extent of polarization in polar solvents, due to their inability to describe local solvation interactions (e.g. solute–solvent hydrogen bonds). However, the effectiveness of the correction scheme in Eq. (5) does not rely on an ability of the dielectric continuum model to describe the full polarization of the solute, but merely the electronic contribution of polarization alone. In other words, it is the ratio of the two dipoles on the right-hand side that matters, not the absolute value of each of them. Indeed, this approach, whereby the purely electronic polarization is represented by a simple dielectric continuum model, has been used successfully in the past [13,14,16,44]. The second remark is that the magnitude of the correction (in other words, the relative contribution of electronic effects to the total polarization) will decrease dramatically with an increase in the dielectric constant of the solvent. For example, the correction factor for water is 1.00005, which means that electronic effects are completely negligible, as is to be expected –

which means, incidentally, that our previous (uncorrected) SCEE calculations for water [17] retain their validity. This is also the case for small alcohols (e.g. the correction is 1.001 for 1-propanol), but for larger alcohols it starts to become significant (e.g. it is ~1.01 for decanol). More importantly, for non-polar solvents like alkanes, electronic effects will play a significant role and cannot be ignored. We will return to this point later in the paper when we discuss methanol solvation in different solvents.

As discussed previously, one of the objectives of the present work was to generalize the SCEE method so that it could also be applied to solutions – i.e. where the solute and solvent are different molecules. In principle, the self-consistent determination of the liquid dipole (Fig. 1) cannot be implemented in those situations, since the molecular identity (and hence the dipole moment) of the QM and MM molecules is different. The extension of SCEE to solution environments, therefore, is based on the assumption that the average dipole moment of the solvent molecules is the same in the solution and in the pure liquid. In other words, we estimate the dipole moment and polarization energies of a solute molecule by surrounding it with MM solvent molecules that have the same dipole moment as the equilibrium value in the pure solvent phase. While this assumption may not be strictly correct in the close vicinity of the solute molecule, it is likely that *on average* these effects will cancel out in the calculation of the solute properties. The assumption, in fact, is consistent with the spirit of the SCEE method, which relies on describing the polarization environment explicitly (through surrounding solvent molecules that are polarized) but in a mean-field sense (by assuming that all surrounding molecules have the same dipole moment). In our previous work [17], we showed that this was a reasonable assumption for pure liquids, since the results obtained using SCEE were statistically equivalent to calculations using a polarizable model for the MM part of the system, where each solvent molecule had a unique dipole moment.

In this work, we further test this hypothesis by calculating the dipole moment of methanol in methanol solvent under the above assumption, and comparing it against the result of the full SCEE procedure. Specifically, we start from a standard pure liquid methanol SCEE calculation, in which we find the equilibrium dipole moment for each individual configuration (as described previously). Averaging over all of these configurations gives us the average liquid dipole moment for methanol (e.g. as reported in Table 2). We then run another QM/MM calculation, but now assuming that the dipole moment of all the surrounding molecules in *all configurations* is equal to the average value. Using this approach, the average dipole moment of the central molecule, i.e. the "solute", is 2.57 ± 0.02 D, which is statistically consistent with the average liquid phase dipole moment of 2.59 ± 0.03 D, further supporting our assumption. This gives us confidence that we can apply our method to calculate dipole moments and polarization energies of heterogeneous systems, i.e. when the solute and solvent are different species.

2.2. Computational details

MD simulations were performed using GROMACS version 5.1.2 [62,63] and the Verlet leap-frog algorithm [64] to integrate the equations of motion with a time step of 2 fs. Simulation boxes were cubic with periodic boundary conditions in all directions and sizes adjusted to be larger than the optimal cutoff radius for the QM/MM clusters, described in Section 2.1 (Table S5 lists the average box sizes and QM cluster radii for each system studied here). Except where noted, liquid phase simulations were run in the *NpT* ensemble, with temperature controlled at 298 K by a Nosé-Hoover thermostat [65,66] with a coupling constant of 2 ps, and pressure controlled at 1 bar by a Parrinello-Rahman barostat [67] with a coupling constant of 1 ps and a compressibility of 4.5×10^{-5} m^3/bar. A cut-off of 1 nm was used for the Lennard-Jones potential, with long-range dispersion corrections added to both energy and pressure, while long-range electrostatic interactions were accounted for by using the particle-mesh Ewald method [68]. Gas-phase MD simulations made use of the same protocol, except that the simulation boxes contained a single molecule, no periodic boundary conditions or cut-off radius were applied (hence replicating a vacuum environment), and no barostat was used, i.e. simulations were run in the *NVT* ensemble. All MD simulations were run for 5 ns, with 200 configurations being harvested from the last 4 ns of each run. The exceptions were the gas-phase simulations of longer alcohols, which were run for 10 ns. We note that the spacing of 20 ps between each successive configuration ensures that they are sufficiently uncorrelated to allow efficient statistical sampling, as discussed in detail by Coutinho et al. [69].

The QM/MM calculations were performed with Gaussian 09 [40] using the B3LYP exchange–correlation functional [35], as this was shown in our previous work [17] and in recent benchmark studies [51,52] to provide good accuracy in predicting molecular dipole moments. Geometry optimizations were carried out with Dunning's cc-pVTZ basis set [24], while subsequent single-point calculations to obtain dipole moments and energies made use of the aug-cc-pVDZ and aug-cc-pVTZ basis sets [25], as described in Section 2.1. It is well known that large basis sets that include diffuse functions are needed to provide accurate values for the electronic properties of molecules [51,52]. Calculations to derive the correction factor described in Eq. (5) made use of the IEFPCM method [57] under the self-consistent reaction field formalism [41] with the default parameters from Gaussian 09 (i.e., scrf=(solvent=*SOLVENT*,iefpcm,read) and ε or $(\varepsilon - \varepsilon_\infty + 1)$ values from Table S1). Single-point calculations to determine the electronic contribution to the polarization energy (Eq. (2)) were carried out with the SCIPCM method [39]. Scripts for processing Gaussian calculations are provided in Supplementary Information. Input files for all MD simulations, as well as the in-house codes used to convert GROMACS output configurations to the Gaussian input file format, are provided as Supplementary Material (see Data Statement).

3. Results and discussion

We applied the SCEE method, with the modifications described above, to five pure alcohol systems: methanol, ethanol, 1-propanol, 2-propanol, and t-butanol. The improvements in the method allowed us to study much larger systems than were previously possible, although the computational time still increases significantly with the size of the solute. This is because the main computational effort is in the QM stage of the process, and that scales with the size of the QM region (i.e. the solute molecule).

The first thing to notice in terms of results (see Table 4) is that the dipole moment of alcohols is significantly enhanced in the liquid phase relative to the gas phase values, by between ~0.8 and ~1.0 D. This level of enhancement is actually quite similar to the one observed for water in our previous study [17], which is perhaps surprising given that the dielectric constant of alcohols is much lower than that of water (see Table S1). This suggests that for hydrogen-bonding liquids like alcohols and water, local interactions have the strongest effect on the degree of polarization, with the dielectric constant of the liquid playing a secondary role. In terms of the energetic contributions to polarization, we again observe a very strong positive distortion contribution, largely cancelled out by a strong negative electronic contribution, leading to

Table 4
Dipole moments and polarization energies for pure alcohols obtained from SCEE calculations. The results are averages over ~200 MD configurations[a] optimized at the B3LYP/cc-pVTZ level of theory, extrapolated to the infinite basis set limit using aug-cc-pVDZ and aug-cc-pVTZ single-point calculations, and corrected for electronic polarization effects. Dipoles are in Debye and Energies in kJ/mol.

	Methanol	Ethanol	1-Propanol	2-Propanol	t-Butanol
μ_G	1.65	1.66 ± 0.01	1.63 ± 0.01	1.69 ± 0.01	1.63 ± 0.01
μ_L	2.61 ± 0.03	2.56 ± 0.03	2.56 ± 0.03	2.71 ± 0.03	2.41 ± 0.03
$\Delta\mu$	0.96 ± 0.03	0.90 ± 0.03	0.93 ± 0.03	1.02 ± 0.03	0.78 ± 0.03
E_{Elec}	−23.7 ± 0.4	−22.3 ± 0.4	−22.7 ± 0.4	−23.9 ± 0.4	−17.1 ± 0.2
E_{Dist}	28.7 ± 1.2	28.5 ± 1.2	29.3 ± 1.0	36.4 ± 1.2	16.0 ± 0.8
E_{Tot}	5.0 ± 1.2	6.2 ± 1.2	6.6 ± 1.0	12.5 ± 1.2	−1.1 ± 0.8

[a] Exceptions are gas-phase methanol that considered the ground state only, and gas-phase 1-propanol, 2-propanol and t-butanol that considered 400 configurations.

relatively small total polarization corrections. However, for most alcohols (with the exception of t-butanol), the distortion contribution seems to somewhat dominate, leading to net positive polarization corrections.

We can look in more detail at the distributions of liquid-phase dipole moments for all the alcohols listed in Table 4. The raw data is shown in Figures 2 and S2, but in Fig. 5 we show only the fitted Gaussian distributions (see equation S1 and Table S6 for fitting parameters) which allow for a clearer comparison between the different systems. As we can see, the distributions for methanol, ethanol and 1-propanol are quite similar, and therefore the average dipole moments are statistically indistinguishable (Table 4). This seems to suggest that for primary alcohols the liquid dipole moment is not strongly dependent on the alkyl chain length, which supports our hypothesis that local solvation effects have great importance on the degree of polarization. All the linear-chain alcohols are able to form strong hydrogen bonds with surrounding molecules; in fact, even though their long-range structure may differ, the geometry and number of hydrogen bonds in the first solvation shell are practically independent of chain length [70–72]. This means they are likely to be polarized to a similar degree in the liquid state, even though the surrounding environment away from the solute becomes less polar in a mean-field sense as the chain length increases.

The results for t-butanol support this interpretation, since the bulkier side chain in that highly branched molecule can inhibit the formation of hydrogen bonds with the central molecule [73], hence lowering the extent of local polarization. As a consequence, the dipole moment distribution of t-butanol in the liquid phase is shifted to significantly lower values (Fig. 5) and the average liquid dipole moment is significantly lower than for the linear alcohols, even though their gas phase dipole moments are very similar. In contrast, the distribution for 2-propanol is shifted to slightly higher

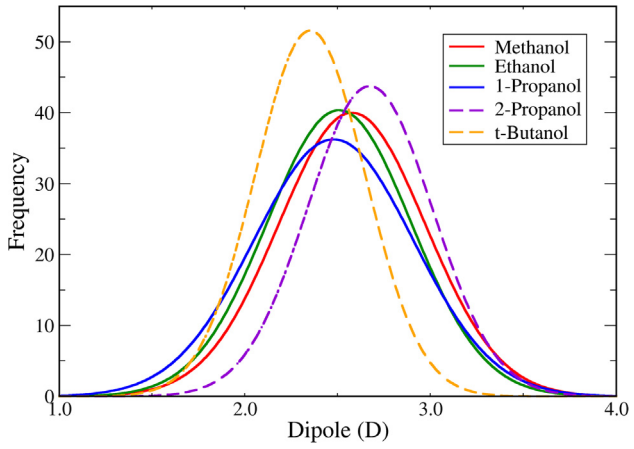

Fig. 5. Fits of a Gaussian function to the dipole moment distributions for the pure alcohol liquids studied here. The raw distributions are shown in Fig. 2 and Figure S2.

dipole moment values, and the average liquid dipole shows a slight but statistically significant increase from the value for 1-propanol. Although this effect is quite subtle, it may be related to an increased strength of the hydrogen bonds in 2-propanol relative to 1-propanol [74], since their first solvation shells are quite similar in the corresponding pure liquids [75,76].

With knowledge of the dipole moment of the pure liquids, we can now apply our extended SCEE method to calculate the dipole moment in heterogeneous systems, i.e. where the solute and solvent are different molecules. We began by calculating the dipole moment of methanol in ethanol and in 1-propanol, and obtained values that were practically indistinguishable from those of the corresponding pure solvents – i.e. we obtained 2.55 ± 0.03 D for methanol in ethanol, compared to 2.56 ± 0.03 D for pure ethanol, and 2.56 ± 0.03 D for methanol in propanol, compared to 2.56 ± 0.03 D for pure propanol. Furthermore, the dipole moment of ethanol in methanol is 2.60 ± 0.03 D, which is essentially the same as the value of 2.61 ± 0.03 D we obtained for pure methanol. These results suggest that, at least for primary alcohols, the liquid dipole moment is independent of the chain length of the solute. This once again reinforces our hypothesis that the degree of polarization of alcohols is primarily due to solute–solvent hydrogen bond formation.

From a methodological point of view, the above observation implies that we should be able to estimate the liquid dipole moment of any primary alkyl alcohol with reasonable accuracy by using methanol as a proxy for that alcohol in the QM region, thus introducing massive savings in computational time and allowing us to extend our approach to much larger molecules. In other words, we have applied the full "pure liquid" SCEE procedure to solutions of methanol in linear primary alcohols of progressively increasing chain length; based on the results discussed above, the calculated equilibrium dipole moment in those solutions constitutes a reasonable estimate of the dipole moment of the pure solvent. The results of these calculations are shown in Fig. 6 as a function of the chain length of the solvent.

As we can see from the closed circles in Fig. 6a, the liquid dipole moment of linear alkyl alcohols seems to decrease slightly for smaller molecules and then seems to stabilize for longer chain lengths (although the noise in the data increases at longer chain lengths as well). This figure also confirms that simple continuum solvation models lead to a very strong underestimation of the degree of polarization, with SCEE yielding dipole moments that are at least 0.5 D larger than values obtained from IEFPCM calculations (open triangles in Fig. 6a). Continuum models describe solvation effects in a mean-field sense, by representing the solvent as a dielectric continuum, and hence are not able to capture any contributions to polarization due to local interactions. In contrast, SCEE can describe both local and long-range contributions to polarization. By comparing the results of the two calculations, we can estimate that local effects, such as hydrogen-bond formation, account for more than half of the dipole enhancement of alcohols in the liquid phase.

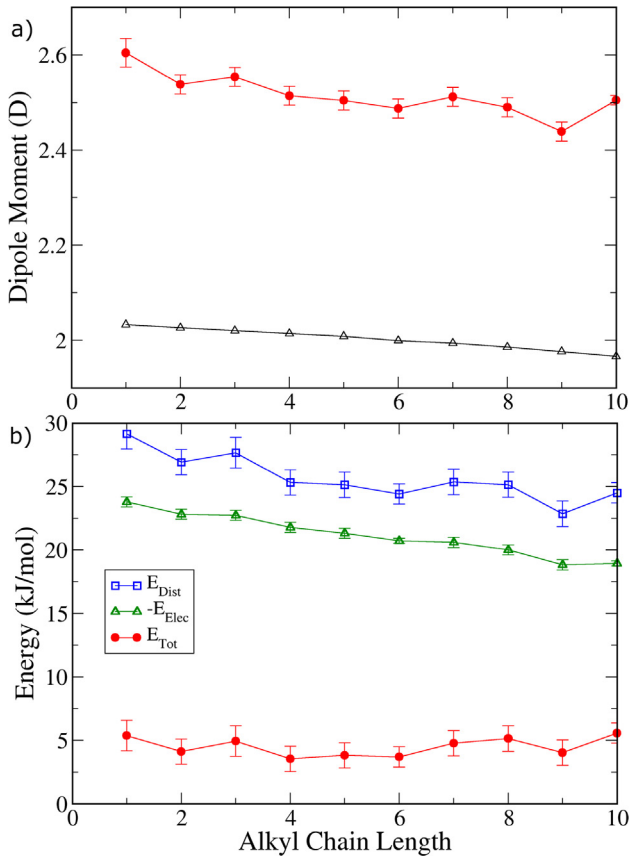

Fig. 6. Change in dipole moment (a) and polarization energies (b) for linear alkyl alcohols as a function of the number of carbon atoms in the chain. In panel (a), the full red circles show results of the full SCEE procedure, while the open black triangles show dipole moments obtained from IEFPCM calculations of a single alcohol molecule in a continuum dielectric solvent. The SCEE results were obtained by using a methanol solute (i.e. QM) molecule as a proxy for the corresponding linear alcohol.

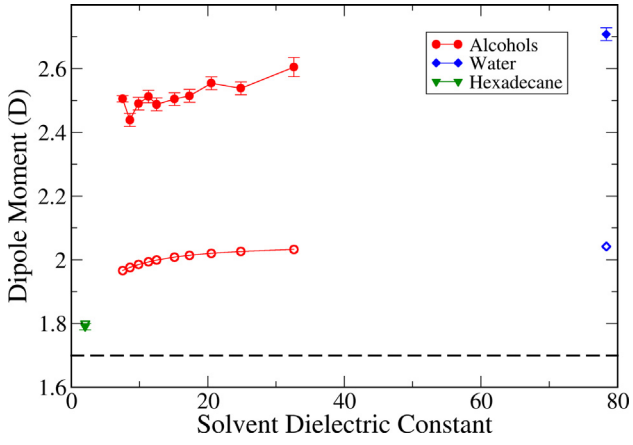

Fig. 7. Dipole moment of methanol in different solvents as a function of the dielectric constant of the solvent. Full symbols are results obtained with SCEE, while open symbols show results obtained with the IEFPCM continuum solvation model. The dashed line shows the dipole moment of methanol in the gas phase.

In Fig. 6b, we show how the energetic components of polarization change with alcohol chain length. Interestingly, although the magnitude of both the distortion and electronic energies decreases with chain length (note that the negative of the electronic energy is shown in Fig. 6b for clarity of representation), they cancel each other out to nearly the same extent for all molecules, leading to an approximately constant net polarization correction. This observation is important for force field development purposes [44], since it suggests that the same correction can be used for a homologous series of molecules with increasing alkyl chain length.

We have also calculated the dipole moment of methanol in water and hexadecane solvents, meant to represent extreme cases of very polar and non-polar environments, respectively. For the water calculations, we used the SPC/E model [53] with a dipole moment of 2.79 D determined from our improved SCEE procedure (see Table S3). For hexadecane, we used our PolCA model [77], which is based on the UA approach and corrects systematic deviations observed in previous UA models for the solvation free energy of long-chain alkanes [78]. Because UA models of alkanes have no point charges (i.e. each CH_x group is considered as a neutral interaction site), the electrostatic contribution of the surrounding solvent to the polarization of the solute is strictly zero. To test whether this assumption is realistic, we also calculated the dipole moment of methanol solvated in the OPLS-AA model [47] of hexadecane, which does have point charges on each solvent atom. The results for the dipole moment of methanol in both solvent models were statistically indistinguishable (1.79 ± 0.01 D for the UA PolCA model and 1.80 ± 0.01 D for the OPLS-AA model), which confirms that the contribution of Coulomb electrostatics to the solvation in hexadecane (and, by implication, in other pure alkane solvents) is negligible.

In Fig. 7, we plot the dipole moment of methanol in water and hexadecane, as well as in the linear alcohols considered previously (i.e. up to 1-decanol), as a function of the dielectric constant of the solvent. We also show, as open symbols, the corresponding results obtained using the IEFPCM continuum solvation model. The SCEE dipole moments for water and all the alcohols are significantly higher than those obtained with the implicit solvent model, once again confirming the importance of capturing local effects to obtain a realistic description of polarization in polar solvents. Conversely, the SCEE dipole moment of methanol in hexadecane is practically the same as that of the continuum model. In completely non-polar solvents like alkanes, local interactions play a negligible role in the polarization process; instead, most of the contribution to the polarization of the solute comes from purely electronic effects, which are well described in both SCEE and IEFPCM. Interestingly, although the methanol dipole moment decreases slightly with increasing alkyl chain length of the alcohol solvent (see also Fig. 6a), it remains well above the IEFPCM values, even for the rather long 1-decanol solvent. Even though the dielectric constant of 1-decanol is ~7.5 and this liquid can be classed as moderately polar at best, it is still able to significantly polarize the methanol solute. This is because 1-decanol is still able to form hydrogen bonds with methanol, which strongly contribute to the solute polarization.

It therefore seems that the degree of polarization of methanol in different solvents can be described by a combination of two effects: i) mean-field polarization due to the bulk dielectric constant of the solvent; 2) local polarization due primarily to hydrogen bond formation in the first solvation shell. For non-polar solvents like alkanes, the former dominates completely, since there are no local effects to speak of. For water and alcohol solvents, however, both effects play a role, but the latter is dominant, such that the liquid dipole curves are displaced vertically by a nearly constant amount from the predictions obtained from dielectric continuum models (which entirely neglect the local solvation effects). It would be interesting to see to which extent this hypothesis is applicable to solvents with intermediate polarity and/or propensity to form hydrogen bonds. We intend to carry out such studies in the near future.

4. Conclusions

In this paper, we have presented an improvement and generalization of our Self-Consistent Electrostatic Embedding method for calculation of liquid-phase dipole moments and polarization energies. These improvements allow for the method to be applied to much larger molecules than previously possible, with reasonable computational resources, as well as to calculate the dipole moments of heterogeneous solutions – i.e. when the solute and solvent are different species. We applied SCEE to calculate the liquid phase dipole moment of several pure alcohols, including primary, secondary and tertiary alcohols, and showed that the dipole moment enhancement in the liquid phase, relative to the gas-phase, is quite significant – in fact, it is approximately the same as observed in pure water, i.e. ∼0.9 D. This is much higher than predictions obtained from dielectric continuum models, due to the fact that the latter neglect local solvation effects in the calculation. In fact, the relative differences in dipole moment between the different alcohol molecules can be qualitatively rationalized on the basis of hydrogen bond formation and strength in the first solvation shell.

When applied to heterogeneous mixtures, our results show, first of all, that the length of the alkyl chain of the solute has a negligible effect on the liquid dipole moment. This allows us to estimate the pure liquid dipole moments of long-chain alcohols by running full SCEE calculations using methanol as a proxy solute. We aim to confirm if this effect is also observed in other homologous series of compounds. Our results for methanol solvation in alcohols, alkanes and water can be described as arising from a combination of two effects: mean-field polarization described by the bulk dielectric constant of the solvent and local polarization caused by hydrogen bond formation. The relative importance of these two effects depends on the nature of the solvent – bulk effects dominate in alkanes, while hydrogen bond formation is much more important in water and alcohols. We intend to test this hypothesis more systematically in the future, by carrying out SCEE calculations over a wider variety of solute/solvent pairs, including molecules with different degrees of polarity and hydrogen bond formation capabilities.

5. Data Statement

All data underpinning this publication are openly available from the University of Strathclyde KnowledgeBase at https://doi.org/10.15129/23867c83-ad24-48a9-aa83-4803ee126e7b.

CRediT authorship contribution statement

Miguel Jorge: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Supervision, Project administration. **José R.B. Gomes:** Conceptualization, Methodology, Software, Validation, Investigation, Resources, Writing – review & editing. **Maria Cecilia Barrera:** Methodology, Validation, Formal analysis, Investigation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

JRBG would like to acknowledge funding from project CICECO-Aveiro Institute of Materials, UIDB/50011/2020, UIDP/50011/2020 and LA/P/0006/2020, financed by national funds through the FCT/MEC (PIDDAC). MCB acknowledges the University of Strathclyde for a PhD studentship.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.molliq.2022.119033.

References

[1] H.B. Aditiya, T.M.I. Mahlia, W.T. Chong, H. Nur, A.H. Sebayang, Second generation bioethanol production: a critical review, Renew. Sustain. Energy Rev. 66 (2016) 631–653, https://doi.org/10.1016/j.rser.2016.07.015.
[2] N.S. Shamsul, S.K. Kamarudin, N.A. Rahman, N.T. Kofli, An overview on the production of bio-methanol as potential renewable energy, Renew. Sustain. Energy Rev. 33 (2014) 578–588, https://doi.org/10.1016/j.rser.2014.02.024.
[3] F. Dalena, A. Senatore, A. Marino, A. Gordano, M. Basile, A. Basile, Methanol production and applications: an overview, in: A. Basile, F. Dalena (Eds.), Methanol, Elsevier, Amsterdam, 2018: pp. 3–28. https://doi.org/10.1016/B978-0-444-63903-5.00001-7
[4] A.M. Bahmanpour, A. Hoadley, A. Tanksale, Critical review and exergy analysis of formaldehyde production processes, Rev. Chem. Eng. 30 (2014), https://doi.org/10.1515/revce-2014-0022.
[5] R.V. Afonso, J.D. Gouveia, J.R.B. Gomes, Catalytic reactions for H_2 production on multimetallic surfaces: a review, J. Phys. Energy 3 (3) (2021) 032016, https://doi.org/10.1088/2515-7655/ac0d9f.
[6] C. Hansch, A. Leo, Exploring QSAR: Fundamentals and Applications in Chemistry and Biology, Volume 1, American Chemical Society, Washington DC, 1995.
[7] C. Meyers, R. Kass, D. Goldenberg, J. Milici, S. Alam, R. Robison, Ethanol and isopropanol inactivation of human coronavirus on hard surfaces, J. Hosp. Infect. 107 (2021) 45–49, https://doi.org/10.1016/j.jhin.2020.09.026.
[8] A. Vrhovšek, O. Gereben, A. Jamnik, L. Pusztai, Hydrogen bonding and molecular aggregates in liquid methanol, ethanol, and 1-propanol, J. Phys. Chem. B 115 (46) (2011) 13473–13488, https://doi.org/10.1021/jp206665w.
[9] D.L. Wertz, R.K. Kruh, Reinvestigation of the structures of ethanol and methanol at room temperature, J. Chem. Phys. 47 (1967) 388–390, https://doi.org/10.1063/1.1711905.
[10] J.K. Vij, C.J. Reid, M.W. Evans, Molecular dynamics of methanol, Mol. Phys. 50 (5) (1983) 935–947, https://doi.org/10.1080/00268978300102771.
[11] N.M. Garrido, A.J. Queimada, M. Jorge, E.A. Macedo, I.G. Economou, 1-Octanol/water partition coefficients of n -alkanes from molecular simulations of absolute solvation free energies, J. Chem. Theory Comput. 5 (9) (2009) 2436–2446, https://doi.org/10.1021/ct900214y.
[12] I. Bakó, P. Jedlovszky, G. Pálinkás, Molecular clusters in liquid methanol: a Reverse Monte Carlo study, J. Mol. Liq. 87 (2-3) (2000) 243–254, https://doi.org/10.1016/S0167-7322(00)00124-0.
[13] A.W. Milne, M. Jorge, Polarization corrections and the hydration free energy of water, J. Chem. Theory Comput. 15 (2) (2019) 1065–1078, https://doi.org/10.1021/acs.jctc.8b0111510.1021/acs.jctc.8b01115.s001.
[14] I.V. Leontyev, A.A. Stuchebrukhov, Electronic polarizability and the effective pair potentials of water, J. Chem. Theory Comput. 6 (10) (2010) 3153–3161, https://doi.org/10.1021/ct1002048.
[15] C. Vega, Water: one molecule, two surfaces, one mistake, Mol. Phys. 113 (9-10) (2015) 1145–1163, https://doi.org/10.1080/00268976.2015.1005191.
[16] I. Leontyev, A. Stuchebrukhov, Accounting for electronic polarization in non-polarizable force fields, Phys. Chem. Chem. Phys. 13 (7) (2011) 2613, https://doi.org/10.1039/c0cp01971b.
[17] M. Jorge, J.R.B. Gomes, A.W. Milne, Self-consistent electrostatic embedding for liquid phase polarization, J. Mol. Liq. 322 (2021) 114550, https://doi.org/10.1016/j.molliq.2020.114550.
[18] C.J.F. Böttcher, Theory of Electric Polarization, Elsevier, Houston, 1952, https://linkinghub.elsevier.com/retrieve/pii/C20090155794.
[19] Y.S. Badyal, M.-L. Saboungi, D.L. Price, S.D. Shastri, D.R. Haeffner, A.K. Soper, Electron distribution in water, J. Chem. Phys. 112 (21) (2000) 9206–9208, https://doi.org/10.1063/1.481541.
[20] R. Car, M. Parrinello, Unified approach for molecular dynamics and density-functional theory, Phys. Rev. Lett. 55 (22) (1985) 2471–2474, https://doi.org/10.1103/PhysRevLett.55.2471.
[21] H.M. Senn, W. Thiel, QM/MM Methods for biomolecular systems, Angew. Chemie Int. Ed. 48 (7) (2009) 1198–1229, https://doi.org/10.1002/anie.200802019.

[22] J. Gao, X. Xia, A priori evaluation of aqueous polarization effects through Monte Carlo QM-MM simulations, Science 258 (5082) (1992) 631–635, https://doi.org/10.1126/science.1411573.
[23] M.E. Martín, M.L. Sánchez, F.J. Olivares del Valle, M.A. Aguilar, A theoretical study of liquid alcohols using averaged solvent electrostatic potentials obtained from molecular dynamics simulations: Methanol, ethanol and propanol, J. Chem. Phys. 116 (4) (2002) 1613–1620, https://doi.org/10.1063/1.1430253.
[24] T.H. Dunning, Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, J. Chem. Phys. 90 (2) (1989) 1007–1023, https://doi.org/10.1063/1.456153.
[25] R.A. Kendall, T.H. Dunning, R.J. Harrison, Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions, J. Chem. Phys. 96 (9) (1992) 6796–6806, https://doi.org/10.1063/1.462569.
[26] M. Pagliai, G. Cardini, R. Righini, V. Schettino, Hydrogen bond dynamics in liquid methanol, J. Chem. Phys. 119 (13) (2003) 6655–6662, https://doi.org/10.1063/1.1605093.
[27] J.-W. Handgraaf, E.J. Meijer, M.-P. Gaigeot, Density-functional theory-based molecular simulation study of liquid methanol, J. Chem. Phys. 121 (20) (2004) 10111–10119, https://doi.org/10.1063/1.1809595.
[28] I.-F. Kuo, C.J. Mundy, M.J. McGrath, J.I. Siepmann, Structure of the methanol liquid−vapor interface: a comprehensive particle-based simulation study, J. Phys. Chem. C 112 (39) (2008) 15412–15418, https://doi.org/10.1021/jp8037126.
[29] N. Sieffert, M. Bühl, M.-P. Gaigeot, C.A. Morrison, Liquid methanol from DFT and DFT/MM molecular dynamics simulations, J. Chem. Theory Comput. 9 (1) (2013) 106–118, https://doi.org/10.1021/ct300784x.
[30] S. Grimme, Semiempirical GGA-type density functional constructed with a long-range dispersion correction, J. Comput. Chem. 27 (15) (2006) 1787–1799, https://doi.org/10.1002/jcc.20495.
[31] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. 132 (15) (2010) 154104, https://doi.org/10.1063/1.3382344.
[32] R.D. Nelson, D.R. Lide, A.A. Maryott, Selected Values of Electric Dipole Moments for Molecules in the Gas Phase, U.S. National Bureau of Standards, Washington DC, 1967.
[33] C.L. Yaws, Thermophysical Properties of Chemicals and Hydrocarbons, 2nd ed., Elsevier, Oxford, England, 2014. https://doi.org/10.1016/C2013-0-12615-3.
[34] D.R. Lide (Ed.), CRC Handbook of Chemistry and Physics, CRC Press, Boca Raton, Internet V, 2005.
[35] A.D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys. 98 (7) (1993) 5648–5652, https://doi.org/10.1063/1.464913.
[36] J.L.F. Abascal, C. Vega, A general purpose model for the condensed phases of water: TIP4P/2005, J. Chem. Phys. 123 (23) (2005) 234505, https://doi.org/10.1063/1.2121687.
[37] W.C. Swope, H.W. Horn, J.E. Rice, Accounting for polarization cost when using fixed charge force fields. I. Method for computing energy, J. Phys. Chem. B. 114 (26) (2010) 8621–8630, https://doi.org/10.1021/jp911699p.
[38] C. Chipot, Rational determination of charge distributions for free energy calculations, J. Comput. Chem. 24 (4) (2003) 409–415, https://doi.org/10.1002/jcc.10207.
[39] J.B. Foresman, T.A. Keith, K.B. Wiberg, J. Snoonian, M.J. Frisch, Solvent Effects. 5. Influence of cavity shape, truncation of electrostatics, and electron correlation on ab initio reaction field calculations, J. Phys. Chem. 100 (40) (1996) 16098–16104, https://doi.org/10.1021/jp960488j.
[40] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A.J. Montgomery, J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N.J. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J. V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian 09, Revision B.01, Gaussian, Inc., Wallingford CT. (2009).
[41] J. Tomasi, B. Mennucci, R. Cammi, Quantum mechanical continuum solvation models, Chem. Rev. 105 (8) (2005) 2999–3094, https://doi.org/10.1021/cr9904009.
[42] M. Orozco, F.J. Luque, D. Habibollahzadeh, J. Gao, The polarization contribution to the free energy of hydration, J. Chem. Phys. 102 (15) (1995) 6145–6152, https://doi.org/10.1063/1.469348.
[43] F.J. Luque, M. Orozco, Semiclassical-continuum approach to the electrostatic free energy of solvation, J. Phys. Chem. B 101 (28) (1997) 5573–5582, https://doi.org/10.1021/jp9617229.
[44] M.C. Barrera, M. Jorge, A polarization-consistent model for alcohols to predict solvation free energies, J. Chem. Inf. Model. 60 (3) (2020) 1352–1367, https://doi.org/10.1021/acs.jcim.9b0100510.1021/acs.jcim.9b01005.s001.
[45] M. Jorge, L. Lue, The dielectric constant: Reconciling simulation and experiment, J. Chem. Phys. 150 (8) (2019) 084108, https://doi.org/10.1063/1.5080927.
[46] J. Cardona, M. Jorge, L. Lue, Simple corrections for the static dielectric constant of liquid mixtures from model force fields, Phys. Chem. Chem. Phys. 22 (38) (2020) 21741–21749, https://doi.org/10.1039/D0CP04034G.
[47] W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, J. Am. Chem. Soc. 118 (45) (1996) 11225–11236, https://doi.org/10.1021/ja9621760.
[48] F. Maseras, K. Morokuma, IMOMM: A new integratedab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states, J. Comput. Chem. 16 (9) (1995) 1170–1179, https://doi.org/10.1002/jcc.540160911.
[49] S. Dapprich, I. Komáromi, K.S. Byun, K. Morokuma, M.J. Frisch, A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives, J. Mol. Struct. Theochem. 461–462 (1999) 1–21, https://doi.org/10.1016/S0166-1280(98)00475-8.
[50] N. Sharma, Ways to Detect and Remove the Outliers, (2018). https://towardsdatascience.com/ways-to-detect-and-remove-the-outliers-404d16608dba.
[51] A.L. Hickey, C.N. Rowley, Benchmarking quantum chemical methods for the calculation of molecular dipole moments and polarizabilities, J. Phys. Chem. A. 118 (20) (2014) 3678–3687, https://doi.org/10.1021/jp502475e.
[52] D. Hait, M. Head-Gordon, How accurate is density functional theory at predicting dipole moments? an assessment using a new database of 200 benchmark values, J. Chem. Theory Comput. 14 (4) (2018) 1969–1981, https://doi.org/10.1021/acs.jctc.7b01252.
[53] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, The missing term in effective pair potentials, J. Phys. Chem. 91 (24) (1987) 6269–6271, https://doi.org/10.1021/j100308a038.
[54] D.G. Truhlar, Basis-set extrapolation, Chem. Phys. Lett. 294 (1-3) (1998) 45–48, https://doi.org/10.1016/S0009-2614(98)00866-5.
[55] J. Toda, M. Fischer, M. Jorge, J.R.B. Gomes, Water adsorption on a copper formate paddlewheel model of CuBTC: A comparative MP2 and DFT study, Chem. Phys. Lett. 587 (2013) 7–13, https://doi.org/10.1016/j.cplett.2013.09.049.
[56] R. Afonso, J. Toda, J.R.B. Gomes, M. Fischer, C. Campbell, M. Jorge, A computational study of the interaction of C2 hydrocarbons with CuBTC, Comput. Mater. Sci. 173 (2020) 109438, https://doi.org/10.1016/j.commatsci.2019.109438.
[57] E. Cancès, B. Mennucci, J. Tomasi, A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics, J. Chem. Phys. 107 (8) (1997) 3032–3041, https://doi.org/10.1063/1.474659.
[58] S. Chalmet, D. Rinaldi, M.F. Ruiz-López, A QM/MM/continuum model for computations in solution: comparison with QM/MM molecular dynamics simulations, Int. J. Quantum Chem. 84 (5) (2001) 559–564, https://doi.org/10.1002/qua.1410.
[59] T.D. Poulsen, P.R. Ogilby, K.V. Mikkelsen, Linear response properties for solvated molecules described by a combined multiconfigurational self-consistent-field/molecular mechanics model, J. Chem. Phys. 116 (9) (2002) 3730–3738, https://doi.org/10.1063/1.1436478.
[60] Anders Osted, Jacob Kongsted, Kurt Mikkelsen, Ove Christiansen, A CC2 dielectric continuum model and a CC2 molecular mechanics model, Mol. Phys. 101 (13) (2003) 2055–2071, https://doi.org/10.1080/0026897031000109338.
[61] L. Jensen, P.T. van Duijnen, J.G. Snijders, A discrete solvent reaction field model within density functional theory, J. Chem. Phys. 118 (2) (2003) 514–521, https://doi.org/10.1063/1.1527010.
[62] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: A message-passing parallel molecular dynamics implementation, Comput. Phys. Commun. 91 (1-3) (1995) 43–56, https://doi.org/10.1016/0010-4655(95)00042-E.
[63] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1–2 (2015) 19–25, https://doi.org/10.1016/j.softx.2015.06.001.
[64] R.W. Hockney, S.P. Goel, J.W. Eastwood, Quiet high-resolution computer models of a plasma, J. Comput. Phys. 14 (2) (1974) 148–158, https://doi.org/10.1016/0021-9991(74)90010-2.
[65] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, J. Chem. Phys. 81 (1) (1984) 511–519, https://doi.org/10.1063/1.447334.
[66] W.G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, Phys. Rev. A. 31 (3) (1985) 1695–1697, https://doi.org/10.1103/PhysRevA.31.1695.
[67] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys. 52 (12) (1981) 7182–7190, https://doi.org/10.1063/1.328693.
[68] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N ·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (12) (1993) 10089–10092, https://doi.org/10.1063/1.464397.
[69] K. Coutinho, R.C. Guedes, B.J. Costa Cabral, S. Canuto, Electronic polarization of liquid water: converged Monte Carlo-quantum mechanics results for the multipole moments, Chem. Phys. Lett. 369 (3-4) (2003) 345–353, https://doi.org/10.1016/S0009-2614(02)02026-2.

[70] A. Jindal, S. Vasudevan, Geometry of OH···O interactions in the liquid state of linear alcohols from ab initio molecular dynamics simulations, Phys. Chem. Chem. Phys. 22 (12) (2020) 6690–6697, https://doi.org/10.1039/D0CP00435A.
[71] J. Lehtola, M. Hakala, K. Hämäläinen, Structure of Liquid Linear Alcohols, J. Phys. Chem. B. 114 (19) (2010) 6426–6436, https://doi.org/10.1021/jp909894y.
[72] M. Tomšič, A. Jamnik, G. Fritz-Popovski, O. Glatter, L. Vlček, Structural properties of pure simple alcohols from ethanol, propanol, butanol, pentanol, to hexanol: comparing monte carlo simulations with experimental SAXS data, J. Phys. Chem. B. 111 (2007) 1738–1751, https://doi.org/10.1021/jp066139z.
[73] B.D.T. Bowron, J.L. Finney, A.K. Soper, The structure of pure tertiary butanol, Mol. Phys. 93 (4) (1998) 531–543, https://doi.org/10.1080/002689798168871.
[74] M.A. González, F.J. Bermejo, E. Enciso, C. Cabrillo, Hydrogen bonding in condensed-phase alcohols: some keys to understanding their structure and dynamics, Philos. Mag. 84 (13-16) (2004) 1599–1607, https://doi.org/10.1080/14786430310001644387.
[75] I. Pethes, L. Pusztai, K. Ohara, L. Temleitner, Temperature-dependent structure of 1-propanol/water mixtures: X-ray diffraction experiments and computer simulations at low and high alcohol contents, J. Mol. Liq. 340 (2021) 117188, https://doi.org/10.1016/j.molliq.2021.117188.
[76] S. Pothoczki, I. Pethes, L. Pusztai, L. Temleitner, D. Csókás, S. Kohara, K. Ohara, I. Bakó, Hydrogen bonding and percolation in propan-2-ol – Water liquid mixtures: X-ray diffraction experiments and computer simulations, J. Mol. Liq. 329 (2021) 115592, https://doi.org/10.1016/j.molliq.2021.115592.
[77] M. Jorge, Predicting hydrophobic solvation by molecular simulation: 2. New united-atom model for alkanes, alkenes, and alkynes, J. Comput. Chem. 38 (6) (2017) 359–369, https://doi.org/10.1002/jcc.24689.
[78] M. Jorge, N.M. Garrido, C.J.V. Simões, C.G. Silva, R.M.M. Brito, Predicting hydrophobic solvation by molecular simulation: 1. Testing united-atom alkane models, J. Comput. Chem. 38 (6) (2017) 346–358, https://doi.org/10.1002/jcc.24690.