

# A Polarization-Consistent Model for Alcohols to Predict Solvation Free Energies

Maria Cecilia Barrera and Miguel Jorge\*

Cite This: *J. Chem. Inf. Model.* 2020, 60, 1352–1367

Read Online

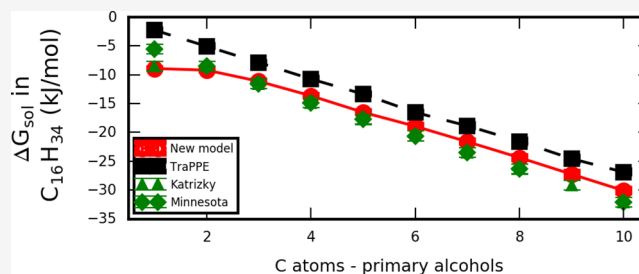
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** Classical nonpolarizable models, normally based on a combination of Lennard-Jones sites and point charges, are extensively used to model thermodynamic properties of fluids, including solvation. An important shortcoming of these models is that they do not explicitly account for polarization effects, i.e., a description of how the electron density responds to changes in the molecular environment. Instead, polarization is implicitly included, in a mean-field sense, into the parameters of the model, usually by fitting to pure liquid properties (e.g., density). This causes problems when trying to describe thermodynamic properties that involve a change of phase (e.g., enthalpy of vaporization), that directly depend on the electronic response of the medium (e.g., dielectric constant), and that require mixing or solvation in different media (e.g., solvation free energies). Fully polarizable models present a natural route for addressing these limitations but at the price of a much higher computational cost. In this work, we combine the best of those two approaches by running fast simulations using nonpolarizable models and applying *post facto* corrections to the computed properties in order to account for the effects of polarization. By applying this new paradigm, a new united-atom force field for alcohols is developed that is able to predict both pure liquid properties, including dielectric constant, and solvation free energies in different solvents with a high degree of accuracy. This paves the way for the development of a generic classical nonpolarizable force field that can predict solvation of drug-like molecules in a variety of solvents.



## 1. INTRODUCTION

Solubility is a very important property in the manufacture and formulation of drugs.<sup>1</sup> For example, drug absorption from the gastrointestinal tract depends strongly on the solubility of orally administered drugs. Nowadays, most drug candidates have a low solubility in water, and as a consequence, their absorption has to be improved by using a suitable formulation approach. About 40% of drugs in the market and nearly 90% of drugs in development are poorly water soluble.<sup>2</sup> Cosolvents can be used to enhance the solubility of nonpolar solutes by several orders of magnitude. These compounds commonly contain both hydrogen-bonding groups, which interact strongly with water, and nonhydrogen-bonding groups. Dimethylacetamide, ethanol, and propylene glycol are widely used in the pharmaceutical industry as cosolvents.<sup>3</sup> Consequently, knowing the solubility of the drug in water and in other solvents, during its formulation process, is extremely important.<sup>4,5</sup> In addition, the octanol/water partition coefficient is a very important property in the pharmaceutical industry, and for example, it can be used to describe a drug's ability to diffuse through lipids.<sup>6</sup> Apart from their practical importance as solvents, alcohols are also interesting from a fundamental point of view,<sup>7–9</sup> as they are the simplest molecules that combine a hydrophobic moiety with a hydrogen-bonding functional group.

Experimental determination of the solubility of drug molecules is very laborious and time consuming, especially when the solubility is very low. As a result, much time and effort could be saved by accurately predicting it using computational methods.<sup>10,11</sup> In particular, molecular dynamics (MD) simulations using classical nonpolarizable models have shown great promise for predicting drug solubilities.<sup>12,13</sup> Despite this promise, however, MD has not yet surpassed the accuracy of much less computationally expensive approaches like quantitative structure–property relationships or group contribution methods in blind tests of solvation free energy predictions.<sup>14</sup> This is most likely due to the inherent shortcomings and lack of transferability of the underlying nonpolarizable molecular models, which cannot explicitly account for polarization effects since the partial charges are assumed to be fixed. Although progress has been made using higher levels of theory, such as fully polarizable models<sup>15</sup> or QM/MM calculations,<sup>16</sup> these approaches are too computationally expensive for routine use. Therefore, it is important to assess to which extent the lack of polarization effects in

Received: October 30, 2019

Published: January 24, 2020

classical simulations can be taken into account by inexpensive correction schemes.

In 1987, Berendsen et al. proposed a correction that accounts for the repolarization, or distortion, of a molecule when it moves from the liquid phase to the gas phase and used this to develop a new water model, SPC/E.<sup>17</sup> Since then, most “modern” water models<sup>18–21</sup> and a few methanol models<sup>22,23</sup> have been parametrized using this correction. However, the issue has been mostly neglected in the development of generic force fields like OPLS-AA,<sup>24</sup> GROMOS,<sup>25</sup> TraPPE-UA,<sup>26</sup> and GAFF.<sup>27</sup> Moreover, it has been argued on the basis of theoretical considerations that not only is the Berendsen distortion correction underestimated but it should be supplemented by another term to represent the interaction of the polarized molecule with the purely electronic degrees of freedom of the solvent.<sup>28</sup> When these two energy terms, which have opposite signs, were estimated from high-level quantum mechanical calculations for liquid water, it was shown that they nearly cancel each other out.<sup>29</sup> It is not yet clear to which extent this effect is observed for other molecules of lower polarity. Furthermore, when polarization effects are represented by a simple dipole moment scaling factor,<sup>30</sup> dielectric constants that are in quantitative agreement with experiments are obtained for a wide variety of molecules, eliminating previously observed systematic deviations.<sup>31</sup> These recent advances suggest that marked improvements in the performance of classical nonpolarizable models can be obtained if accurate polarization corrections are employed in a consistent fashion in force field parametrization and validation stages. In this paper, we demonstrate that this can indeed be achieved for the particular case of alcohol molecules.

Apart from the generic force fields mentioned above, there have been several attempts to parametrize specific models for alcohols. For example, Weerasinghe and Smith<sup>22</sup> developed a methanol model from Kirkwood–Buff integrals aimed at describing pure liquid and aqueous solution properties. The Berendsen correction was employed in calculations of the enthalpy of vaporization, but no corrections were applied to the dielectric constant or to thermodynamic properties of mixing. The more recent OPLS/2016 model<sup>23</sup> was reparametrized including solid–fluid experimental data. Several properties were used for the validation of the model including the static dielectric constant, which was calculated using polarization scaling corrections.<sup>30,31</sup> Interestingly, while the calculation of the enthalpy of vaporization applied the model-dependent Berendsen correction, this was not used when computing other vapor–liquid coexistence properties, such as phase diagrams and vapor pressure. Another modification of OPLS was the one proposed by Kulschewski and Pleiss.<sup>32</sup> They modified the partial charges of the hydroxyl group in the OPLS-AA force field for different alcohols (including methanol to octanol) to consider the influence of the alkyl tail on the electronic structure of the hydroxyl group. Their simulated densities, self-diffusion coefficients, and dielectric constants are closer to the experimental values than the predictions from the original OPLS-AA force field. However, these modifications increase the complexity of the model and reduce its transferability.<sup>32</sup> Moreover, no attempt was made to include polarization corrections in the calculations. An improved GROMOS force field for alcohols was obtained by fitting to the density, enthalpy of vaporization, and free energy of solvation in water and in cyclohexane,<sup>33</sup> but again, no polarization corrections were used. In conclusion,

although there are several force fields for alcohols available in the literature, none of them, to the best of our knowledge, has been parametrized using consistent polarization corrections. Here, we develop a new united-atom model for alcohols taking into consideration polarization corrections at the parametrization and validation stage for all relevant properties. United-atom (UA) models have fewer interaction sites than all-atom models, and therefore, they are less computationally expensive and simpler to parametrize. Furthermore, our starting point is a modified TraPPE-UA force field for hydrocarbons<sup>34</sup> (i.e., alkanes, alkenes, and alkynes) that corrects small but systematic deviations when predicting solvation free energies in hydrophobic solvents using generic force fields.<sup>35</sup> In particular, we show that very accurate predictions of solvation free energies in solvents of different polarity can be obtained when polarization effects are consistently accounted for.

## 2. METHODOLOGY

**2.1. Bulk Properties.** Simulations were run with Gromacs 5.1.2.<sup>36</sup> First, 399 molecules of methanol were inserted in a cubic box of approximately 3 nm each side, and a steepest descent minimization run was performed. After this, a 100 ps NVT simulation and a 100 ps NPT simulation, using the Berendsen coupling scheme,<sup>37</sup> were performed to reach the desired temperature (298.15 K) and pressure (1 bar). Subsequently, a 25 ns NPT simulation was run where the temperature was kept at 298.15 K using the V-rescale thermostat,<sup>38</sup> with a time constant of 0.1 ps, and the pressure was controlled using the Parrinello–Rahman barostat,<sup>39</sup> with a time constant of 2 ps and an isothermal compressibility of  $4.5 \times 10^{-5} \text{ bar}^{-1}$ . A leapfrog algorithm<sup>40</sup> was used to integrate Newton’s equations of motion with a time step of 2 fs. All bonds were constrained using the LINCS algorithm.<sup>41</sup> The Verlet scheme was chosen for neighbor searching, with a cutoff radius of 1 nm for both van der Waals and electrostatic interactions, and long-range dispersion corrections for energy and pressure were applied. Long-range electrostatic interactions were calculated using PME<sup>42</sup> with a Fourier spacing of 0.16 nm. Additionally, periodic boundary conditions were always applied unless stated otherwise.

The same procedure was followed when simulating ethanol (473 molecules), propanol (361 molecules), butanol (281 molecules), pentanol (216 molecules), hexanol (180 molecules), heptanol (147 molecules), octanol (128 molecules), nonanol (111 molecules), decanol (103 molecules), 2-propanol (367 molecules), 2-butanol (290 molecules), 2-pentanol (233 molecules) 2-methyl-2-propanol (289 molecules), and 2-methyl-2-butanol (228 molecules). The number of molecules in each simulation box was selected to maintain an approximately constant box size (ca. 3 nm), in order to minimize the use of computational resources. We have previously shown that the results (with the exception of the self-diffusion coefficient; see below) are independent of system size provided long-range corrections are employed.<sup>35</sup> In total, 10 independent simulations were run for each compound, and error bars were calculated to give a 95% confidence interval using a Student’s *t* equal to 2.262.<sup>43</sup> In some cases, the error bars were too small to be plotted, and therefore, they are not visible in the graphs.

The enthalpies of vaporization were calculated using eq 1

$$\Delta H_{\text{vap}} = U_{\text{gas}} - U_{\text{liq}} + RT + C_{\text{pol}} \quad (1)$$

where  $U_{\text{gas}}$  is the molar potential energy in the vapor phase,  $U_{\text{liq}}$  is the molar potential energy in the liquid phase,  $R$  is the ideal gas constant, and  $T$  is the temperature. This equation assumes that the kinetic energy per molecule is the same in the gas phase and in the liquid phase, an assumption which has been shown to be very accurate for small molecules.<sup>44</sup> To estimate the potential energy in the vapor phase (i.e., the intramolecular energy of an isolated molecule), a 50 ns simulation of a single molecule in vacuum with no boundary conditions was run for each alcohol, and the first 10 ns were discarded. For these simulations, the cutoff scheme selected for neighbor searching was “group”, and the cutoff radii were all set to 0.

A correction term ( $C_{\text{pol}}$ ) was added to the enthalpy calculation to account for polarization effects. This term was calculated using eq 2<sup>45</sup>

$$C_{\text{pol}} = -\frac{(\mu_{\text{l}} - \mu_{\text{g}})^2}{2\alpha} + \frac{6(\epsilon_{\text{el}} - 1)^2 \mu_{\text{l}}^2}{\pi(2\epsilon_{\text{el}} + 1)(\epsilon_{\text{el}} + 2)\alpha} \quad (2)$$

Here,  $\mu_{\text{l}}$  and  $\mu_{\text{g}}$  are the liquid and gas phase dipole moments, respectively. Additionally,  $\alpha$  and  $\epsilon_{\text{el}}$  are the electronic polarizability of the molecule in the gas phase and the high-frequency dielectric permittivity of the medium, respectively. The first term on the right-hand side of eq 2 is the distortion energy, which is favorable when moving from the liquid phase to the gas phase (as in a vaporization process). The second term is the interaction energy of the molecule with the surrounding electronic continuum, described here by a simple Kirkwood–Onsager model<sup>46–48</sup> for a dipole in a spherical cavity. The radius of the cavity is treated self-consistently so that it is eliminated from the equations for the polarization energy.<sup>45</sup> Equation 2 contains only the dipole moment contributions to the polarization energies. Although it has been shown that quadrupole (and potentially higher moment) contributions can be non-negligible,<sup>29,49</sup> it was not possible to calculate those contributions here due to the absence of sufficiently accurate data for the quadrupole moments and quadrupole polarizabilities of alcohols in the liquid phase. Note also that quantum vibration corrections were not included in the calculation of the enthalpy of vaporization, as they are already negligible for methanol.<sup>22</sup> Finally, it is important to clarify that the correction  $C_{\text{pol}}$  employed here is independent of the parameters of the model, unlike the correction proposed by Berendsen et al. which depends on the effective dipole of the model ( $\mu_{\text{model}}$ ), as can be seen from eq 3.<sup>17</sup>

$$C_{\text{Berendsen}} = -\frac{(\mu_{\text{model}} - \mu_{\text{g}})^2}{2\alpha} \quad (3)$$

The theory of Leontyev and Stuchebrukhov<sup>45</sup> is based on decoupling the contributions to the dielectric response of the solvent coming from the “fast” electronic degrees of freedom and those arising from the “slow” nuclear degrees of freedom, following the Born–Oppenheimer approximation. As such, the liquid state can be thought of as nuclear charges  $q_i$  moving in a polarizable electronic continuum with a high-frequency dielectric permittivity ( $\epsilon_{\text{el}}$ ) equal to  $n^2$  (where  $n$  is the refraction index of the medium), while the relative permittivity of the gas phase is close to 1 ( $\epsilon_{\text{el}} \approx 1$  for a perfect vacuum). This purely electronic continuum interacts with the polarized molecule and also screens its charges, resulting in effective charges that are much lower than those of the real liquid, i.e.,

$q_i^{\text{eff}} = q_i/\epsilon_{\text{el}}^{0.5}$ . Hence, in order to consistently account for polarization effects, the value of  $\mu_{\text{l}}$  to be used in eq 2 should be the dipole of the real liquid, which is normally much higher than the dipole moment of nonpolarizable fixed-charge models due to the charge screening effect.<sup>29,45</sup> As mentioned earlier, the original Berendsen correction uses the dipole moment of the model as a proxy for the real liquid dipole, hence leading to strongly underestimated distortion corrections.<sup>45</sup> In previous work,<sup>29</sup> we used the real dipole moment of liquid water estimated from experiments<sup>50</sup> and high-level *ab initio* MD calculations.<sup>51</sup> Although, to the best of our knowledge, there is no reliable experimental estimate of the liquid phase dipole moment for alcohols, we were able to find a few AIMD studies that report the liquid dipole of methanol. Pagliai et al.<sup>52</sup> reported a value of 2.64 D from simulations on a rather small box of 26 methanol molecules using the BLYP exchange-correlation functional. The same functional was used by Handgraaf et al.<sup>53</sup> with a larger box of 64 molecules, yielding a value of 2.59 D. In a study of the methanol vapor/liquid interface with a box of 120 molecules, Kuo et al.<sup>54</sup> report average bulk liquid dipole moments around 2.7 D, using both BLYP and PBE functionals. Finally, a more recent study by Sieffert et al.<sup>55</sup> report a range of values obtained with several functionals in a box of 64 molecules, ranging from 2.58 to 2.84 D. Importantly, some of the calculations were carried out with dispersion corrections, which are known to play important roles in determining the structure of liquids.<sup>51</sup> The average value of those calculations is 2.7 D, which we take to be the best estimate of the dipole moment of liquid methanol. Unfortunately, we were unable to find similar studies for larger alcohols; therefore, an alternative approach was sought. Namely, we used an equation proposed by Leontyev and Stuchebrukhov,<sup>45</sup> again based on the theory of continuum dielectrics in a spherical cavity, where  $\epsilon_{\text{sol}}$  is the static dielectric constant of the solvent.

$$\mu_{\text{l}} = \mu_{\text{g}} \left( 1 - \frac{12(\epsilon_{\text{el}} - 1)(\epsilon_{\text{sol}} - 1)}{\pi(\epsilon_{\text{el}} + 2)(2\epsilon_{\text{sol}} + 1)} \right)^{-1} \quad (4)$$

Using experimental values for  $\mu_{\text{g}}$  and  $\epsilon_{\text{sol}}$  and estimating  $\epsilon_{\text{el}}$  from the square of the experimental refraction index at the sodium D-line frequency, we obtain a value for methanol of 2.7 D, in precise agreement with the average of AIMD simulations mentioned above. This expression also provides a good estimate of the dipole moment of liquid water.<sup>29</sup> This gives confidence that eq 4 can provide reasonable estimates of the liquid phase dipole moments for the remaining alcohol molecules. Table S1 reports details of these calculations for all the molecules considered in this work, together with the corresponding polarization corrections obtained from eq 2.

The diffusion constants were calculated using the Einstein relation (eq 5). The left-hand side of eq 5 was obtained by linear regression of the mean square displacement (MSD), where the times were weighted according to the number of reference points. In every case, the fitting was done between  $t = 100$  ps and  $t = 500$  ps, where  $t$  is the time from the reference positions and not simulation time. The estimated error was the difference of the diffusion coefficients obtained from fits over the two halves of the fit interval.<sup>56</sup>

$$\lim_{t \rightarrow \infty} \langle \left\| \mathbf{r}_i(t) - \mathbf{r}_i(0) \right\|^2 \rangle_{i \in \Lambda} = 6D_{\Lambda} t \quad (5)$$

Here,  $D_A$  is the self-diffusion constant of particles of type A, and the left-hand side of eq 5 is the limit of the mean square displacement when time tends to infinity.

It has been shown that the apparent self-diffusion coefficients depend significantly on the system size, and thus, it is important to correct for the deviations observed when comparing to an infinite system.<sup>57</sup> The diffusion coefficient depends linearly on the inverse box length ( $1/L$ ), and consequently, the diffusion constant for an infinite system can be calculated by extrapolating a straight-line fit<sup>57</sup> (Figure S1). For this reason, simulations were run for four different box lengths (3, 4, 5, 6 nm), and a correction term was calculated for each alcohol molecule by subtracting the simulated diffusion constant obtained using a cubic box of 3 nm each side from the intercept of the fitted line. It was assumed that these corrections would be similar for other parameter sets, and therefore, they were used during force field optimization and validation stages. To corroborate this hypothesis, new correction terms were calculated using the final optimized force field, and the values obtained were consistent with the ones calculated with TraPPE (see Table S2 for a comparison).

The static dielectric constant was estimated with Gromacs through eq 6<sup>58</sup>

$$\epsilon_{sim} = 1 + \frac{\langle M^2 \rangle - \langle M \rangle^2}{3\epsilon_0 k_B T \langle V \rangle} \quad (6)$$

where  $k_B$  is the Boltzmann constant,  $V$  is the volume of the simulation box and  $M$  is the total dipole moment of the system. The dielectric constant obtained using Gromacs was then corrected for polarization effects using eq 7.<sup>31</sup>

$$\epsilon_{scaled} = \epsilon_{el} + (\epsilon_{sim} - 1)k^2 \quad (7)$$

where  $\epsilon_{sim}$  is the dielectric constant obtained from eq 6, and  $k$  is the ratio between the dipole moment of the real liquid (estimated from eq 4) and the dipole moment of the nonpolarizable model. This equation, derived elsewhere,<sup>31</sup> takes into account the purely electronic response of the liquid, which is not accounted for in nonpolarizable models, as well as the charge screening caused by the presence of this electronic continuum.

**2.2. Free Energy of Solvation.** The free energy of solvation ( $\Delta G_{sol}$ ) was calculated by alchemical transformations using standard protocols,<sup>59</sup> with full details provided in the Supporting Information. Here, we provide only a summary of the most important methodological considerations.

First of all, we computed  $\Delta G_{sol}$  via a one-step transformation using the option `couple-intramol = "no"` in Gromacs. This means that the decoupled state of the solute molecule corresponds to the proper vacuum state without periodicity effects and not to a molecule without intramolecular interactions.<sup>56</sup> This implies the assumption that the contribution of the intramolecular degrees of freedom of the solute to the solvation free energy are the same in the gas phase and in the liquid phase. We have previously shown that this assumption is accurate for solvation of alkanes up to hexadecane.<sup>35</sup> To further check the accuracy of this assumption for alcohols, the free energy of solvation of decanol obtained from two independent one-step simulations was compared to the results obtained using a full thermodynamic cycle (i.e., a standard two-step simulation), and no significant differences were observed (see Table S3 in the Supporting Information).

The free energies were sampled using Bennett's acceptance ratio (BAR) method.<sup>60,61</sup> It is more efficient and insightful to calculate electrostatic and Lennard-Jones (LJ) free energy contributions separately,<sup>59,62</sup> and this was the procedure followed here. The LJ component of the free energy of solvation was calculated using 15 lambda values (0, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.7, 0.8, 0.9, and 1), while seven lambda values were used to obtain the electrostatic component (0, 0.3, 0.6, 0.7, 0.8, 0.9, and 1). The intermediate lambdas were chosen based on their relative entropy ( $s$ ), which is a measure of how one probability distribution diverges from a second.<sup>63</sup> Table S4 shows the relative entropies for the LJ and electrostatic component of methanol as an example. The error in the free energy calculations was estimated using block averaging and is reported as error bars representing approximately a 95% confidence interval on the mean.

Free energy simulations were run using a leapfrog stochastic dynamics integrator,<sup>64</sup> and therefore, the temperature was kept constant using the Langevin method. A soft-core function was used to avoid instabilities close to the noninteracting state.<sup>65</sup> The soft-core parameters *sc-power* and *sc-sigma* were 1 and 0.3, respectively. *Sc-alpha* was 0.5 for the LJ term since this seems to be the ideal value for transformations of chargeless molecules,<sup>62</sup> while *sc-alpha* = 0 was used for the electrostatic component. All other parameters were the same as the ones used to calculate bulk properties except for the Fourier spacing, which was 0.12, and the time constant for pressure coupling, which in this case was 5 ps. Long-range van der Waals interactions were calculated using a cutoff method with long-range corrections. In some situations, it is necessary to use PME when calculating van der Waals interactions;<sup>66</sup> however, as can be seen from Figure S5 in the Supporting Information, both methods are equivalent for the alcohols studied here.

The length of the simulation differed from molecule to molecule, and it was chosen by plotting free energy of self-solvation vs simulation time to ensure convergence (see Figures S3 and S4 in the Supporting Information). When calculating the LJ term of the free energy of self-solvation, simulations were run for 5 ns (methanol to propanol), 7 ns (butanol to heptanol), and 10 ns (octanol to decanol), while longer simulation times were needed to calculate the component of the free energy of self-solvation that is due to electrostatic interactions. The simulation times in this case were 5 ns (methanol), 15 ns (ethanol and propanol), 35 ns (butanol and pentanol), and 50 ns (hexanol to decanol). Additionally, 10 ns simulations were run to obtain the free energy of solvation of primary alcohols in hexadecane and the free energy of solvation of alkanes in octanol. In all cases, the first 500 ps were discarded.

Polarization corrections were added to the solvation free energies calculated from MD simulations. These corrections were calculated from eqs 2 and 4, but  $C_{pol}$  has the opposite sign, since the solvation process moves a molecule from the vapor to the condensed phase (i.e., the direction is the opposite of the vaporization process). Note that when applying those equations to solvation, the experimental values of  $\mu_g$  and  $\alpha$  correspond to the solute molecule, while values of  $\epsilon_{sol}$  and  $\epsilon_{el}$  correspond to the surrounding solvent; hence,  $\mu_l$  is the estimated dipole moment of the solute in the solvent of interest. Results for the polarization corrections in all systems studied here are reported in Section 3.1. Experimental values for the free energies of solvation were taken from the Minnesota Solvation Database<sup>67</sup> and the Katritzky database.<sup>68</sup>

Additionally, some free energies of self-solvation were calculated from the experimental vapor pressure and density at 298 K, using eq 8<sup>67</sup>

$$\Delta G_s = -2.303RT \log \left( \frac{\rho/M_w}{P_v/24.45} \right) \quad (8)$$

where 24.45 atm is the pressure of an ideal gas at 1 molar concentration and 298 K,  $P_v$  is the vapor pressure in atm,  $\rho$  is the density in g/L, and  $M_w$  is the molecular weight.

**2.3. Force Field Parameters.** The new model developed here is based on the TraPPE-UA (Transferable Potentials for Phase Equilibria-United Atom) force field<sup>69</sup> as this force field performed better than OPLS-UA and GROMOS at predicting hydrophobic solvation.<sup>35</sup> Only nonbonded parameters were modified, while all bonded parameters were kept unchanged. This force field treats the  $\text{CH}_x$  groups as pseudoatoms, located at the sites of the carbon atoms, while it models all other atoms explicitly. Nonbonded interactions are described by pairwise-additive LJ 12-6 potentials and Coulombic interactions of partial charges

$$u(r_{ij}) = 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (9)$$

where  $\epsilon_{ij}$ ,  $\sigma_{ij}$ ,  $q_i$ , and  $q_j$  are the LJ well depth, LJ size, and partial charges, respectively, for the pair of atoms  $i$  and  $j$  separated by a distance  $r_{ij}$ , and  $\epsilon_0$  is the vacuum permittivity. The LJ parameters for the interaction between two different atoms are calculated using the Lorentz–Berthelot combining rules (eqs 10 and 11)

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}) \quad (10)$$

$$\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \quad (11)$$

To describe intramolecular interactions, the TraPPE force field uses fixed bonds lengths, a harmonic potential for bond angle bending (eq 12), and torsional potentials to restrict the dihedral rotations around bonds (eq 13)

$$u_{\text{bend}} = \frac{k_\theta}{2}(\theta - \theta_0)^2 \quad (12)$$

where  $\theta$ ,  $\theta_0$ , and  $k_\theta$  are the instantaneous bending angle, the equilibrium bending angle, and the force constant, respectively.

$$u_{\text{tors}} = c_0 + c_1[1 + \cos(\phi)] + c_2[1 - \cos(2\phi)] + c_3[1 + \cos(3\phi)] \quad (13)$$

Here,  $\phi$  is the dihedral angle, and  $c_0$ ,  $c_1$ ,  $c_2$ , and  $c_3$  are the Fourier coefficients. Intramolecular 1–4 Lennard-Jones and Coulomb interactions are excluded.

The TraPPE-UA LJ parameters were determined by fitting the single component vapor–liquid phase diagram of methanol, ethanol, 1-propanol, 2-propanol, 2-butanol, 2-methylpropan-2-ol, 1-pentanol, pentane-1,5-diol, and 1-octanol, while point charges were taken from the OPLS-UA force field<sup>24,70</sup> without modifications.

The LJ parameters for alkanes, alkenes, and alkynes have previously been optimized to correct for the small but systematic deviations observed when predicting their solvation free energies in hydrophobic solvents using the original TraPPE-UA force field.<sup>34</sup> These parameters are listed in

Table 1, along with the new parameters found for alcohols in this work (see Section 2.4). Bonded parameters, which are

**Table 1. Lennard-Jones Parameters and Partial Charges for the New PolCA United-Atom Force Field for Alcohols**

nonbonded <sup>a</sup> (eq 9)	$\sigma$ (nm)	$\epsilon$ (kJ/mol)	partial charge ( $q$ )
<b><math>\text{CH}_x</math>–(O)–H</b>	0.2853	0.7733	–0.646
<b>O–(H)</b>	0	0	+0.406
<b>(<math>\text{CH}_3</math>)–OH</b>	0.379	0.833	+0.240
<b>(<math>\text{CH}_3</math>)–<math>\text{CH}_x</math></b>	0.379	0.833	0
<b><math>\text{CH}_x</math>–(<math>\text{CH}_2</math>)–OH</b>	0.399	0.392	+0.240
<b>(<math>\text{CH}_x</math>)<sub>2</sub>–(<math>\text{CH}_2</math>)</b>	0.399	0.392	0
<b>(<math>\text{CH}_x</math>)<sub>2</sub>–(<math>\text{CH}</math>)–OH</b>	0.438	0.085	+0.240
<b>(<math>\text{CH}_x</math>)<sub>3</sub>–(<math>\text{CH}</math>)</b>	0.473	0.085	0
<b>(<math>\text{CH}_x</math>)<sub>3</sub>–(<math>\text{C}</math>)–OH</b>	0.585	0.00426	+0.240
<b>(<math>\text{CH}_x</math>)<sub>4</sub>–(<math>\text{C}</math>)</b>	0.646	0.00426	0

<sup>a</sup>Nonbonded parameters correspond to the sites in bold.

identical to those of TraPPE-UA, are reported in Table S5 for completeness, while the original TraPPE-UA nonbonded parameters are reported in Table S7.

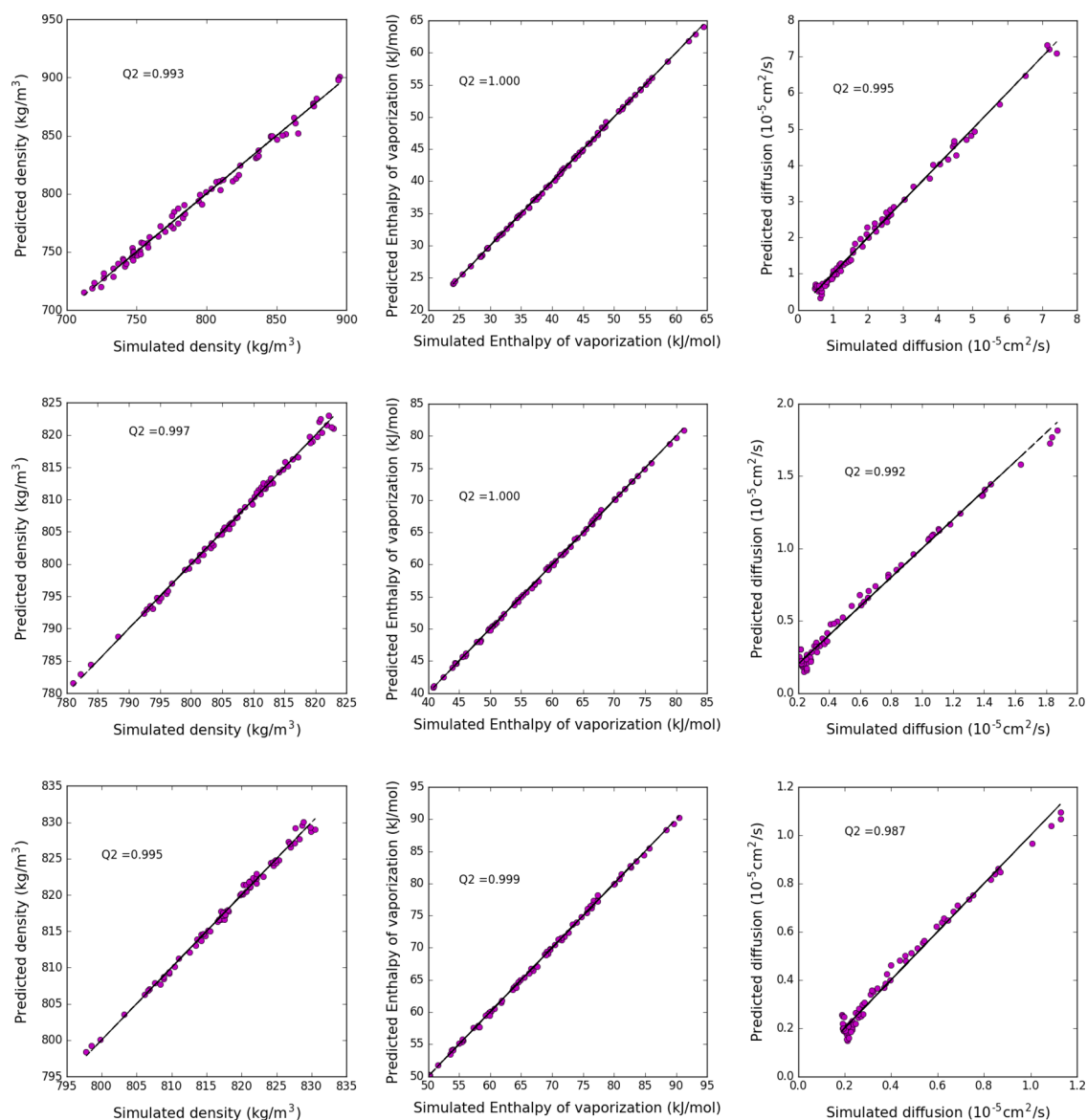
**2.4. Force Field Optimization.** In this paper, we developed a new force field for alcohols by incorporating polarization corrections on all relevant properties (see Section 2.1) at the parametrization stage. To our knowledge, it is the first time that this approach has been applied in force field development. We have named our new model PolCA, standing for “Polarization-Consistent Approach”. Below, we describe in detail the strategy used to optimize the potential parameters for PolCA.

The density, diffusion constant, and enthalpy of vaporization of methanol, 1-pentanol, and 1-heptanol were simulated using different LJ parameters for the oxygen atom and different partial charges for the hydroxyl group. The partial charge of the  $\alpha$ -carbon was obtained as minus the sum of the other two partial charges to ensure neutrality of the molecule. The parameters for the alkyl chain groups were taken from the model we have proposed for alkanes, alkenes, and alkynes<sup>34</sup> (Table 1). Additionally, the LJ parameters of the alkane methyl and methylene groups were used for the  $\alpha$ -carbons without modification.

Meta-models, that predict molecular simulation results for a given set of parameters, were used to decrease computational cost and allow for a more extensive exploration of the force field parameter space. This technique has already been tested by Cailliez et al.<sup>71</sup> and proven to be successful. They also found that it was much more convenient and efficient to model each individual property used for the calibration process instead of directly modeling the full objective function.<sup>71</sup> For this reason, a learning set was built for each molecule used in the optimization routine, and then, each property was fitted to a second-order model using eq 14<sup>72</sup>

$$f(X) = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i < j} \beta_{ij} x_i x_j + \sum_{i=1}^k \beta_{ii} x_i^2 \quad (14)$$

where  $k$  is the number of parameters and  $X = (x_1, x_2, \dots, x_k)$ . Coded values of the force field parameters were used instead of using the unmodified parameters as this resulted in better prediction of the properties. Equation 15 was used to convert  $\sigma$ ,  $\epsilon$ , oxygen’s partial charge ( $q_O$ ), and hydrogen’s partial charge ( $q_H$ ) into their corresponding coded values.



**Figure 1.** Plot of the simulated values vs the values predicted using the metamodels. The first column is the density, the second column the enthalpy of vaporization, and the last column the diffusion. The rows are methanol, pentanol, and heptanol, from top to bottom.

$$x_{\text{coded}} = \frac{2(x - x_c)}{x_{\text{max}} - x_{\text{min}}} \quad (15)$$

where  $x_{\text{min}}$  and  $x_{\text{max}}$  are the lowest and highest values for each control variable ( $\sigma$ ,  $\varepsilon$ ,  $q_{\text{O}}$ , and  $q_{\text{H}}$ ), respectively, and  $x_c$  is the corresponding parameter taken from the TraPPE force field ( $\sigma = 0.302$ ,  $\varepsilon = 0.773$ ,  $q_{\text{O}} = -0.7$ , and  $q_{\text{H}} = 0.435$ ).<sup>69</sup> The learning set was created by assigning different levels to each factor and then simulating all possible combinations (except for combinations of  $\sigma = 0.294$  or  $0.298$  with  $\varepsilon = 0.7$  or  $0.773$ ). The levels for each variable were  $\sigma = 0.290$ ,  $0.294$ ,  $0.298$ , and  $0.302$ ;  $\varepsilon = 0.7$ ,  $0.773$ , and  $0.846$ ;  $q_{\text{O}} = -0.6$ ,  $-0.7$ , and  $-0.8$ , and  $q_{\text{H}} = 0.37$ ,  $0.435$ , and  $0.5$ .

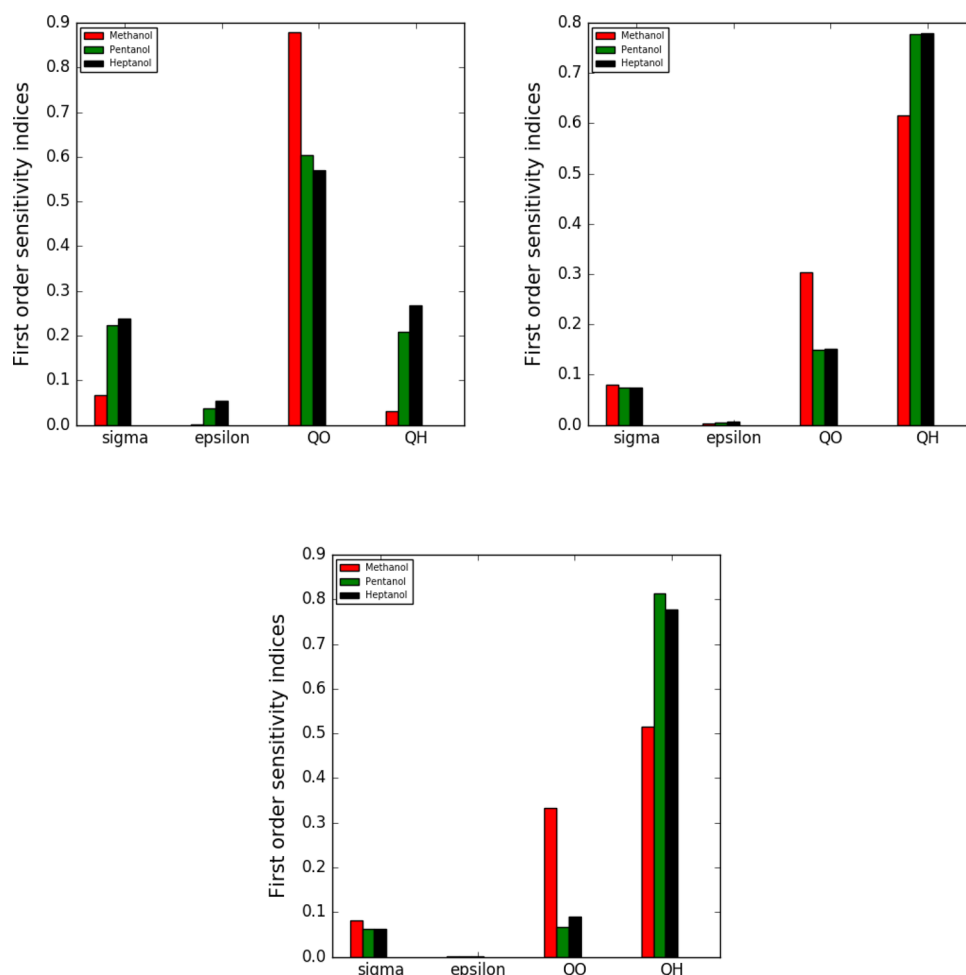
In order to obtain a good estimation of the objective function's minimum, it is important to have a good correlation between the simulated values and the values obtained with the meta-models. This was checked by plotting predicted vs simulated values for each calibration property, and as can be seen from Figure 1, there is a very good match in all cases. The

predictivity coefficients shown in the plots were calculated using eq 16<sup>71</sup>

$$Q_2 = 1 - \frac{\sum_{i=1}^N [y_{\text{sim}}(X_i) - f(X_i)]^2}{\sum_{i=1}^N [y_{\text{sim}}(X_i) - \bar{y}]^2} \quad (16)$$

where  $N$  is the number of points in the sample,  $y_{\text{sim}}(X_i)$  is the simulated value,  $f(X_i)$  is the value predicted by the metamodel, and  $\bar{y}$  is the mean value of  $y_{\text{sim}}(X_i)$  on the whole sample. Additionally, once an optimum was found, new simulations were run using these parameters, and the values predicted by the meta-models were shown to be in agreement with the simulated values.

An ideal force field is one where the difference between experimental and simulated values is zero. Consequently, an objective function was created that takes into account these differences (eq 17)



**Figure 2.** First order sensitivity indices for different properties. Top: density (left) and enthalpy of vaporization (right). Bottom: diffusion constant. The different colors represent the different molecules: methanol (red), 1-pentanol (green), 1-heptanol (black).

$$F(X) = \sum_{j=1}^3 ((f(X) - y_{\text{exp}})_{\rho_j}^2 + (f(X) - y_{\text{exp}})_{D_{A_j}}^2 + (f(X) - y_{\text{exp}})_{\Delta H_j}^2) \quad (17)$$

where  $j = 1$  corresponds to methanol,  $j = 2$  to pentanol, and  $j = 3$  to heptanol. Additionally,  $f(X)$  is the value predicted using the meta-model,  $y_{\text{exp}}$  is the experimental value, and  $\rho$ ,  $D_A$  and  $\Delta H$  represent the density, diffusion, and enthalpy of vaporization, respectively. This function was minimized using an optimized steepest descent algorithm, similar to the one used by Di Pierro and Elber.<sup>73</sup>

The point that describes the original TraPPE force field was the initial guess. From this point, the algorithm searched along the opposite direction of the gradient of the objective function to find a new point. The step length, the distance to move along the specified direction, was found using 100 equidistant trial steps and choosing the one that returns the lowest function value. The maximum step length ( $t_{\text{max}}$ ) tried was obtained from eq 18<sup>73</sup>

$$t_{\text{max}} = \frac{F(X_k)}{\nabla F(X_k) \cdot \nabla F(X_k)} \quad (18)$$

where  $X_k$  is the current iterate, and  $\nabla F(X_k)$  is the gradient of the function evaluated at  $X_k$ .

The optimization routine was done in two parts. First, an optimum was found using larger trial steps (from  $0.01 t_{\text{max}}$  to  $t_{\text{max}}$ ), and then, this optimum was chosen as the starting point for a new optimization algorithm with smaller steps (from  $0.0001 t_{\text{max}}$  to  $t_{\text{max}}$ ). The maximum number of iterations was 4000 for the first part and 100 for the second part because the initial point for the second optimization routine tends to be close to the optimum, and thus, not many extra iterations are needed.

**Sensitivity Analysis.** A variance-based global sensitivity analysis was performed to determine the most influential parameters for each target property. A global sensitivity analysis can be used to identify which factors make no significant contribution to the variance of the output and, therefore, can be fixed to any given value within their range of variation.<sup>74</sup> Sobol' indices<sup>75</sup> can be used to rank the input variables based on their importance. First-order indices  $S_i$  give the influence of each parameter taken alone, while total sensitivity indices  $S_T$  consider the total effect of an input parameter.<sup>76</sup> The importance of interaction effects for a specific parameter depends on the difference between the first-order and total indices. If these two indices are close to each other, interaction effects are not important for that parameter. Furthermore, the sum of all  $S_i$  is equal to 1 for additive models (no interaction between parameters) and less than 1 for

**Table 2.** Comparison between Polarization Corrections Used in This Article ( $C_{\text{pol}}$ ) and the Correction Previously Proposed by Berendsen et al. ( $C_{\text{Berendsen}}$ ) for TraPPE and PolCA Models<sup>a</sup>

solute	solvent	$C_{\text{Distortion}}$	$C_{\text{Electronic}}$	$C_{\text{pol}}$	$C_{\text{Berendsen}}$	
					TraPPE	PolCA
methanol	methanol	-9.251	8.842	-0.409	-2.896	-1.264
ethanol	ethanol	-7.297	7.074	-0.224	-1.907	-0.848
propanol	propanol	-6.169	6.073	-0.096	-1.455	-0.658
butanol	butanol	-5.079	5.101	0.022	-1.233	-0.576
pentanol	pentanol	-4.568	4.702	0.134	-0.890	-0.389
hexanol	hexanol	-3.611	3.860	0.249	-0.899	-0.426
heptanol	heptanol	-3.425	3.731	0.306	-0.637	-0.273
octanol	octanol	-2.783	3.175	0.393	-0.628	-0.284
nonanol	nonanol	-2.155	2.568	0.413	-0.730	-0.370
decanol	decanol	-1.848	2.307	0.459	-0.661	-0.335
2-propanol	2-propanol	-4.948	4.938	-0.010	-1.942	-1.019
2-butanol	2-butanol	-4.560	4.659	0.098	-1.455	-0.746
2-pentanol	2-pentanol	-4.042	4.251	0.209	-0.995	-0.470
tert-butanol	tert-butanol	-4.002	4.347	0.344	-1.160	-0.496
methanol	hexadecane	-1.860	8.730	6.870	-2.896	-1.264
ethanol	hexadecane	-1.168	5.483	4.314	-1.907	-0.848
propanol	hexadecane	-0.851	3.993	3.142	-1.455	-0.658
butanol	hexadecane	-0.658	3.087	2.429	-1.233	-0.576
pentanol	hexadecane	-0.572	2.682	2.111	-0.890	-0.389
hexanol	hexadecane	-0.458	2.152	1.693	-0.899	-0.426
heptanol	hexadecane	-0.429	2.014	1.585	-0.637	-0.273
octanol	hexadecane	-0.367	1.722	1.355	-0.628	-0.284
nonanol	hexadecane	-0.299	1.403	1.104	-0.730	-0.370
decanol	hexadecane	-0.271	1.271	1.000	-0.661	-0.335
methane	octanol	0.000	0.000	0.000	0.000	0.000
ethane	octanol	0.000	0.000	0.000	0.000	0.000
propane	octanol	-0.017	0.020	0.002	0.000	0.000
butane	octanol	-0.005	0.006	0.001	0.000	0.000
pentane	octanol	-0.016	0.018	0.002	0.000	0.000
hexane	octanol	-0.013	0.015	0.002	0.000	0.000
heptane	octanol	-0.012	0.013	0.002	0.000	0.000
octane	octanol	-0.010	0.012	0.001	0.000	0.000
nonane	octanol	-0.009	0.010	0.001	0.000	0.000
decane	octanol	-0.008	0.009	0.001	0.000	0.000

<sup>a</sup>Corrections are for transfer from liquid to gas and are expressed in kJ/mol.

nonadditive models. For this sensitivity analysis, only first-order indices were used.

First-order indices can be evaluated using Monte Carlo simulations to estimate the mean value, the total variance ( $D$ ), and the partial variance due to variable  $x_i$  ( $D_i$ ) (eqs 19–22<sup>76</sup>)

$$f_0 = \frac{1}{N_{\text{sim}}} \sum_{m=1}^{N_{\text{sim}}} f(x_m) \quad (19)$$

$$D = \frac{1}{N_{\text{sim}}} \sum_{m=1}^{N_{\text{sim}}} f(x_m)^2 - f_0^2 \quad (20)$$

$$D_i = \frac{1}{N_{\text{sim}}} \sum_{m=1}^{N_{\text{sim}}} f(x_{(\sim i)m}^{(1)}, x_{im}^{(1)}) f(x_{(\sim i)m}^{(2)}, x_{im}^{(1)}) - f_0^2 \quad (21)$$

$$S_i = \frac{D_i}{D} \quad (22)$$

where  $N_{\text{sim}}$  is the number of samples,  $x_m$  denotes the  $m$ th sample point, and  $x_{(\sim i)m}$  is the  $m$ th sample point without the variable  $i$ . Two matrices of random numbers, called (1) and

(2), and order  $N_{\text{sim}} \times k$  ( $k$  is the number of input variables) need to be generated first. The superscripts (1) and (2) indicate which matrix the sample point needs to be taken from.

The results of the sensitivity analysis for the four fitting parameters considered and for the three target properties (density, enthalpy of vaporization, and self-diffusion coefficient) are shown in Figure 2. The charges on the hydroxyl group are seen to have the strongest influence for all three target properties. This is not surprising, given the importance of hydrogen bonds in the structure and thermodynamics of alcohols.<sup>77–79</sup> The oxygen LJ diameter is of secondary importance, except for density, where it plays a major role. Again, this is to be expected, since density is strongly sensitive to intermolecular packing and hence to excluded volume. At the other extreme,  $\varepsilon$  is shown not to be a relevant factor for any of the target properties of alcohols. As such, to reduce the degrees of freedom of the force field optimization and thus avoid overfitting, we have decided to keep  $\varepsilon$  constant and equal to the TraPPE value in the remainder of this paper.



### 3. RESULTS AND DISCUSSION

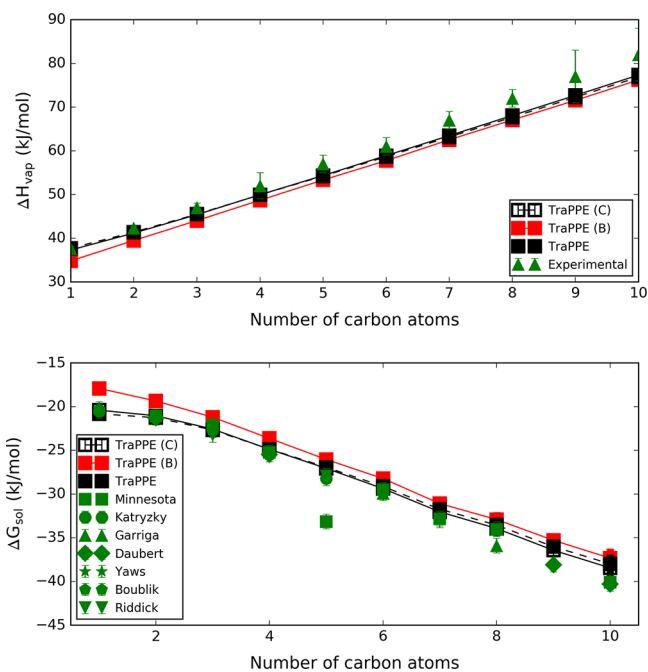
**3.1. TraPPE with Polarization Corrections.** We begin by comparing the performance of the TraPPE-UA model for alcohols with and without applying the polarization corrections described in Section 2.1. This will assess how useful our analytical corrections are when applied to a well-established, previously developed, force field. We note that there is no *a priori* guarantee that the corrections will work well, given that the force field was developed with the standard approach that does not explicitly account for polarization effects. We focus only on properties for which polarization corrections are required, i.e., phase change properties and the dielectric constant.

In Table 2, we report, for all solute/solvent pairs considered in this work, the two contributions to the polarization correction in eq 2, i.e., the negative distortion term and the positive electronic polarization term as well as the final value of the correction. We note that the sign of these terms corresponds to the transfer of a molecule from the liquid phase to the gas phase, as in calculating the enthalpy of vaporization; for the opposite direction, as in solvation free energy calculations, the signs are reversed. We also present the values of the distortion correction estimated using the Berendsen approach (eq 3) for both the original TraPPE model and for the new PolCA model. We reiterate the fact that the Berendsen correction has different values for the two models because it uses the dipole moment of the model as a proxy for the dipole moment of the liquid (for the alcohol molecules presented on this table, the dipole moment of the PolCA model is 2.07 D, see below, while the dipole moment of TraPPE is 2.26). In contrast, our polarization corrections are model independent and can therefore be applied to correct simulation results obtained with any force field.

An interesting observation from Table 2 is that for the pure alcohol systems (i.e., where the solute and solvent are both alcohols), the distortion and electronic contributions nearly cancel out, leaving total polarization corrections that are close to zero (they vary between  $-0.41$  and  $0.46$  kJ/mol). This near cancellation between the two contributions has previously been observed for water.<sup>29,45</sup> Further research is needed to determine if this is a universal feature for molecules with functional groups of high polarity. In contrast, the Berendsen correction is always negative because it only considers distortion effects and is significantly larger in magnitude, particularly for the TraPPE model.

In Figure 3, we compare the predictions of the TraPPE model for the enthalpy of vaporization and self-solvation free energy of alcohols with and without applying polarization corrections. The original TraPPE model, having been parametrized against vapor–liquid coexistence of pure alcohols, performs quite well for these systems. As expected from the values in Table 2, applying our polarization corrections to pure alcohol systems has a very small effect, albeit in the direction of improving agreement with experimental data (see Table 3 for a quantitative comparison of the root-mean-square deviation (RMSD) for each approach). However, applying the Berendsen correction worsens the agreement with experiments, particularly for the smaller alcohols for which those corrections are largest in magnitude.

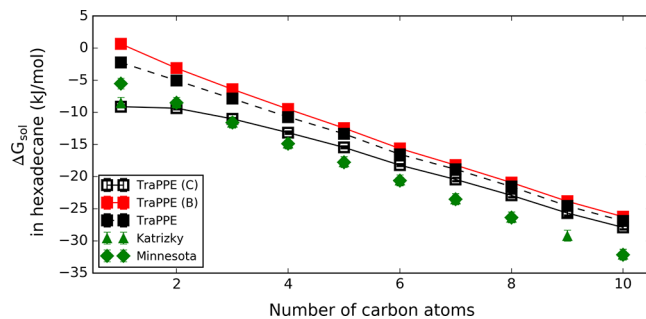
A completely different behavior is observed for solvation of alcohols in hexadecane, as shown in Figure 4. In this case, the original TraPPE systematically overestimates the solvation free



**Figure 3.** Enthalpy of vaporization (top) and free energy of self-solvation (bottom) of primary alcohols obtained with the TraPPE model: without any polarization corrections (“TraPPE”; full black symbols), with the corrections proposed here (“TraPPE(C)”; open black symbols), and with the Berendsen corrections (“TraPPE(B)”; red symbols). The green symbols are experimental values, taken from NIST,<sup>80</sup> for the enthalpy of vaporization and from refs 67, 68, and 81–85 for the self-solvation free energy. The estimated average uncertainty for the experimental free energy was taken as 0.84 kJ/mol.<sup>67</sup>

**Table 3.** RMSD of Dielectric Constant, Enthalpy of Vaporization, and Free Energy of Solvation of Primary Alcohols Using TraPPE, TraPPE with Berendsen Corrections (TraPPE (B)), and TraPPE with Polarization Corrections (TraPPE (C))

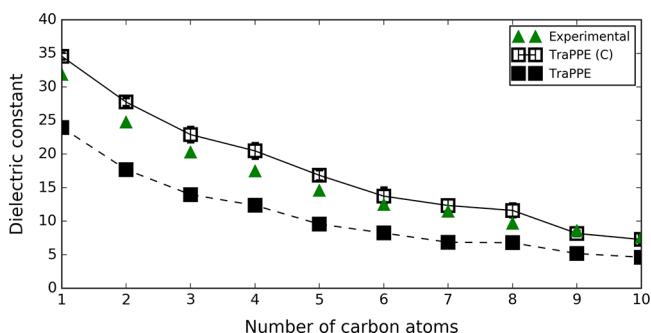
	TraPPE	TraPPE (B)	TraPPE (C)
$\Delta H_{\text{vap}}$ (kJ/mol)	3.20	4.08	2.96
dielectric constant	5.24	N/A	2.06
$\Delta G_{\text{self-sol}}$ (kJ/mol)	1.74	2.53	1.59
$\Delta G_{\text{sol}}$ in hexadecane (kJ/mol)	4.29	5.45	2.83



**Figure 4.** Free energy of solvation of primary alcohols in hexadecane obtained with the TraPPE model: without any polarization corrections (“TraPPE”; full black symbols and dashed line), with the corrections proposed here (“TraPPE(C)”; open black symbols), and with the Berendsen corrections (“TraPPE(B)”; red symbols). The green symbols are experimental values.<sup>67,68</sup>

energy (i.e., predicts less favorable solvation than observed experimentally), showing poor transferability from a polar to a nonpolar environment. This has been a well-documented limitation of fixed-charge force fields, as discussed in the Introduction. Applying the Berendsen correction only makes matters worse, as it shifts the solvation free energies to even less negative values. The reason for this becomes apparent when analyzing Table 2. The dipole moment of the alcohol solutes does not change much when they are moved from the gas phase to an alkane solvent, as the latter is nonpolar and hence does not significantly polarize the solute. However, the background electronic continuum is much the same as for polar solvents; in fact, the electronic response accounts for practically the entire dielectric constant of alkane solvents.<sup>31</sup> As a consequence, the distortion component of polarization is actually quite small for solvation in alkanes and is thus significantly overestimated by the Berendsen approach. In contrast, the electronic contribution retains a similar magnitude as for the pure alcohol systems and thus dominates the polarization effects. This leads to total polarization corrections for the solvation free energy of polar molecules like alcohols, in nonpolar solvents like alkanes, that are quite substantial. Applying these corrections to the TraPPE results significantly improves agreement with experiments (see also Table 3 for a quantitative comparison).

Finally, we analyze the effect of polarization corrections on the dielectric constant predictions in Figure 5. It is clear that



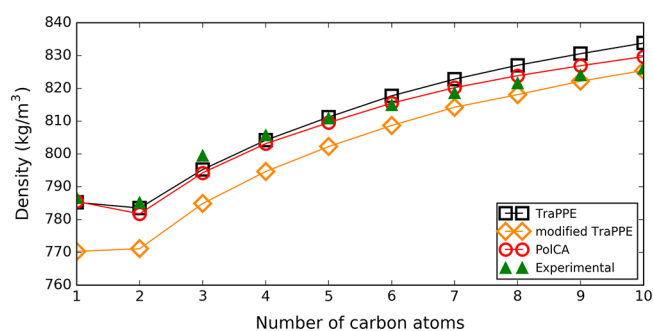
**Figure 5.** Dielectric constant of primary alcohols obtained with the TraPPE model with and without polarization corrections.

the original TraPPE model significantly underpredicts the experimental dielectric constant for all the alcohol molecules. By approximately accounting for both nuclear and electronic polarization effects, eq 7 leads to a much improved agreement with experiments across the entire range. This reinforces the conclusions of a previous study, where the corrections were seen to practically eliminate systematic deviations between simulated and experimental dielectric constants over a wide range of compounds and molecular models.<sup>31</sup> An important difference is that in our previous study the scaling factor  $k$  was adjusted empirically, whereas in the present work, it was derived from first-principles using eq 4.

**3.2. New PolCA Model.** The analysis in the previous section conclusively demonstrates that polarization corrections, as defined by eqs 2 and 7, can significantly improve the performance of an existing model for predicting the enthalpy of vaporization, dielectric constant, and solvation free energies in both polar and nonpolar solvents. However, the TraPPE model did not include any polarization corrections during parametrization. The reason why it works so well for pure alcohols

is because the overall polarization correction for those systems turns out to be very small (Table 2). In this section, we develop a new polarization-consistent model that includes polarization corrections from its inception, as described in Section 2.4.

The TraPPE force field can predict very well the density of small primary alcohols (from methanol to pentanol), but from hexanol to decanol, it overestimates this property (Figure 6).



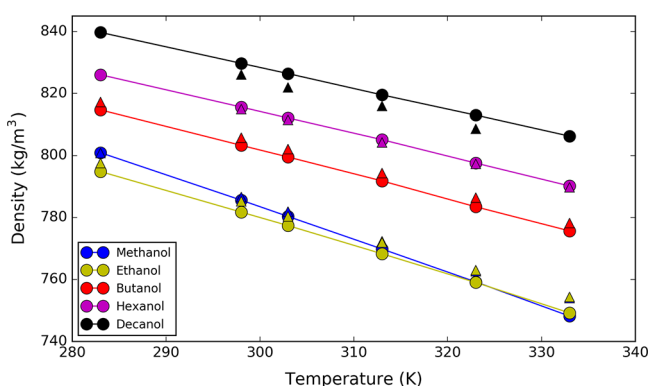
**Figure 6.** Densities of primary alcohols at 298.15 K obtained using the PolCA force field (red circles), original TraPPE force field<sup>69</sup> (black squares), and modified TraPPE force field<sup>34</sup> (yellow diamonds). The green symbols are experimental values.<sup>86</sup>

Also, the difference between simulation and experimental values seems to increase with the length of the alkyl chain, suggesting that the predictions will get worse as we move toward even heavier alcohols. As a preliminary test, the density of each alcohol was calculated using the set of parameters we have previously proposed for hydrocarbons<sup>34</sup> combined with the original parameters for the hydroxyl group taken from TraPPE-UA. We call this model “modified TraPPE”, and the corresponding nonbonded parameters are provided in Table S8. The predicted densities of small alcohols obtained this way were too low, as can be seen in Figure 6. However, from octanol to decanol, the values obtained were closer to the experimental densities compared to those obtained using TraPPE-UA. This can be explained by the fact that the original TraPPE parameters were obtained from calculations of the vapor–liquid equilibria (including saturated liquid densities) of methanol and ethanol,<sup>69</sup> while the new parameters were designed to fit the density, enthalpy of vaporization, and free energy of solvation of alkanes.<sup>34</sup> As the number of carbon atoms in an alcohol increases, the nonpolar interactions become more important, and the behavior of the alcohol tends to be more similar to that of the alkane with the same number of carbon atoms. This is the reason why a new set of LJ parameters for the oxygen and new partial charges for the hydroxyl group and the  $\alpha$ -carbon had to be found. As described in Section 2.4, this was achieved by optimizing those parameters to simultaneously match the density, enthalpy of vaporization, and diffusion coefficient of methanol, 1-pentanol, and 1-heptanol.

The PolCA model performs better than TraPPE at predicting the density of alcohols from hexanol to decanol, while still giving good predictions for smaller alcohols. The root-mean-square deviation of the new model with respect to the experimental data is 2.79 kg/m<sup>3</sup>, while for TraPPE it is 4.23 kg/m<sup>3</sup>.

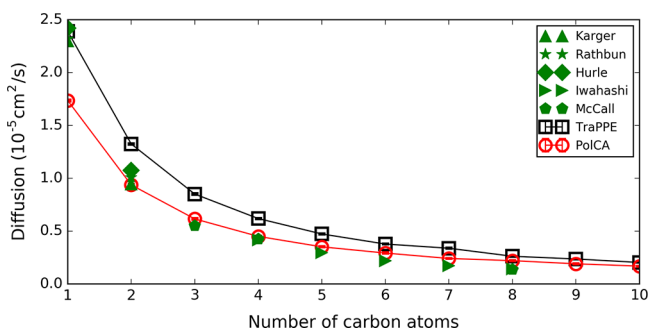
Additionally, simulations were run at five other temperatures (283, 303, 313, 323, and 333 K) for methanol, ethanol, butanol, hexanol, and decanol to validate the new PolCA

model over this range of temperatures. Figure 7 shows that PolCA very accurately captures the temperature dependence of the density for all the alcohols.



**Figure 7.** Simulated density of methanol, ethanol, butanol, hexanol, and decanol at different temperatures using the PolCA force field versus their experimental values. Triangles were used for the experimental values, while circles were chosen for the simulated values. Experimental values were obtained from refs 87–91.

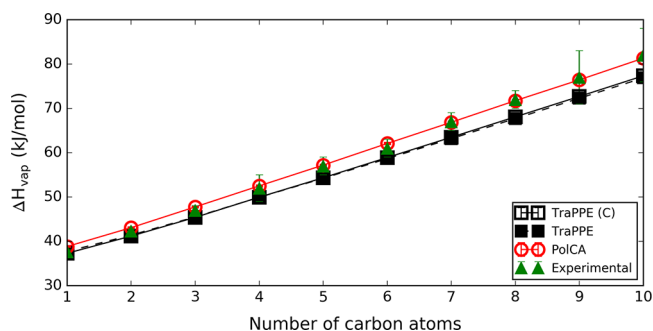
TraPPE overestimates the self-diffusion constant of primary alcohols, except for the diffusion of methanol which is very close to the experimental value, while PolCA underestimates the self-diffusion constant of methanol but accurately predicts this property for the other primary alcohols (Figure 8). The



**Figure 8.** Self-diffusion constant of primary alcohols at 298.15 K obtained using the PolCA model (red circles) and the original TraPPE force field<sup>69</sup> (black squares). The green symbols are experimental values obtained from refs 92–96.

root-mean-square deviation (RMSD) of the new PolCA model with respect to the experimental data is  $0.291 \times 10^{-5} \text{ cm}^2/\text{s}$ , while for TraPPE it is  $0.207 \times 10^{-5} \text{ cm}^2/\text{s}$ . These deviations are quite small, considering the inherent simplifications of both models, and we believe our model strikes a good compromise. It is, of course, possible to improve the accuracy of the fit by increasing the number of degrees of freedom of the model (e.g., varying  $\epsilon$  or scaling the charges of methanol relative to the other alcohols), but that would violate our aim of keeping the model as general and as transferable as possible.

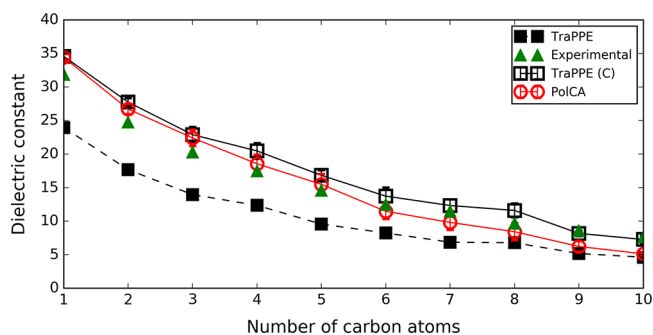
Both TraPPE and PolCA do a very good job at predicting the enthalpy of vaporization. Our model slightly overpredicts this property for methanol, but from ethanol on, it matches the experimental data almost exactly (Figure 9). It is important to mention that experimental values were taken from the NIST Web site,<sup>80</sup> and the error bars for nonanol and decanol are quite large (6 kJ/mol). Therefore, the values obtained using



**Figure 9.** Enthalpy of vaporization of primary alcohols at 298.15 K obtained using the new PolCA model (red circles) and the TraPPE force field with and without polarization corrections (open black squares and filled black symbols, respectively). The green symbols are experimental values obtained from NIST Web site.<sup>80</sup>

TraPPE fall within the error bars but are a lot lower than the average. The RMSD of PolCA is 0.71 kJ/mol, and the RMSD of TraPPE (C) is 2.96 kJ/mol.

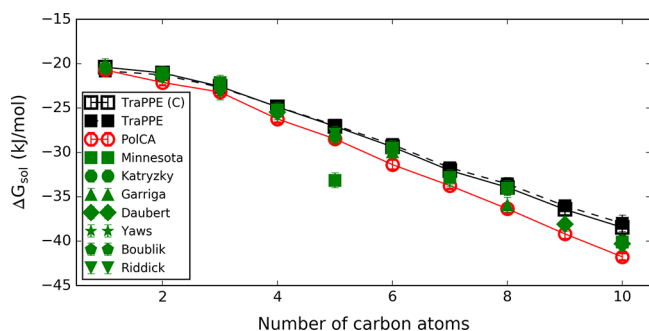
TraPPE and PolCA are both able to accurately predict the dielectric constant once the corrections defined in eq 7 are applied (Figure 10). The root-mean-square deviations for



**Figure 10.** Dielectric constant of primary alcohols at 298.15 K obtained using our new PolCA model (red circles) and the TraPPE force field with corrections (open squares) and without corrections (dashed lines), as a function of the number of carbons. The green symbols are experimental values.

PolCA and TraPPE (C) are 1.84 and 2.06, respectively. This level of agreement is particularly remarkable considering that the simulation results constitute pure predictions; i.e., the dielectric constant was not part of the training set for the parametrization of either TraPPE or PolCA. It further emphasizes the need to consistently account for polarization effects when comparing predictions of nonpolarizable force fields against experimental dielectric constants.

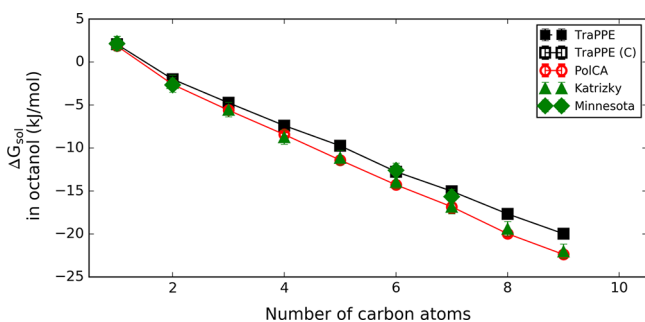
Figure 11 compares the experimental data for the free energy of self-solvation of primary alcohols (i.e., when the solute and solvent are the same molecule) against predictions from TraPPE and PolCA. First of all, we note that the value for pentanol reported in the Minnesota Solvation Database<sup>67</sup> is most likely in error, as it disagrees with the value reported by Katritzky et al.<sup>68</sup> and with estimates based on the vapor pressure and deviates significantly from the trend for the remaining alcohols. With this in mind, and excluding this outlier from the analysis, the RMSD for TraPPE (C) is 1.59 kJ/mol, while it is 1.45 kJ/mol for PolCA. This shows that both models do a good job at predicting this property, although the TraPPE (C) values are on the upper (i.e., less



**Figure 11.** Free energy of self-solvation of primary alcohols at 298.15 K obtained using our new PolCA model (red circles), TraPPE, and TraPPE with corrections (filled squares and open black squares, respectively), as a function of the number of carbons. Experimental values for the free energy were extracted from the Minnesota Solvation Database,<sup>67</sup> the Katritzky database,<sup>68</sup> and from refs 81–85. The estimated average uncertainty for the values extracted from the Minnesota Database is approximately 0.84 kJ/mol;<sup>67</sup> the same uncertainty was used for the other experimental values.

favorable) end of the range of the experimental data, while those of our new model are on the lower end (i.e., more favorable). The polarization corrections for the self-solvation free energy are the same as for the enthalpy of vaporization but with opposite sign; as such, their overall magnitude is also quite small.

The good agreement observed for the free energies of self-solvation reinforces our confidence that the PolCA model is performing well for properties of the pure liquid alcohols. A much more stringent test is to assess its performance in mixtures with different components. Here, we restrict our comparison to mixtures with alkanes because those are the only other molecules for which our polarization-consistent approach has been applied.<sup>34</sup> Figure 12 shows the free energy

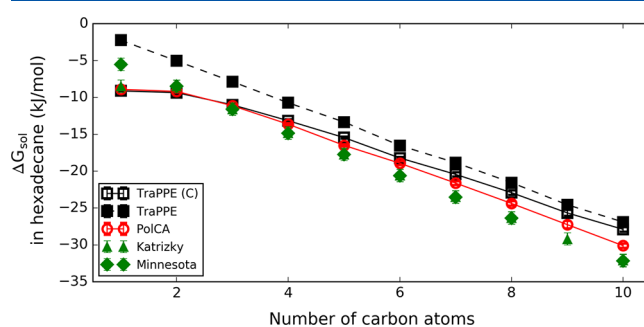


**Figure 12.** Free energy of solvation of alkanes in octanol at 298.15 K as a function of the number of carbons in the solute. Experimental values for the free energy were extracted from the Minnesota Solvation Database<sup>67</sup> and from the Katritzky database.<sup>68</sup> The estimated average uncertainty for the values extracted from the Minnesota Database is approximately 0.84 kJ/mol;<sup>67</sup> the same uncertainty was used when plotting Katritzky's values.

of solvation of a series of alkanes in octanol; i.e., we are testing the ability of the model to describe alcohols as solvents. The PolCA model yields excellent agreement with experimental data (RMSD of 0.90 kJ/mol), although TraPPE also performs quite well (RMSD of 1.10 kJ/mol). It is worth emphasizing that in this case no polarization corrections need to be applied. The alkanes are nonpolar molecules and hence, the electro-

static contribution to the solvation free energy is effectively zero<sup>97</sup> (Table S1).

A more interesting comparison is presented in Figure 13, which shows the free energy of solvation of primary alcohols in



**Figure 13.** Free energy of solvation of primary alcohols in hexadecane at 298.15 K as a function of the number of carbons in the solute. Experimental values for the free energy were extracted from the Minnesota Solvation Database<sup>67</sup> and from the Katritzky database.<sup>68</sup> The estimated average uncertainty for the values extracted from the Minnesota Database is approximately 0.84 kJ/mol;<sup>67</sup> the same uncertainty was used when plotting Katritzky's values.

a hexadecane solvent. Notice that for methanol there is a discrepancy in the two experimental data points available. The performance of both models is identical for small alcohols, but our PolCA model yields improved predictions for larger alcohols due to the reparametrized alkane force field parameters relative to TraPPE. In both cases, the simulation predictions for methanol are in agreement with the experimental data of Katritzky<sup>68</sup> but deviate from that of the Minnesota Solvation Database.<sup>67</sup> Overall, Figure 13 shows that PolCA leads to consistent predictions of solvation of alcohols in alkanes and that polarization corrections are the key to achieve this good performance. The RMSD for PolCA is 1.833 kJ/mol, while for TraPPE (C) the RMSD is 2.828 kJ/mol.

The root-mean-square deviation of each property can be found in Table 4.

**Table 4.** RMSD of Each Property for PolCA and TraPPE-UA

	PolCA	TraPPE (C)
density (kg/m <sup>3</sup> )	2.79	4.23
diffusion (10 <sup>-5</sup> cm <sup>2</sup> /s)	0.29	0.21
ΔH <sub>vap</sub> (kJ/mol)	0.71	2.96
dielectric constant	1.84	2.06
ΔG <sub>self-sol</sub> (kJ/mol)	1.45	1.59
ΔG <sub>sol</sub> in octanol (kJ/mol)	0.90	1.10
ΔG <sub>sol</sub> in hexadecane (kJ/mol)	1.83	2.83

**3.3. Secondary and Tertiary Alcohols.** To increase the generality of the PolCA model, we have extended it to describe also secondary and tertiary alcohols. As a first attempt, the LJ parameters of the alkane CH and C pseudoatoms were used for the α-carbons without modification (i.e., identical to the corresponding parameters in pure alkanes<sup>34</sup>), but this resulted in a model that significantly underpredicted the density. This was somewhat expected since the same effect was noticed by the developers of the TraPPE-UA force field, who had to reduce σ of the α-CH and α-C by 7.48% and 9.38%, respectively.<sup>69</sup> This could be explained by the fact that a C–O

Table 5. Simulated Properties for Nonprimary Alcohols Using the PolCA Model<sup>a</sup>

	2-propanol	2-butanol	2-pentanol	tert-butanol	tert-amyl
density (kg/m <sup>3</sup> )	779.0 ± 0.1	798.8 ± 0.2	804.6 ± 0.3	775.4 ± 0.1	808.3 ± 0.1
	<b>781.31</b> <sup>98</sup>	<b>802.43</b> <sup>98</sup>	<b>805.9</b> <sup>99</sup>	<b>780.79</b> <sup>100</sup>	<b>806.9</b> <sup>99</sup>
$\Delta H_{\text{vap}}$ (kJ/mol)	45.68 ± 0.03	50.88 ± 0.04	55.5 ± 0.1	44.00 ± 0.03	49.6 ± 0.2
	<b>45 ± 3</b>	<b>49 ± 1</b>	<b>53 ± 1</b>	<b>46 ± 1</b>	<b>50 ± 1</b>
dielectric constant	19 ± 1	15 ± 2	12 ± 2	3.08 ± 0.01	2.80 ± 0.01
	<b>19.43</b> <sup>101</sup>	<b>16.60</b> <sup>101</sup>	<b>13.71</b> <sup>99</sup>	<b>12.49</b> <sup>101</sup>	<b>5.78</b> <sup>99</sup>
$\Delta G_{\text{self-sol}}$ (kJ/mol)		-24.2 ± 0.4		-19.6 ± 0.3	
		<b>-23.1 ± 0.8</b> <sup>82</sup>		<b>-21.0 ± 0.8</b> <sup>82</sup>	

<sup>a</sup>Experimental values are in bold. The values for  $\Delta H_{\text{vap}}$  were taken from the NIST website,<sup>80</sup> and in the case of 2-butanol, the average was taken over eight of the nine data points because one seemed to be an outlier.

bond is shorter than a C–C bond and that the hydroxyl oxygen has an electron withdrawing effect over the carbon.<sup>69</sup> As a second attempt, therefore, we decided to also scale down the sigma values for  $\alpha$ -CH and  $\alpha$ -C pseudoatoms by the same factor (i.e., reducing them by 7.48% and 9.38% with respect to the corresponding alkane pseudoatoms). The performance of this model for a few nonprimary alcohols is presented in Table S, from which we can see that the performance of the model is quite satisfactory overall, particularly for secondary alcohols. For tertiary alcohols, the performance deteriorates slightly, and the model is unable to predict the dielectric constant of these molecules.

#### 4. CONCLUSIONS

In this work, we have proposed a new nonpolarizable force field for alcohols that was parametrized taking into consideration polarization effects. Nonpolarizable force fields, which are normally parametrized to match pure liquid properties, tend to fail at predicting the free energy of solvation of polar molecules in nonpolar solvents since they do not explicitly account for changes in the electron distribution of a molecule due to its environment. We have shown in this article that a consistent account of polarization effects through simple and computationally inexpensive *post facto* corrections is able to circumvent this limitation. In fact, our new PolCA model is able to predict pure liquid properties and solvation free energies in different solvents with a high degree of accuracy, including the free energy of solvation of primary alcohols in hexadecane. The static dielectric constant of pure alcohols was also accurately predicted after corrections were added to take into account the difference between the simulated dipole of the molecule and the true dipole moment in the liquid phase.

The new parameters for the alcohol functional group are consistent with a recent united-atom model for aliphatic hydrocarbons<sup>34</sup> that eliminated systematic deviations in solvation free energy predictions from existing UA force fields. When compared to the benchmark TraPPE-UA force field for alcohols, our force field performs better at predicting the density, enthalpy of vaporization, dielectric constant, free energy of self-solvation, free energy of solvation in hexadecane, and free energy of solvation of alkanes in octanol. The RMSD of the diffusion coefficient is slightly higher for PolCA than it is for TraPPE, but this is likely due to an inherent limitation of the UA approach, since our model sacrifices the diffusion of methanol to predict the diffusion of ethanol to decanol with

more accuracy than TraPPE. Indeed, the improved performance of PolCA is seen most significantly for larger alcohol molecules, precisely due to the elimination of systematic errors in the description of the alkane moieties. Nevertheless, it is important to highlight that TraPPE-UA already performs quite well, especially for smaller alcohols, once *post facto* polarization corrections are added. Even though this model was not parametrized taking into account polarization corrections, it performs well because the corrections for properties involving a change of phase in pure alcohols are quite small for alcohols with a short alkyl chain.

In summary, our paper demonstrates that consistently considering polarization effects in the prediction of properties that involve changes in the electrostatic environment leads to improved predictions and increased transferability. Indeed, remarkable improvements were observed when polarization corrections were applied to the predictions of both our new PolCA model and the existing TraPPE-UA force field for alcohols, indicating that such corrections offer a simple route for improving the transferability of generic nonpolarizable models. We expect this work to pave the way for a systematic parametrization of classical nonpolarizable models that adequately accounts for polarization effects, thus achieving a significant increase in accuracy with negligible increase in computational cost. Nevertheless, there are still several aspects of the polarization-consistent approach that can be improved, extended, or generalized. For example, we used an analytical expression to estimate the electronic contribution to the polarization energy that employs the assumption of a spherical solute cavity. Although this was shown to work well even for long-chain alcohols, it is likely that a more sophisticated approach will be needed for nonspherical molecules that contain more than one polar functional group. Furthermore, correction terms that are similar in spirit to those proposed here will still need to be developed for other properties that involve changes in phase, such as melting enthalpies, vapor–liquid equilibrium curves, critical properties, vapor pressures, and surface tensions. In particular, it should be noticed that our approach cannot be applied in its present form to simulations that involve direct coexistence between two phases. Efforts to address all these limitations are currently underway in our research group.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.9b01005>.

Table with all polarization corrections used in this work, finite size corrections for the diffusion, convergence of the free energy calculations, table with all force field parameters, and data from the simulations. (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

Miguel Jorge – Department of Chemical and Process Engineering, University of Strathclyde, Glasgow G1 1XJ, United Kingdom; [orcid.org/0000-0003-3009-4725](https://orcid.org/0000-0003-3009-4725); Email: [miguel.jorge@strath.ac.uk](mailto:miguel.jorge@strath.ac.uk)

### Author

Maria Cecilia Barrera – Department of Chemical and Process Engineering, University of Strathclyde, Glasgow G1 1XJ, United Kingdom; [orcid.org/0000-0002-4554-375X](https://orcid.org/0000-0002-4554-375X)

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jcim.9b01005>

### Notes

All data underpinning this publication are openly available from the University of Strathclyde KnowledgeBase at 10.15129/81a27bde-cf07-4946-9ab6-911138e83e04.

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

M. C. Barrera would like to thank the Funds for Women Graduates (FfWG) for partially funding this work and the University of Strathclyde for a Ph.D. grant.

## ■ REFERENCES

- (1) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. Alchemical free energy methods for drug discovery: progress and challenges. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150–160.
- (2) Loftsson, T.; Brewster, M. E. Pharmaceutical applications of cyclodextrins: basic science and product development. *J. Pharm. Pharmacol.* **2010**, *62*, 1607–1621.
- (3) Lee, Y.-C.; Zocharski, P. D.; Samas, B. An intravenous formulation decision tree for discovery compound formulation development. *Int. J. Pharm.* **2003**, *253*, 111–119.
- (4) Lüder, K.; Lindfors, L.; Westergren, J.; Nordholm, S.; Kjellander, R. In silico prediction of drug solubility. 3. Free energy of solvation in pure amorphous matter. *J. Phys. Chem. B* **2007**, *111*, 7303–7311.
- (5) Savjani, K. T.; Gajjar, A. K.; Savjani, J. K. Drug solubility: importance and enhancement techniques. *ISRN Pharm.* **2012**, *2012*, 1.
- (6) Leo, A.; Hoekman, D. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*; American Chemical Society, 1995; Vol. 1.
- (7) Vrhovsek, A.; Gereben, O.; Jamnik, A.; Pusztai, L. Hydrogen bonding and molecular aggregates in liquid methanol, ethanol, and 1-propanol. *J. Phys. Chem. B* **2011**, *115*, 13473–13488.
- (8) Wertz, D.; Kruh, R. Reinvestigation of the structures of ethanol and methanol at room temperature. *J. Chem. Phys.* **1967**, *47*, 388–390.
- (9) Vij, J.; Reid, C.; Evans, M. Molecular dynamics of methanol: Far infra-red laser spectroscopy, interferometry and computer simulation. *Mol. Phys.* **1983**, *50*, 935–947.

(10) Westergren, J.; Lindfors, L.; Höglund, T.; Lüder, K.; Nordholm, S.; Kjellander, R. In silico prediction of drug solubility: 1. Free energy of hydration. *J. Phys. Chem. B* **2007**, *111*, 1872–1882.

(11) Lüder, K.; Lindfors, L.; Westergren, J.; Nordholm, S.; Kjellander, R. In silico prediction of drug solubility: 2. Free energy of solvation in pure melts. *J. Phys. Chem. B* **2007**, *111*, 1883–1892.

(12) Liu, S.; Cao, S.; Hoang, K.; Young, K. L.; Paluch, A. S.; Mobley, D. L. Using MD simulations to calculate how solvents modulate solubility. *J. Chem. Theory Comput.* **2016**, *12*, 1930–1941.

(13) Paluch, A. S.; Parameswaran, S.; Liu, S.; Kolavennu, A.; Mobley, D. L. Predicting the excess solubility of acetanilide, acetaminophen, phenacetin, benzocaine, and caffeine in binary water/ethanol mixtures via molecular simulation. *J. Chem. Phys.* **2015**, *142*, 044508.

(14) Mobley, D. L.; Wymer, K. L.; Lim, N. M.; Guthrie, J. P. Blind prediction of solvation free energies from the SAMPL4 challenge. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 135–150.

(15) Schnieders, M. J.; Baltrusaitis, J.; Shi, Y.; Chatterjee, G.; Zheng, L.; Yang, W.; Ren, P. The structure, thermodynamics, and solubility of organic crystals from simulation with a polarizable force field. *J. Chem. Theory Comput.* **2012**, *8*, 1721–1736.

(16) Fox, S. J.; Pittcock, C.; Tautermann, C. S.; Fox, T.; Christ, C.; Malcolm, N.; Essex, J. W.; Skylaris, C.-K. Free energies of binding from large-scale first-principles quantum mechanical calculations: application to ligand hydration energies. *J. Phys. Chem. B* **2013**, *117*, 9478–9485.

(17) Berendsen, H.; Grigera, J.; Straatsma, T. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

(18) Abascal, J. L.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.

(19) Horn, H. W.; Swope, W. C.; Pitner, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665–9678.

(20) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building water models: a different approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863–3871.

(21) Wang, L.-P.; Martinez, T. J.; Pande, V. S. Building force fields: An automatic, systematic, and reproducible approach. *J. Phys. Chem. Lett.* **2014**, *5*, 1885–1891.

(22) Weerasinghe, S.; Smith, P. E. A Kirkwood-Buff derived force field for methanol and aqueous methanol solutions. *J. Phys. Chem. B* **2005**, *109*, 15080–15086.

(23) Gonzalez-Salgado, D.; Vega, C. A new intermolecular potential for simulations of methanol: The OPLS/2016 model. *J. Chem. Phys.* **2016**, *145*, 034508.

(24) Jorgensen, W. L. Optimized intermolecular potential functions for liquid alcohols. *J. Phys. Chem.* **1986**, *90*, 1276–1284.

(25) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656–1676.

(26) Martin, M. G.; Siepmann, J. I. Transferable potentials for phase equilibria. 1. United-atom description of n-alkanes. *J. Phys. Chem. B* **1998**, *102*, 2569–2577.

(27) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(28) Leontyev, I.; Stuchebrukhov, A. Electronic continuum model for molecular dynamics simulations. *J. Chem. Phys.* **2009**, *130*, 085102.

(29) Milne, A. W.; Jorge, M. Polarization Corrections and the Hydration Free Energy of Water. *J. Chem. Theory Comput.* **2019**, *15*, 1065–1078.

(30) Vega, C. Water: one molecule, two surfaces, one mistake. *Mol. Phys.* **2015**, *113*, 1145–1163.

(31) Jorge, M.; Lue, L. The dielectric constant: Reconciling simulation and experiment. *J. Chem. Phys.* **2019**, *150*, 084108.

(32) Kulschewski, T.; Pleiss, J. A molecular dynamics study of liquid aliphatic alcohols: simulation of density and self-diffusion coefficient using a modified OPLS force field. *Mol. Simul.* **2013**, *39*, 754–767.

- (33) Horta, B. A.; Fuchs, P. F.; van Gunsteren, W. F.; Hunenberger, P. H. New interaction parameters for oxygen compounds in the GROMOS force field: Improved pure-liquid and solvation properties for alcohols, ethers, aldehydes, ketones, carboxylic acids, and esters. *J. Chem. Theory Comput.* **2011**, *7*, 1016–1031.
- (34) Jorge, M. Predicting hydrophobic solvation by molecular simulation: 2. New united-atom model for alkanes, alkenes, and alkynes. *J. Comput. Chem.* **2017**, *38*, 359–369.
- (35) Jorge, M.; Garrido, N. M.; Simões, C. J.; Silva, C. G.; Brito, R. M. Predicting hydrophobic solvation by molecular simulation: 1. Testing united-atom alkane models. *J. Comput. Chem.* **2017**, *38*, 346–358.
- (36) Bekker, H.; Berendsen, H.; Dijkstra, E.; Achterop, S.; Van Drunen, R.; Van der Spoel, D.; Sijbers, A.; Keegstra, H.; Reitsma, B.; Renardus, M. Gromacs: A parallel computer for molecular dynamics simulations. *Phys. Comput.* **1993**, *92*, 252–256.
- (37) Berendsen, H. J.; Postma, J. v.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (38) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
- (39) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (40) Hockney, R.; Goel, S.; Eastwood, J. Quiet high-resolution computer models of a plasma. *J. Comput. Phys.* **1974**, *14*, 148–158.
- (41) Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (42) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An  $N \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (43) Mendenhall, W.; Beaver, R. J.; Beaver, B. M. *Introduction to Probability and Statistics*; Cengage Learning, 2012.
- (44) Wang, J.; Hou, T. Application of molecular dynamics simulations in molecular property prediction. 1. density and heat of vaporization. *J. Chem. Theory Comput.* **2011**, *7*, 2151–2165.
- (45) Leontyev, I.; Stuchebrukhov, A. Electronic polarizability and the effective pair potentials of water. *J. Chem. Theory Comput.* **2010**, *6*, 3153–3161.
- (46) Kirkwood, J. G. Theory of solutions of molecules containing widely separated charges with special application to zwitterions. *J. Chem. Phys.* **1934**, *2*, 351–361.
- (47) Kirkwood, J. G. The dielectric polarization of polar liquids. *J. Chem. Phys.* **1939**, *7*, 911–919.
- (48) Onsager, L. Electric moments of molecules in liquids. *J. Am. Chem. Soc.* **1936**, *58*, 1486–1493.
- (49) Swope, W. C.; Horn, H. W.; Rice, J. E. Accounting for polarization cost when using fixed charge force fields. I. Method for computing energy. *J. Phys. Chem. B* **2010**, *114*, 8621–8630.
- (50) Badyal, Y.; Saboungi, M.-L.; Price, D.; Shastri, S.; Haefner, D.; Soper, A. Electron distribution in water. *J. Chem. Phys.* **2000**, *112*, 9206–9208.
- (51) Schienbein, P.; Marx, D. Liquid–Vapor Phase Diagram of RPBE-D3 Water: Electronic Properties along the Coexistence Curve and in the Supercritical Phase. *J. Phys. Chem. B* **2018**, *122*, 3318–3329.
- (52) Pagliai, M.; Cardini, G.; Righini, R.; Schettino, V. Hydrogen bond dynamics in liquid methanol. *J. Chem. Phys.* **2003**, *119*, 6655–6662.
- (53) Handgraaf, J.-W.; Meijer, E. J.; Gaigeot, M.-P. Density-functional theory-based molecular simulation study of liquid methanol. *J. Chem. Phys.* **2004**, *121*, 10111–10119.
- (54) Kuo, I.-F. W.; Mundy, C. J.; McGrath, M. J.; Siepmann, J. I. Structure of the methanol liquid–vapor interface: A comprehensive particle-based simulation study. *J. Phys. Chem. C* **2008**, *112*, 15412–15418.
- (55) Sieffert, N.; Bühl, M.; Gaigeot, M.-P.; Morrison, C. A. Liquid methanol from DFT and DFT/MM molecular dynamics simulations. *J. Chem. Theory Comput.* **2013**, *9*, 106–118.
- (56) Abraham, M.; van der Spoel, D.; Lindahl, E.; Hess, B. *GROMACS User Manual*, version 5.0.7, 2015. [www.gromacs.org](http://www.gromacs.org) (accessed January 2020).
- (57) Yeh, I.-C.; Hummer, G. System-size dependence of diffusion coefficients and viscosities from molecular dynamics simulations with periodic boundary conditions. *J. Phys. Chem. B* **2004**, *108*, 15873–15879.
- (58) Neumann, M. Dipole moment fluctuation formulas in computer simulations of polar systems. *Mol. Phys.* **1983**, *50*, 841–858.
- (59) Pohorille, A.; Jarzynski, C.; Chipot, C. Good practices in free-energy calculations. *J. Phys. Chem. B* **2010**, *114*, 10235–10253.
- (60) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245–268.
- (61) Bruckner, S.; Boresch, S. Efficiency of alchemical free energy simulations. I. A practical comparison of the exponential formula, thermodynamic integration, and Bennett’s acceptance ratio method. *J. Comput. Chem.* **2011**, *32*, 1303–1319.
- (62) Steinbrecher, T.; Joung, I.; Case, D. A. Soft-core potentials in thermodynamic integration: Comparing one- and two-step transformations. *J. Comput. Chem.* **2011**, *32*, 3253–3263.
- (63) Wu, D.; Kofke, D. A. Phase-space overlap measures. I. Fail-safe bias detection in free energies calculated by molecular simulation. *J. Chem. Phys.* **2005**, *123*, 054103.
- (64) Goga, N.; Rzepiela, A.; De Vries, A.; Marrink, S.; Berendsen, H. Efficient algorithms for Langevin and DPD dynamics. *J. Chem. Theory Comput.* **2012**, *8*, 3637–3649.
- (65) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; Van Gunsteren, W. F. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem. Phys. Lett.* **1994**, *222*, 529–539.
- (66) Wennberg, C. L.; Murtola, T.; Hess, B.; Lindahl, E. Lennard-Jones lattice summation in bilayer simulations has critical effects on surface tension and lipid properties. *J. Chem. Theory Comput.* **2013**, *9*, 3527–3537.
- (67) Marenich, A. V.; Kelly, C. P.; Thompson, J. D.; Hawkins, G. D.; Chambers, C. C.; Giesen, D. J.; Winget, P.; Cramer, C. J.; Truhlar, D. G. *Minnesota Solvation Database*, University of Minnesota, 2012.
- (68) Katritzky, A. R.; Oliferenko, A. A.; Oliferenko, P. V.; Petrukhin, R.; Tatham, D. B.; Maran, U.; Lomaka, A.; Acree, W. E. A general treatment of solubility. 1. The QSPR correlation of solvation free energies of single solutes in series of solvents. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1794–1805.
- (69) Chen, B.; Potoff, J. J.; Siepmann, J. I. Monte Carlo calculations for alcohols and their mixtures with alkanes. Transferable potentials for phase equilibria. 5. United-atom description of primary, secondary, and tertiary alcohols. *J. Phys. Chem. B* **2001**, *105*, 3093–3104.
- (70) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. Optimized intermolecular potential functions for liquid hydrocarbons. *J. Am. Chem. Soc.* **1984**, *106*, 6638–6646.
- (71) Cailliez, F.; Bourasseau, A.; Pernot, P. Calibration of forcefields for molecular simulation: Sequential design of computer experiments for building cost-efficient kriging metamodels. *J. Comput. Chem.* **2014**, *35*, 130–149.
- (72) Khuri, A. I.; Mukhopadhyay, S. *Response surface methodology* **2010**, *2*, 128–149.
- (73) Di Pierro, M.; Elber, R. Automated optimization of potential parameters. *J. Chem. Theory Comput.* **2013**, *9*, 3311–3320.
- (74) Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S. *Global Sensitivity Analysis: The Primer*; John Wiley & Sons, 2008.
- (75) Sobol’, I. M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.* **2001**, *55*, 271–280.
- (76) Sudret, B. Global sensitivity analysis using polynomial chaos expansions. *Reliab. Eng. Syst. Safte.* **2008**, *93*, 964–979.

- (77) Stephenson, S.; Offeman, R.; Robertson, G.; Orts, W. Hydrogen-bond networks in linear, branched and tertiary alcohols. *Chem. Eng. Sci.* **2007**, *62*, 3019–3031.
- (78) MacCallum, J. L.; Tieleman, D. P. Structures of neat and hydrated 1-octanol from computer simulations. *J. Am. Chem. Soc.* **2002**, *124*, 15085–15093.
- (79) Lehtola, J.; Hakala, M.; Hamalainen, K. Structure of liquid linear alcohols. *J. Phys. Chem. B* **2010**, *114*, 6426–6436.
- (80) NIST Standard Reference Simulation Website, 2018. <https://webbook.nist.gov/> (accessed December 1, 2018).
- (81) Garriga, R.; Martinez, S.; Pérez, P.; Gracia, M. Vapour pressures at several temperatures between 288.15 and 323.15 K of di-n-propylether with 1-hexanol or 1-octanol. Application of the ERAS model. *Fluid Phase Equilib.* **1998**, *147*, 195–206.
- (82) Daubert, T. E.; Danner, R. P. *Physical and Thermodynamic Properties of Pure Chemicals: Data Compilation*; Hemisphere Publishing Corp: New York, 1989.
- (83) Yaws, C. *Handbook of Vapor Pressure*; Gulf Publishing Co.: Houston, TX, 1994.
- (84) Boublík, T.; Fried, V.; Hála, E. *The Vapour Pressures of Pure Substances*; Elsevier Science Publishing Co., Inc.: New York, 1984.
- (85) Riddick, J.; Bunger, W.; Sakano, T. *Organic Solvents*, 4th ed.; Techniques of Chemistry Series; Wiley, Volume II, 1985.
- (86) García, B.; Aparicio, S.; Navarro, A. M.; Alcalde, R.; Leal, J. M. Measurements and modeling of thermophysical behavior of (C1- C4) Alkylbenzoate/(C1- C11) alkan-1-ol mixed solvents. *J. Phys. Chem. B* **2004**, *108*, 15841–15850.
- (87) Mikhail, S.; Kimel, W. Densities and Viscosities of Methanol-Water Mixtures. *J. Chem. Eng. Data* **1961**, *6*, 533–537.
- (88) Kurata, F.; Yergovich, T. W.; Swift, G. W. Density and viscosity of aqueous solutions of methanol and acetone from the freezing point to 10. deg. *J. Chem. Eng. Data* **1971**, *16*, 222–226.
- (89) Mokhtarani, B.; Sharifi, A.; Mortaheb, H. R.; Mirzaei, M.; Mafi, M.; Sadeghian, F. Density and viscosity of 1-butyl-3-methylimidazolium nitrate with ethanol, 1-propanol, or 1-butanol at several temperatures. *J. Chem. Thermodyn.* **2009**, *41*, 1432–1438.
- (90) Alaoui, F.; Montero, E.; Bazile, J.-P.; Comuñas, M.; Boned, C. Liquid density of oxygenated additives to bio-fuels: 1-Hexanol at pressures up to 140 MPa and from 293.15 to 403.15 K. *Fluid Phase Equilib.* **2012**, *320*, 43–48.
- (91) Saleh, M. A.; Akhtar, S.; Begum, S.; Ahmed, M. S.; Begum, S. K. Density and viscosity of 1-alkanols. *Phys. Chem. Liq.* **2004**, *42*, 615–623.
- (92) Karger, N.; Vardag, T.; Lüdemann, H.-D. *J. Chem. Phys.* **1990**, *93*, 3437–3444.
- (93) Rathbun, R.; Babb, A. Self-diffusion in liquids. iii. temperature dependence in pure liquids. *J. Phys. Chem.* **1961**, *65*, 1072–1074.
- (94) Hurle, R. L.; Eastal, A. J.; Woolf, L. A. Self-diffusion in monohydric alcohols under pressure. Methanol, methan (2 H) ol and ethanol. *J. Chem. Soc., Faraday Trans. 1* **1985**, *81*, 769–779.
- (95) Iwashashi, M.; Ohbu, Y.; Kato, T.; Suzuki, Y.; Yamauchi, K.; Yamaguchi, Y.; Muramatsu, M. The dynamical structure of normal alcohols in their liquids as determined by the viscosity and self-diffusion measurements. *Bull. Chem. Soc. Jpn.* **1986**, *59*, 3771–3774.
- (96) McCall, D. W.; Douglass, D. C. Self-Diffusion in the Primary Alcohols. *J. Chem. Phys.* **1960**, *32*, 1876–1877.
- (97) Garrido, N. M.; Jorge, M.; Queimada, A. J.; Macedo, E. A.; Economou, I. G. Using molecular simulation to predict solute solvation and partition coefficients in solvents of different polarity. *Phys. Chem. Chem. Phys.* **2011**, *13*, 9155–9164.
- (98) Awwad, A. M.; Alsayouri, H. M.; Abu-Daibes, M. A.; Jbara, K. A. Densities and volumetric properties of (N-(2-hydroxyethyl) morpholine+ ethanol,+ 1-propanol,+ 2-propanol,+ 1-butanol, and+ 2-butanol) at (293.15, 298.15, 303.15, 313.15, and 323.15) K. *J. Chem. Thermodyn.* **2008**, *40*, 592–598.
- (99) D'Aprano, A.; Donato, D. I.; Agrigento, V. Static dielectric constant, viscosity, and structure of pure isomeric pentanols. *J. Solution Chem.* **1981**, *10*, 673–680.
- (100) Martinez, S.; Garriga, R.; Pérez, P.; Gracia, M. Densities and viscosities of binary mixtures of butanone with butanol isomers at several temperatures. *Fluid Phase Equilib.* **2000**, *168*, 267–279.
- (101) Chmielewska, A.; Zurada, M.; Klimaszewski, K.; Bald, A. Dielectric properties of methanol mixtures with ethanol, isomers of propanol, and butanol. *J. Chem. Eng. Data* **2009**, *54*, 801–806.