Article

Predicting the solvation of organic compounds in aqueous environments: from alkanes and alcohols to pharmaceuticals

Panatpong Hutacharoen, Simon Dufal, Vasileios Papaioannou, Ravi M. Shanker, Claire S Adjiman, George Jackson, and Amparo Galindo

Just Accepted

ACS Publications

Predicting the solvation of organic compounds in aqueous environments: from alkanes and alcohols to pharmaceuticals

Panatpong Hutacharoen,[†] Simon Dufal,[†] Vasileios Papaionanou,[†] Ravi M. Shanker,[‡] Claire S. Adjiman,[†] George Jackson,[†] and Amparo Galindo[*,†]

†Department of Chemical Engineering, Centre for Process System Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom
‡Pfizer Research and Development, Groton Laboratories, Eastern Point Road, Groton, Connecticut 06340, United States

E-mail: a.galindo@imperial.ac.uk

Phone: +44 (0)20 7594 5606

Abstract

The development of accurate models to predict the solvation, solubility, and partitioning of non-polar and amphiphilic compounds in aqueous environments remains an important challenge. We develop state-of-the-art group-interaction models that deliver an accurate description of the thermodynamic properties of alkanes and alcohols in aqueous solutions. The group-contribution statistical associating fluid theory based on potentials of variable Mie form (SAFT-γ Mie) is shown to provide accurate predictions of the phase equilibria, including liquid-liquid equilibria, solubility, free energies of solvation, and other infinite-dilution properties. The transferability of the model is further exemplified with predictions of octanol-water partitioning and solubility for a

range of organic and pharmaceutically relevant compounds. Our SAFT-γ Mie platform is reliable, even for the prediction of challenging properties, such as mutual solubilities of water and organic compounds which can span over ten orders of magnitude, while remaining generic in its applicability to a wide range of compounds and thermodynamic conditions. The work sheds light on contradictory findings related to alkane-water solubility data and the suitability of models that do not account explicitly for polarity.

Introduction

The complex intermolecular solvation interactions between non-polar moieties and water, and in particular the hydrophobic effect,[1–3] play a central role in determining the structure and properties of biomolecules, their compartmentalization and subsequent organization in the living cell.[4] The impact of the hydrophobic effect can be seen in the macroscopic properties of simple mixtures of water and alkanes which exhibit extreme non-ideal fluid-phase behaviour, with very limited miscibility over broad ranges of thermodynamic conditions.[5] The mutual solubilities in the two (water-rich and alkane-rich) coexisting phases are highly asymmetric; the solubility of an alkane in the water-rich phase is several orders of magnitude lower than the solubility of water in the alkane-rich phase. Interestingly, while the solubility of the alkanes in the aqueous phase[6] at conditions of three-phase coexistence presents a minimum at a temperature of ~ 303 K (a phenomenon believed to be related to the hydrophobic effect), the solubility of water in the organic phase increases monotonically with temperature.[7] In mixtures including more functionalized molecules, such as alcohols and larger organic molecules, a delicate balance between the hydrophilic and the hydrophobic interactions of these molecules with water determines unexpected changes in phase behaviour, self-organisation, and segregation that lead to membrane and micelle formation, protein folding, and ligand-protein binding. This balance is also relevant in industrial applications, in refineries, petrochemical, pharmaceutical and biotechnological processes given the ubiquitous presence of water.

In the quest to understand and control the macroscopic behaviour of complex aqueous mixtures, the development of theoretical approaches, grounded in statistical mechanics, that capture the balance of the different intermolecular interactions and deliver bulk properties predictively and accurately remains a key challenge. This can be tackled by starting with simpler systems, such as mixtures of water and alkanes, in which the hydrophobic effect plays a central role, and mixtures of water and alcohols where the impact of hydrophilicity is crucial, then progressing to multifunctional compounds. The phase behaviour and solvation properties (*e.g.*, solubility, partition coefficient, and free energy of solvation) offer a route to determine and probe the effectiveness of models and theories and provide a direct link to applications such as drug development.[8–10] A number of approaches can be adopted, such as direct molecular simulation methods,[11–13] activity coefficient models,[14–17] the conductor-like screening model for real solvent (COSMO-RS),[18] and equations of state.[19,20] From a theoretical perspective equations of state (EoSs) are a particularly useful tool that can access a wide range of states and properties in a computationally efficient way. Classical cubic equations of state[21,22] can be used to describe the fluid-phase behaviour of aqueous systems. Though a degree of success can be achieved with appropriate mixing rules,[23–25] cubic EoSs are either limited accuracy or cannot be applied to aqueous mixtures. Their limited success has been ascribed to their inability to take into account for association and solvation effects explicitly. A breakthrough in modelling aqueous systems was the development of equations of state that explicitly take hydrogen bonding into account, including approaches such as the statistical associating fluid theory (SAFT)[26,27] and the cubic plus association (CPA)[28] EoSs; both stem from the first-order thermodynamic perturbation theory (TPT1) for associating fluids of Wertheim.[29–32] A recent review on the application of molecular-based equations of state for water and aqueous solutions can be found in Ref. 20.

A more challenging task involves the development of group-contribution (GC) equations of state[33] to capture the complex behaviour and subtleties of aqueous systems. GC approaches are based on the premise that the properties of a molecule can be determined from

3

appropriate contributions of the chemically-distinct functional groups that the compound comprises. A major advantage of GC approaches lies in the possibility of describing the properties of a large number of chemical compounds with a relatively small number of group parameters and in allowing property prediction to be carried out without the need for experimental data of the target compounds or mixtures. The GC concept was applied early on to obtain activity coefficients of a broad range of solutes and solvents in semi-empirical methods based on the number of carbon atoms in the solute.[34] Subsequent work, such as the analytical solution of group (ASOG)[15,16] and the universal functional activity coefficient (UNIFAC) models,[14–16,35] focussed on the development of GC activity coefficient approaches with more theoretically sound models. Mengarelli et al.[36] have extended the traditional UNIFAC model by incorporating an association term, based on the Wertheim TPT1 framework and demonstrated a good description of the vapour-liquid equilibria (VLE) of the ethanol + water mixture. More recently, Soares et al.[37] have proposed a functional-segment activity coefficient (F-SAC) model that is based on the concept of functional groups as employed in UNIFAC, but with the interaction energy between groups derived from the COSMO-RS theory. The model has been used to model aqueous hydrocarbons, and was shown to provide an improvement over UNIFAC-type models in the prediction of the infinite-dilution activity coefficient (γ^∞) and liquid-liquid equilibria (LLE) of most systems considered.[38,39] Interestingly, the use of the F-SAC model with an additional parameter for the computation of the association energy allows one to capture the minimum solubility of hydrocarbons in the water-rich phase.[39]

Our focus is devoted to the use of GC EoSs cast within the SAFT framework to model aqueous systems. A further advantage of approaches based on molecular-based EoSs such as SAFT is that they can be employed to develop force field for use in molecular simulation.[40] Given these benefits, there is a growing body of work applying GC EoSs to aqueous systems. The GC-CPA[41] and group contribution plus association (GCA)[42] EoSs have been used to describe the mutual solubilities of hydrocarbons and waterčiteHajiw2014,Pereda2009 as well

as the fluid-phase behaviour of several alcohol + water mixtures.[43,44] A SAFT-based GC approach including multipolar interactions has been used by Nguyen-Huynh *et al.*[45] to model aqueous mixtures. A good overall description of the LLE and VLE of alkane + water and alcohol + water mixtures was acheived, although large deviations from the experimental data was found for the calculated aqueous solubilities of both alkanes and alcohols at ambient temperature. The GC-polar perturbed chain (PPC)-SAFT EoS has also been applied to the prediction of infinite-dilution properties, *e.g.*, Henry's law constants and octanol-water partition coefficients of hydrocarbons and oxygenated compounds.[46,47] Recently, Ahmed *et al.*[48] have used a temperature-dependent diameter for pure water in the GC-PPC SAFT EoS with an additional "non-additive hard sphere" contribution to the Helmholtz energy to attain a satisfactory prediction of the solubility.

Especially relevant to our current work is the recasting of the SAFT EoS for interaction potentials of variable range (VR)[49,50] as a group-contribution EoS (SAFT-γ).[51,52] The fluid-phase behaviour of aqueous solutions of alkanes and alcohols over wide ranges of thermodynamic conditions has been studied in this framework based on the use of square-well (SW) potential;[53] SAFT-γ was found to provide a simultaneous prediction of both the VLE and LLE of the aqueous mixtures using a unique set of transferable group interaction parameters, although the accuracy was seen to deteriorate in the prediction of the alkane solubility in the water-rich phase. The SAFT-γ SW group parameters required to model these mixtures have subsequently been revised,[54] leading to a better description of the properties of the water-rich phase.

Despite this progress, there is, as yet, no generic model that can access the infinitely dilute regime as well as the properties of concentrated aqueous solutions. In attempting to develop such models in the framework of an EoS, the need to incorporate explicitly dipolar interactions to treat water, and other polar molecules, has been the focus of some debate. The original SAFT,[7,55–57] hard-sphere SAFT,[58] SAFT-VR,[59] soft-SAFT,[60] and PC-SAFT[61–63] have been used successfully to study the fluid-phase behaviour of alkane + water mixtures,

5

especially in the correlation of low-pressure phase equilibria (VLE), the prediction of critical boundaries,[58,64] and the description of water solubility in the hydrocarbon-rich phase at high pressure (LLE).[7,56] The models are however noticeably less accurate in the description of the solubilities of hydrocarbons in the water-rich phase, which can be extremely small. These larger deviations have been attributed to the use of an inappropriate model of water and to the inadequacies of some SAFT models in accounting for the high polarity of the aqeuous medium.[60,65,66] The argument, however, remains controversial as a number of studies have demonstrated that SAFT approaches which do not incorporate an explicit dipole (*e.g.*, Refs. 67 and 54) can provide a good description of the *n*-alkane solubility in water; it is worth noting that this is often at the expense of accuracy in the calculated concentration of water in the alkane-rich phase. Al-Saif *et al.*[68] and de Villiers *et al.*[69] have used PC-SAFT with the incorporation of dipole interactions, while Folas *et al.*,[19] Kontogeorgis *et al.*,[70] and Liang *et al.*[67] have used the original PC-SAFT and CPA without the explicit consideration for such interactions to model the VLE and LLE of alcohol + water mixtures. Interestingly, the results for the phase equilibrium calculations are comparable for both approaches, with and without an explicit consideration of the polar interaction.

In our current work, we develop novel models for the accurate prediction of vapour-liquid, liquid-liquid and vapour-liquid-liquid equilibria (VLLE) of aqueous solutions of alkanes and alcohols within the SAFT-γ Mie group contribution approach,[71] based on the SAFT-VR EoS for Mie (generalized Lennard-Jones) potentials of variable range (SAFT-VR Mie).[72] We show that the SAFT-γ Mie EoS provides a very accurate platform for the description of the fluid-phase equilibria over a broad range of compositions of the mixtures, including the infinite-dilution regime, and confirm the predictive capability of the method by calculating partition coefficients and solubilities of a number of pharmaceutical compounds that are particularly challenging due to their highly multifunctional nature. New insights into the modelling of polarity and hydrophobicity in aqueous systems are provided throughout our discussion.

Theory and molecular models

The SAFT-γ Mie group-contribution approach

In the SAFT-γ Mie approach[71,73] molecules are represented as heteronuclear molecules formed from fused spherical segments which correspond to the various chemical moieties and which can be decorated with short-range associating sites if necessary to mediate directional interactions. The Mie intermolecular potential[74] is implemented together with a high-temperature perturbation expansion to third order[72] to provide a high level of accuracy in the description of fluid-phase equilibria and thermodynamic derivative properties. An example of a molecule represented in the SAFT-γ Mie treatment acn be seen in Fig. 1, where the CH_3, CH_2, and CH_2OH functional groups characterizing n-butan-1-ol are shown. A given group k is formed by a number ν_k^* of spherical segments, and a shape factor S_k is used to characterize the contribution of each segment to the overall free energy of the molecule. Two segments k and l are modelled as interacting via Mie[74] potentials of variable repulsive and attractive range:

$$\Phi_{kl}^{\text{Mie}}(r_{kl}) = C_{kl}\varepsilon_{kl} \left[\left(\frac{\sigma_{kl}}{r_{kl}} \right)^{\lambda_{kl}^{\text{r}}} - \left(\frac{\sigma_{kl}}{r_{kl}} \right)^{\lambda_{kl}^{\text{a}}} \right], \tag{1}$$

where r_{kl} is the distance between the centres of the segments, σ_{kl} the segment diameter, ε_{kl} the depth of the potential well (the dispersion energy), and λ_{kl}^{r} and λ_{kl}^{a} the repulsive and attractive exponents of the segment-segment interaction, respectively. The prefactor C_{kl} is a function of the exponents and ensures that the minimum of the interaction is at $-\varepsilon_{kl}$:

$$C_{kl} = \frac{\lambda_{kl}^{\text{r}}}{\lambda_{kl}^{\text{r}} - \lambda_{kl}^{\text{a}}} \left(\frac{\lambda_{kl}^{\text{r}}}{\lambda_{kl}^{\text{a}}} \right)^{\frac{\lambda_{kl}^{\text{a}}}{(\lambda_{kl}^{\text{r}} - \lambda_{kl}^{\text{a}})}}. \tag{2}$$

In common with other SAFT approaches, hydrogen bonding and strongly polar interactions can be treated through the incorporation of short-range square-well association sites, which are placed on any given segment. A segment can have a number $N_{\text{ST},k}$ of different
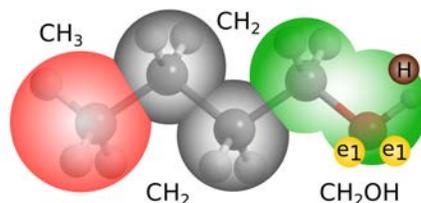
Figure 1: Example of the SAFT-γ decomposition of a molecule into functional groups: n-butan-1-ol is composed of one CH_3 group (shaded red), two CH_2 groups (shaded gray), and one CH_2OH group comprising two fused spherical segments (shaded green) with association sites (brown and yellow) to mimic hydrogen-bonding interactions.

site types, with $n_{k,a}$ sites of type $a = 1, ..., N_{\mathrm{ST},k}$. The association interaction between two square-well sites, one of type a placed on segment k, and a second of type b on segment l, is given by:

$$\Phi_{kl,ab}^{\mathrm{HB}}(r_{kl,ab}) = \begin{cases} -\varepsilon_{kl,ab}^{\mathrm{HB}} & \text{if } r_{kl,ab} \leq r_{kl,ab}^{\mathrm{c}} \\ 0 & \text{if } r_{kl,ab} > r_{kl,ab}^{\mathrm{c}} \end{cases}, \tag{3}$$

where $r_{kl,ab}$ is the centre-centre distance between association sites a and b, $-\varepsilon_{kl,ab}^{\mathrm{HB}}$ is the association energy, and $r_{kl,ab}^{\mathrm{c}}$ the cut-off range of the association interaction. The cut-off range $r_{kl,ab}^{\mathrm{c}}$ can be equivalently described in terms of a bonding volume $K_{kl,ab}$.[75] Each site is positioned at a distance $r_{kk,aa}^{\mathrm{d}}$ or $r_{ll,bb}^{\mathrm{d}}$ from the centre of the segment on which it is placed.

Once the relevant parameters are determined the total Helmholtz free energy A of a mixture of associating heteronuclear chain molecules formed from Mie segments can be obtained from the appropriate contributions of the different groups that make up the molecules as the sum of four separate contributions,[71] as in other SAFT approaches:[26,27]

$$A = A^{\mathrm{ideal}} + A^{\mathrm{mono.}} + A^{\mathrm{chain}} + A^{\mathrm{assoc.}}, \tag{4}$$

where A^{ideal} is the ideal free energy of the molecules, $A^{\mathrm{mono.}}$ the contribution to the free energy due to the segment-segment repulsion and dispersion interactions, A^{chain} the free energy due to the formation of chains of Mie segments, and $A^{\mathrm{assoc.}}$ the contribution due to association. The detailed expressions of the theory and each of the four contributions are

given in Ref. 71 and in Ref. 76 for additional details of the expressions used to describe the association interactions. A summary of the key aspects of the approach can also be found in Ref. 73.

Property calculations

The SAFT-γ Mie equation of state expressed in Eq.(4) provides an algebraic form of the Helmholtz free energy as a function of volume V, temperature T, and the composition vector $\mathbf{N}$ $(N_1, N_2, ...)$ of the mixture. Other properties can be determined through standard thermodynamic relations.[77,78] The pressure p, residual chemical potential μ_i^{res}, and fugacity coefficient $\widehat{\varphi}_i$ of component i in the mixture can be obtained from the Helmholtz free energy as

$$p = - \left. \frac{\partial A(T, V, \mathbf{x})}{\partial V} \right|_{T, \mathbf{N}}, \tag{5}$$

$$\mu_i^{\mathrm{res}}(T, p, \mathbf{x}) = \left. \frac{\partial A^{\mathrm{res}}(T, V, \mathbf{x})}{\partial N_i} \right|_{T, V, N_{j \neq i}} - RT \ln Z(T, p, \mathbf{x}), \tag{6}$$

and

$$\ln \widehat{\varphi}_i(T, p, \mathbf{x}) = \frac{\mu_i^{\mathrm{res}}(T, p, \mathbf{x})}{RT}, \tag{7}$$

respectively, where $A^{\mathrm{res}} = A - A^{\mathrm{ideal}}$ is the residual free energy, $Z = pv_p/(RT)$ is the compressibility factor, with $v_p = V_p/N$ representing the molar volume corresponding to the specified pressure, N is the total number of molecules and $\mathbf{x} = \mathbf{N}/N$. Through the pressure, chemical potential, and fugacity coefficient, the fluid-phase behaviour and solution properties can be calculated. In our current work, we determine the groups parameters using fluid-phase behaviour and limited excess property of mixing data, and assess the performance of the models developed by predicting the solvation properties of n-alkanes and n-alkan-1-ols in aqueous solutions, and examine several pharmaceutically relevant systems.

Fluid-phase equilibria

At a given pressure, temperature, and total composition, the conditions of phase equilibria are solved using the solvers available in the gPROMS software package,[79] and the HELD flash algorithm[80,81] is used to confirm the stability of the equilibrium phases.

Henry's law constant

The Henry's law constant $K_{\mathrm{H}_{i,j}}$ of solute i in solvent j is obtained from measurements of partial pressures of the highly dilute solute over its solution.[77] It is often used as an estimate of solubility, as larger values of $K_{\mathrm{H}_{i,j}}$ usually correspond to lower solubility of the solute in the solvent and vice versa.[64] $K_{\mathrm{H}_{i,j}}$ can be calculated from the fugacity coefficient as

$$K_{\mathrm{H}_{i,j}}(T, p_j^{sat}) = \widehat{\varphi}_{i,j}^{\infty}(T, p_j^{sat}) \cdot p_j^{\mathrm{sat}}, \tag{8}$$

where $\widehat{\varphi}_{i,j}^{\infty}$ is the liquid-phase fugacity coefficient of the infinitely dilute solute i ($i.e.$, $x_i \to 0$) in the mixture at the corresponding saturation pressure of the solution p_j^{sat} for the specified temperature T; in practice, this pressure is calculated at the fluid-phase equilibrium conditions of the mixture for a limiting concentration of the solute using a composition $x_i = 10^{-10}$.

Solvation Gibbs free energy

The solvation Gibbs free energy is defined as the change in Gibbs free energy in transferring a solute molecule from an ideal gas phase to a solution at infinite dilution at constant temperature and pressure. By definition, it is equivalent to the residual chemical potential of a solute i ($cf.$ Eq. 7) at infinite dilution in a solvent j,[64] $i.e.$,

$$\Delta G_{i,j}^{\mathrm{sol}}(T, p) = \mu_{i,j}^{\mathrm{res},\infty}(T, p) = RT \ln \widehat{\varphi}_{i,j}^{\infty}(T, p). \tag{9}$$

The calculations are performed via a single-phase calculation with specified $T = 298.15$

K and $p = 0.100$ MPa. Alternatively, the solvation Gibbs free energy can also be calculated directly from the Henry's constant, based on Eqs. (8) and (9), as[82]

$$\Delta G_{i,j}^{\text{sol}}(T,p) = RT \ln\left[K_{H_{i,j}}(T, p_j^{sat})/p_j^{\text{sat}}\right] + \int_{p_j^{\text{sat}}}^{p} v_{i,j}(T,p)\, dp \approx RT \ln\left[K_{H_{i,j}}(T, p_j^s at)/p_j^{\text{sat}}\right],$$

(10)

where $v_{i,j}$ is the partial molar volume of i in the solution. Note that the contribution of the integral $\int_{p_j^{\text{sat}}}^{p} v_{i,j}(T,p)\, dp$ is generally negligible.[83] A prefactor $M_w/1000$ (where M_w is molar mass of water in g/mol) is used to convert the Henry's constant to a hypothetical one-molal reference solution as employed in the experimental study.[83]

Infinite-dilution activity coefficient

The infinite-dilution activity coefficient $\gamma_{i,j}^{\infty}$ provides an additional measure of the behaviour of a solute molecule i in the solvent environment j. It is an excess property of mixing calculated from the ratio of the fugacity coefficient of the solute in solution at infinite dilution, $\widehat{\varphi}_{i,j}^{\infty}$ (*i.e.*, $x_i = 10^{-10}$), and the fugacity coefficient of the pure solute, φ_i^o (*i.e.*, $x_i \rightarrow 1$), at the same temperature and pressure (via a single-phase calculation):[78]

$$\gamma_{i,j}^{\infty}(T,p) = \frac{\widehat{\varphi}_{i,j}^{\infty}(T,p)}{\varphi_i^o(T,p)}.$$

(11)

Octanol-water partition coefficient

The octanol-water partition coefficient $K_{i,\text{OW}}$ is defined as the ratio of molar concentrations of an infinitely dilute solute i ($C_{i,\alpha}$) distributed in the two equilibrium LLE phases formed by mixtures of n-octan-1-ol and water. It can be related to the ratio of the infinite-dilution activity coefficients of the solute in the water-rich (WR) and octanol-rich (OR) phases with

the molar volumes of the two phases:

$$K_{i,\mathrm{OW}}(T,p) = \lim_{x_i \to 0} \frac{C_{i,\mathrm{OR}}}{C_{i,\mathrm{WR}}} = \frac{v_{\mathrm{WR}}\gamma_{i,\mathrm{WR}}^{\infty}(T,p)}{v_{\mathrm{OR}}\gamma_{i,\mathrm{OR}}^{\infty}(T,p)}, \tag{12}$$

where v_{WR} and v_{OR} denote the molar volumes of the water-rich and octanol-rich phases, respectively, obtained using the SAFT-γ Mie EoS at $T = 298.15$ K and $p = 0.101$ MPa. The infinite-dilution activity coefficients of the solute in the two phases ($\gamma_{i,\mathrm{WR}}^{\infty}$ and $\gamma_{i,\mathrm{OR}}^{\infty}$) are calculated at the same T and p with the specified solute compositions of $x_i = 10^{-10}$ and x_{OR} or x_{WR} at the corresponding calculated octanol-water LLE (the dilute component i is assumed to have no effect on the octanol-water LLE).

Model development and phase equilibrium calculations

To study the fluid-phase behaviour of aqueous mixtures of alkanes and alcohols, the SAFT-γ Mie parameters characterizing the interactions between CH_3, CH_2, CH_2OH and H_2O functional groups are required. The parameter estimation procedure is at the very heart of any group-contribution methodology. In most cases the functional group parameters are estimated from the description of target experimental data of compounds that contain the relevant groups of interest. The procedure is initiated with a chemical family containing "simple" functional groups (*e.g.*, CH_3 and CH_2 of the *n*-alkanes) and compounds that are composed of a single group (*e.g.*, H_2O). Once the parameters for these groups are developed, they are transferred to study other chemical families, comprising some of the established groups, and including additional functional groups. It is important to stress that the heteronuclear formulation of the SAFT-γ Mie approach allows one to estimate unlike parameter between different chemical groups based on pure-component data alone for compounds containing the appropriate chemical functionality. An example for this is the characterization of the CH_3 and CH_2 groups, in which pure-component data of the *n*-alkane homologous series from ethane to *n*-decane are used to estimate the parameters.[71] Vapour-pressure and satu-

rated liquid-density data over a temperature range spanning from the triple point to $\sim 90\%$ of the experimental critical temperature are typically used. Selected thermodynamic data for mixtures can also be employed in the characterization in order to improve the relaibility of the group parameters or to obtain unlike-interaction parameters between a functional group (*e.g.*, CH_2) and a molecular group (*e.g.*, H_2O). A key advantage of group contribution approaches is that it is not necessary to have experimental data for the specific compound of interest when the interactions between the various functional groups are known.

Each functional group k is characterized by the number ν_k^* of identical spherical segments forming the group, the shape factor of the segments S_k, the diameter σ_{kk}, the energy of the segment interaction ε_{kk}, the exponents of the Mie potential λ_{kk}^r and λ_{kk}^a, and the parameters characterizing the association sites - site association interactions, the number $N_{ST,k}$ of different site types, and the number of sites of each type ($n_{k,a}$, $n_k, b, ...$), the energy $\varepsilon_{kk,ab}^{HB}$ and bonding volume parameters $K_{kk,ab}$ for the association between sites of the same or of different type. The interactions between groups of different types k and l are characterized by the corresponding unlike parameters σ_{kl}, ε_{kl}, λ_{kl}^r, λ_{kl}^a, $\varepsilon_{kl,ab}^{HB}$ and $K_{kl,ab}$. In order to reduce the complexity of the parameter estimation procedure a number of parameters, typically ν_k^*, $N_{ST,k}$, $n_{k,a}$, and λ_{kk}^a are pre-assigned fixed values based on the chemical nature of each group, and the unlike interaction parameters σ_{kl} and λ_{kl}^a (and often λ_{kl}^r) are determined by means of appropriate combining rules.[71] The unlike segment diameter σ_{kl} is obtained using the Lorentz-like arithmetic mean of the like diameters:[84] $\sigma_{kl} = \dfrac{\sigma_{kk} + \sigma_{ll}}{2}$. The exponents of the unlike segment-segment interaction can be obtained as: $\lambda_{kl} = 3 + \sqrt{(\lambda_{kk} - 3)(\lambda_{ll} - 3)}$. In our current work, we use previously developed parameters for the CH_3, CH_2, and H_2O groups[71,73] and apply them to obtain new CH_2OH group like- and unlike-interaction parameters. The like and unlike group parameters are summarized in Tables 1 and 2, respectively. The CH_3 and CH_2 functional groups have been examined extensively for their applicability and transferability for a variety of thermodynamic properties and systems.[71] The parameters for the model for H_2O which includes four association sites (two sites of type H and two of

type e_1) to meditate hydrogen bonding are reported in Ref. 76. In addition to the development of the CH_2OH functional group, we refine the group parameters for the interaction between the alkyl CH_3 and CH_2 groups and water reported previously[73] in order to obtain a more accurate description and improved transferability in the modelling aqueous solutions of n-alkane. Novel $CH_2OH - H_2O$ unlike interaction parameters are developed in our current work in order to describe aqeuous mixtures of n-alkan-1-ols.

Table 1: Like group parameters for use within the SAFT-γ Mie 71 group-contribution approach: ν_k^* is the number of segments constituting group k, S_k the shape factor, λ_{kk}^r the repulsive exponent, λ_{kk}^a the attractive exponent, σ_{kk} the segment diameter of group k, and ε_{kk} the dispersion energy of the Mie potential characterizing the interaction of two k groups (with k_B denoting the Boltzmann constant); $N_{ST,k}$ represents the number of association site types on group k, with $n_{k,H}$ and n_{k,e_1} denoting the number of association sites of type H and e_1, respectively. The superscript † indicates the parameters developed in Ref. 71, § those developed in our current work, and ‡ those developed in Ref. 76.

Group k	ν_k^*	S_k	λ_{kk}^r	λ_{kk}^a	$\sigma_{kk}/\text{Å}$	$(\varepsilon_{kk}/k_B)/\text{K}$	$N_{ST,k}$	$n_{k,H}$	n_{k,e_1}
CH_3 †	1	0.57255	15.050	6.0000	4.0773	256.77	-	-	-
CH_2 †	1	0.22932	19.871	6.0000	4.8801	473.39	-	-	-
CH_2OH §	2	0.58538	22.699	6.0000	3.4054	407.22	2	1	2
H_2O ‡	1	1.0000	17.020	6.0000	3.0063	266.68	2	2	2

The alkanol CH_2OH group and the CH_2OH - CH_3 and CH_2OH - CH_2 unlike interaction parameters

In general, the identification/definition of groups in GC approaches is heuristic; the final choice is usually the combination of groups that results in the best representation of the experimental data and the trade-off of delivering transferability. In a group-contribution context, 1-alcohols can be modelled as having the alcohol functional group represented with either an OH or a CH_2OH group. In a recent publication on the modelling of n-alkan-1-ols within the SAFT-γ SW approach,[53] it was shown that the description of this chemical family using a CH_2OH group resulted in more realiable predictions of the fluid-phase behaviour of binary mixtures (including aqueous solutions), whilst retaining an excellent description of the

14

Table 2: Unlike group dispersion interaction energy ε_{kl}, repulsive exponent $\lambda^{\mathrm{r}}_{kl}$, and site-site association energy $\varepsilon^{\mathrm{HB}}_{kl,ab}$, and bonding volume $K_{kl,ab}$ for use within the SAFT-γ Mie [71] group-contribution approach. The unlike segment diameter σ_{kl} is obtained from combining rules (CR)[71] and all unlike attractive exponents are set at the London value of $\lambda^{\mathrm{a}}_{kl} = 6.0000$. The superscript $\dagger$ indicates the parameters developed in Ref. 71, $\S$ those developed in our current work, and $\ddagger$ those developed in Ref. 76.

Group k	Group l		$(\varepsilon_{kl}/k_{\mathrm{B}})$/K	$\lambda^{\mathrm{r}}_{kl}$	site a of group k	site b of group l	$(\varepsilon^{\mathrm{HB}}_{kl,ab}/k_{\mathrm{B}})$/K	$K_{kl,ab}$/Å^3
CH$_3$	CH$_3$	$\dagger$	256.77	15.050	-	-	-	-
CH$_3$	CH$_2$	$\dagger$	350.77	CR	-	-	-	-
CH$_3$	CH$_2$OH	$\S$	333.20	CR	-	-	-	-
CH$_3$	H$_2$O	$\S$	358.18	100.00	-	-	-	-
CH$_2$	CH$_2$	$\dagger$	473.39	19.871	-	-	-	-
CH$_2$	CH$_2$OH	$\S$	423.17	CR	-	-	-	-
CH$_2$	H$_2$O	$\S$	423.63	100.00	-	-	-	-
CH$_2$OH	CH$_2$OH	$\S$	407.22	22.699	e$_1$	H	2097.9	62.309
CH$_2$OH	H$_2$O	$\S$	353.37	CR	e$_1$	H	2153.2	147.40
CH$_2$OH	H$_2$O	$\S$	353.37	CR	H	e$_1$	621.68	425.00
H$_2$O	H$_2$O	$\ddagger$	266.68	17.020	e$_1$	H	1985.4	101.69

pure-component VLE data employed in the development of the group parameters. Following the idea of including the first neighbouring methanediyl group in the definition of the alcohol group (as initially suggested by Wu and Sandler[85]), we CH$_2$OH group in our current work.

The CH$_2$OH group (*cf.* Fig. 1) is modelled with two identical segments ($\nu^{*}_{\mathrm{CH_2OH}} = 2$) and two association site types ($N_{\mathrm{ST,CH_2OH}} = 2$): two sites of type e$_1$ ($n_{\mathrm{CH_2OH,e_1}} = 2$) represent the two electron lone-pairs on the oxygen atom, and one site of type H ($n_{\mathrm{CH_2OH,H}} = 1$) represents the hydrogen atom (corresponding to the 3B association scheme according to Huang and Radosz[86]), where only sites of different type are allowed to interact (*i.e.*, $\varepsilon^{\mathrm{HB}}_{\mathrm{CH_2OH-CH_2OH,e_1e_1}} = \varepsilon^{\mathrm{HB}}_{\mathrm{CH_2OH-CH_2OH,HH}} = 0$). Of the remaining parameters, $\lambda^{\mathrm{a}}_{\mathrm{CH_2OH-CH_2OH}} = 6$ is fixed, so that the task involves the determination of the $S_{\mathrm{CH_2OH}}$, $\lambda^{\mathrm{r}}_{\mathrm{CH_2OH-CH_2OH}}$, $\sigma_{\mathrm{CH_2OH-CH_2OH}}$, $\varepsilon_{\mathrm{CH_2OH-CH_2OH}}$, $\varepsilon^{\mathrm{HB}}_{\mathrm{CH_2OH-CH_2OH,e_1H}}$, $K^{\mathrm{HB}}_{\mathrm{CH_2OH-CH_2OH,e_1H}}$ like parameters and $\varepsilon_{\mathrm{CH_2OH-CH_3}}$ and $\varepsilon_{\mathrm{CH_2OH-CH_2}}$ unlike parameters. These parameters are characterized by comparison with appropriate experimental data. The vapour pressure ($N_{p_{\mathrm{vap}}} = 214$), saturated liquid density ($N_{\rho_{\mathrm{sat}}} = 336$) of the pure n-alkan-1-ols from ethanol to n-decan-1-ol[87] are employed in this regard, as well as some mixture data (molar excess enthalpy of mixing data for n-pentan-1-ol + n-heptane[88]

$(N_{h^E} = 25)$ and LLE data for the n-tetradecane + ethanol[89]($N_{x_{C_{14}}} = 7$ and $N_{y_{C_{14}}} = 9$)). The precise form of the objective function used in the parameter estimation is provided in the Appendix.

The estimated parameters are presented in Tables 1 and 2. The deviation of the theoretical description withy respect to the experimental (exp) data is expressed by means of the percentage average absolute deviation (%AAD) as

$$\%\text{AAD } R_j = \frac{1}{N_{R_j}} \sum_{i=1}^{N_{R_j}} \left| \frac{R_{j,i}^{\text{exp}} - R_{j,i}^{\text{calc}}}{R_{j,i}^{\text{exp}}} \right| \times 100, \tag{13}$$

$$j = \{p_{\text{vap}}, \rho_{\text{sat}}, x_{C_{14}}, y_{C_{14}}\},$$

where N_{R_j} is the number of data points of a property R_j, $R_{j,i}^{\text{exp}}$ represents the experimental value, and $R_{j,i}^{\text{calc}}$ the calculated value for the same property, at the conditions of the ith experimental point. The SAFT-γ Mie provides a good overall description of the VLE with a %AAD of 2.48% for the vapour pressures[87] and 1.82% for the saturated liquid densities[87] of the pure compounds (*cf.* Table 3). For the mixture data,[89] the %AADs of the mole fraction of n-tetradecane in the ethanol-rich and alkane-rich phases are 5.13% and 1.19%, respectively. The average absolute error for the molar excess enthalpy of mixing for the n-pentan-1-ol + n-heptane[88] is 0.20 kJ/mol.

As mentioned earlier, it is possible to employ exclusively pure-component data to characterize the functional group parameters within the SAFT-γ Mie approach. Considering the same model to treat the CH_2OH group, the optimized parameters estimated using pure component vapour-liquid equilibrium data from ethanol to n-decan-1-ol[87] are: $\nu^*_{CH_2OH} = 1$, $S_{CH_2OH} = 0.92122$, $\sigma_{CH_2OH} = 3.6569$ Å, $\lambda^a_{CH_2OH-CH_2OH} = 6.0000$, $\lambda^r_{CH_2OH-CH_2OH} = 11.548$, $\varepsilon_{CH_2OH-CH_2OH}/k_B = 415.91$ K, $\varepsilon^{HB}_{CH_2OH-CH_2OH,e_1H}/k_B = 2510.5$ K, and $K_{CH_2OH-CH_2OH,e_1H} = 19.886$ Å^3. The unlike dispersion energies with the alkyl functional groups on the n-alkan-1-ol are: $\varepsilon_{CH_2OH-CH_3}/k_B = 233.06$ K and $\varepsilon_{CH_2OH-CH_2}/k_B = 391.12$ K. This model of CH_2OH is found to describe accurately the pure-component data for the series of n-alkan-1-ols con-

Table 3: Percentage average absolute deviations (%AAD) for the vapour pressures $p_{\mathrm{vap},i}(T)$ and the saturated liquid densities $\rho_{\mathrm{sat},i}(T)$ obtained with the SAFT-γ Mie group-contribution approach 71 compared to experiment[87] (where $N_{p_{\mathrm{vap},i}}$ and $N_{\rho_{\mathrm{sat},i}}$ are the number of vapour-pressure and saturated-liquid density data points used for each of the n-alkan-1-ols considered in the parameter estimation).

Compound i	T range/K	$N_{p_{\mathrm{vap}},i}$	%AAD $p_{\mathrm{vap},i}(T)$	T range/K	$N_{\rho_{\mathrm{sat}},i}$	%AAD $\rho_{\mathrm{sat},i}(T)$
ethanol	231-463	30	2.55	159-463	39	2.08
propan-1-ol	280-483	25	5.85	169-483	38	1.42
n-butan-1-ol	295-506	26	2.76	186-506	39	1.17
n-pentan-1-ol	278-508	29	0.71	278-508	29	0.95
n-hexan-1-ol	310-428	17	1.40	273-547	38	1.54
n-heptan-1-ol	343-445	14	1.53	273-563	38	1.77
n-octan-1-ol	296-549	31	3.59	263-583	39	2.00
n-nonan-1-ol	366-481	15	1.08	293-596	38	2.41
n-decan-1-ol	301-526	27	2.81	293-613	38	3.01
average	-	-	**2.48**	-	-	**1.82**

sidered, with %AADs for the entire series of 1.56% for the vapour pressures and 1.31% for the saturated liquid densities, which are slightly smaller than those presented in Table 3. One could be tempted to select these as the optimal set of parameters. However, when the two sets of parameters obtained for the interactions involving the CH_2OH group are compared in predictions of the fluid-phase behaviour of binary mixtures containing n-alkanes and n-alkan-1-ols a different conclusion emerges.

Thee fluid-phase behaviour of the mixtures n-heptane + n-pentan-1-ol[90] and n-undecane + n-tetradecan-1-ol[91] is presented in Fig. 2. It is apparent that both models for the CH_2OH group lead to a good description of the VLE of moderate and mixtures of alkanes and long alkanols. The predicted fluid-phase equilibrium of mixtures of alkanes + short alkanols (with a carbon number ≤ 4) obtained with the CH_2OH model estimated solely from pure component data, however, exhibits a region of liquid-liquid immiscibility at high pressure which is not observed experimentally: the predictions for the mixtures n-butane + ethanol[92] and n-heptane + n-butan-1-ol[93] are shown in Fig. 3. The two CH_2OH models are also used for the description of the LLE of the mixture n-tetradecane + ethanol[89] (Fig. 4). Large devi-

ations from the experimental data are observed for the predictions of the phase boundaries when the CH_2OH model estimated from pure component data alone is used; a significantly better description is achieved with the CH_2OH model developed from both pure-component and mixture data. The inability to capture accurately the properties of small polar compounds and their mixtures is a common drawback of group-contribution approaches. The polarization caused by strongly polar groups undermines the underlying assumption of the GC approach that the contribution made by each group to the thermodynamics properties is independent of the molecule on which the group resides.[94] This type of proximity effect becomes more pronounced in the prediction of mixtures involving small polar compounds. Although it is difficult to account for this in the GC approach, techniques involving the use of second-order group interactions[54] or the inclusion of mixture data to develop model parameters allows one to aleviate the problem. We show here that with the inclusion of mixture data (specifically, the LLE of n-tetradecane + ethanol[89] and molar excess enthalpy of mixing of n-heptane + n-pentan-1-ol[88]) in the parameter estimation improves the transferability of the CH_2OH model significantly, as testified in Figs. 2 and 3. It is worth highlighting the accurate prediction obtained for the mixture n-undecane + n-tetradecan-1-ol[91] ($cf.$ Fig. 2b), noting that these two compounds where not included in the estimation of the group parameters.

The alkyl-water CH_3 - H_2O and CH_2 - H_2O unlike interaction parameters

In our current work we re-evaluate the unlike interactions between the alkyl methyl CH_3 and methylene CH_2 segments and H_2O by estimating both ε_{kl} and λ_{kl}^{r} parameters to experimental mixture data. It is not possible to obtain these interactions using pure component data since H_2O is a molecular group. Models for the CH_3 – H_2O and CH_2 – H_2O interactions were developed in a previous study[73] by estimating only the unlike dispersion interaction ε_{kl} using the experimental liquid-liquid equilibrium data for the n-heptane + water. The rest of the parameters (including the range of the repulsive exponent λ_{kl}^{r}) were determined with
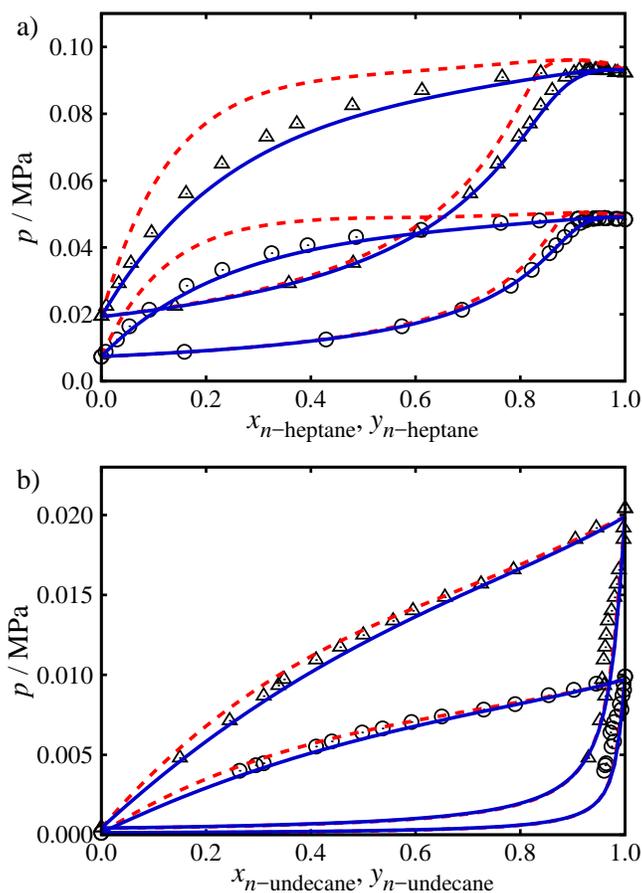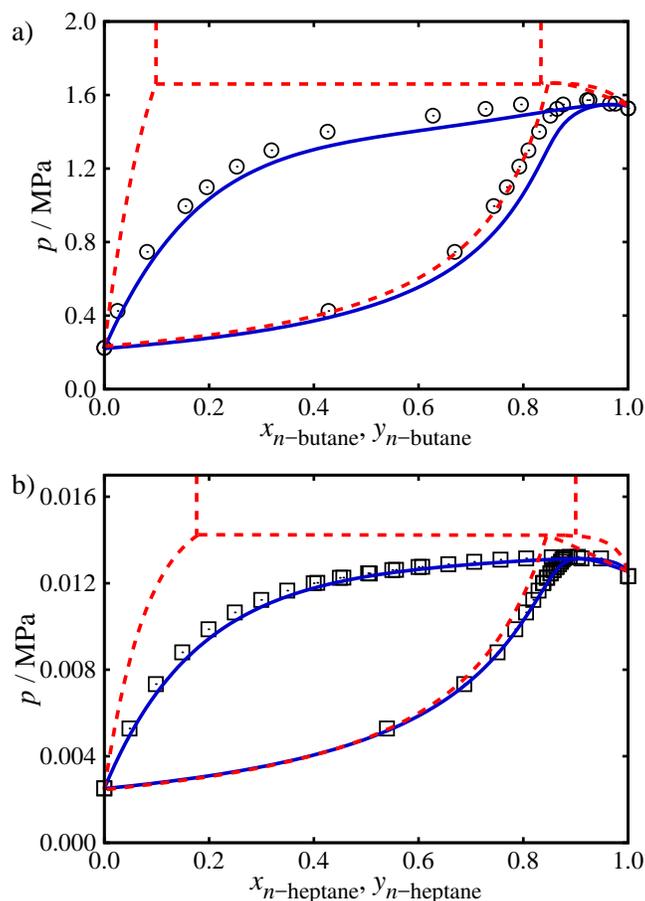
Figure 2: Isothermal pressure–mole fraction $(P–x)$ phase diagrams of the vapour-liquid equilibria of mixtures of n-alkanes and n-alkan-1-ols: (a) n-heptane + n-pentan-1-ol at temperatures of $T = 348.15$ K (circles)[90] and $T= 368.15$ K (triangles)[90] , and (b) n-undecane + n-tetradecan-1-ol mixture at temperatures of $T = 393.15$ K (circles)[91] and $T= 413.15$ K (triangles)[91] . The symbols represent the experimental data, the dashed curves the SAFT-γ Mie predictions obtained with a CH_2OH model based on pure-component data alone, and the continuous curves the SAFT-γ Mie predictions obtained with a CH_2OH model developed from pure-component and mixture data (the parameters are reported in Tables 1 and 2).

Figure 3: Isothermal pressure–mole fraction (P–x) phase diagrams of the vapour-liquid equilibria of mixtures of n-alkanes and short n-alkan-1-ols: (a) n-butane + ethanol at a temperatures of $T = 373.27$ K,[92] and (b) n-heptane + n-butan-1-ol mixture at a temperature of $T = 313.15$ K[93] . The symbols represent the experimental data, the dashed curves the SAFT-γ Mie predictions obtained with a CH_2OH model based on pure-component data alone, and the continuous curves the SAFT-γ Mie predictions obtained with a CH_2OH model developed from pure-component and mixture data (the parameters are reported in Tables 1 and 2).
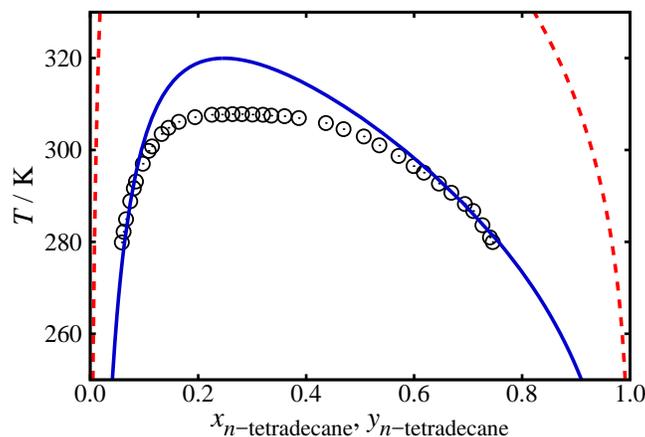
Figure 4: Isobaric temperature–mole fraction (T–x) phase diagram of the liquid-liquid equilibria of n-tetradecane + ethanol at a pressure of $p = 0.101$ MPa[89] . The symbols represent the experimental data, the dashed curves the SAFT-γ Mie predictions obtained with a CH_2OH model based on pure-component data alone, and the continuous curves the SAFT-γ Mie predictions obtained with a CH_2OH model developed from pure-component and mixture data (the parameters are reported in Tables 1 and 2).

combining rules. The previous models[73] provided a good description of the solubilities of the n-alkanes in the water-rich phase only for the medium-chain-length compounds. The solubilities predicted for mixtures with other alkanes were not accurate enough. As well as considering both ε_{kl} and λ^r_{kl} for parameter estimation in our current work, we use three-phase equilibrium (VLLE) solubility data[95,96] for aqeuous mixtures of n-pentane and n-octane over the temperature range of $280 - 400$ K to refine the unlike parameters. Experimental data[97–100] for the coexisting liquid compositions at 298.00 K and pressures along three-phase coexistence for aqueous mixtures of n-hexane to n-decane aqueous mixtures are also included in the parameter estimation ($N_{total} = 42$). The objective function used in the parameter estimation is presented in the Appendix.

The estimated values for the $CH_3 - H_2O$ and $CH_2 - H_2O$ unlike group interaction parameters are summarized in Table 2 and the SAFT-γ Mie description of the three-phase equilibria (VLLE) of the n-alkane + water mixtures are illustrated in Fig. 5. The relatively large values obtained for the unlike dispersion energy parameters ($\varepsilon_{CH_3-H_2O}/k_B = 358.18$ K and $\varepsilon_{CH_2-H_2O}/k_B = 423.63$ K) are compensated by the more repulsive nature of the unlike

potentials ($\lambda^{r}_{CH_3-H_2O} = 100.00$ and $\lambda^{r}_{CH_2-H_2O} = 100.00$), making the overall attractive energy weaker between water and alkyl segments. Despite the use of a relatively small number of experimental mutual solubility data points, the model can be used to describe the solubilities of the shortest alkane (ethane) to the longest alkane considered (*n*-decane) accurately over a wide range of temperature, as is apparent from Fig. 5. Since the solubility of water in the alkane-rich phase is not sensitive function of length of the alkane chain, only the solubilities of water in the *n*-pentane-rich and *n*-decane-rich phases are displayed in Fig. 5 as representative examples of the adequacy of the model. While the description obtained for the solubility of water in the alkane-rich phase is not as accurate as the description of the solubility of the *n*-alkanes in the water-rich phase, the overall quality of the representation is very satisfactory. The solubilities of ethane to *n*-decane in water span over six orders of magnitude in mole fraction over 200 K range of temperatures. It is impressive that the group-contribution method can provide such an accurate description of this behaviour with only a few adjustable parameters.

The aqueous solubilities of various *n*-alkanes at 298 K are assessed separately in Fig. 6. The SAFT-γ Mie predictions are found to be in good agreement with experimental[95–102] and simulation data[12] with the exception of some of the data for the longer alkanes (carbon number > 11) data[103]. The data of Sutton *et al.*[103] exhibit a plateau in the aqueous *n*-alkanes solubilities for alkanes longer than *n*-dodecane, while in the more recent study of Tolls *et al.*[102] a gradual decrease in the solubilities is found for all *n*-alkane studied (up to *n*-pentadecane). Tsonopoulos[6] had previously rationalized the existence of the solubility plateau as related to a "collapsed" conformation of the long alkanes that reduces the contact of the alkanes with water and results in a limiting behaviour for the solubility with increasing carbon number. The exprimental determination of the solubility of highly hydrophobic compounds in water is notoriously difficult given the working solubility range of a part per billion or trillion; this could explain the discrepancies found in the experimental studies. In this context theoretical studies can be useful in the validation of experimental data. Ferguson

et al.[12] have performed molecular dynamics simulations of the solubility of aqueous solutions of n-alkanes, using the TraPPE force field[104] for the n-alkanes and the SPC/E model[105] of water. They do not find plateau in the dependence of the solubility on increasing the carbon number (*cf.* Fig. 6) in agreement with our SAFT-γ Mie predictions. The alkane solubility calculations using previously reported SAFT-γ Mie parameters[73] are also shown in Fig. 6 for reference. It is clear that the new model proposed in our current work provides a more accurate represntation of the behaviour.
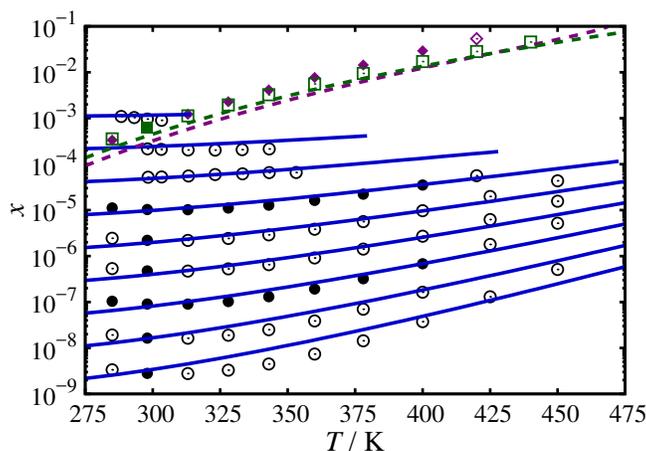
Figure 5: Solubilities of water in the n-alkane-rich liquid phase (dashed curves) and of the n-alkanes (C$_2$-C$_{10}$) in the water-rich liquid phase (continuous curves) obtained with the SAFT-γ Mie group-contribution approach at conditions of three-phase (orthobaric) equilibria as a function of temperature. The diamonds and squares denote experimental correlated data for the solubilities of water in n-pentane-rich and n-decane-rich phases, respectively. The circles correspond to experimental and experimental correlated data for various alkane solubilities in water, ranging from ethane to n-decane. The filled symbols represent the mutual solubility data employed in the estimation of the unlike CH$_3$ – H$_2$O and CH$_2$ – H$_2$O interaction parameters. The data for ethane + water, propane + water and n-butane + water are taken from the work of Mokraoui *et al.*[101] . The data for n-pentane + water to n-decane + water are taken from the IUPAC-NIST solubility data series. [95–100]

The estimated CH$_3$ – H$_2$O and CH$_2$ – H$_2$O interaction parameters are further assessed for transferability by predicting the fluid-phase behaviour of n-alkane + water binary mixtures over a wide range of conditions. Constant temperature pressure-composition slices are shown in Figs. 7 for three mixtures: n-butane + water (at $T = 477.59$ K);[106] n-hexane + water (at $T = 473.15$ K);[107–109] and n-hexadecane + water (at $T = 523.15$ K)[108,110] . Overall, the
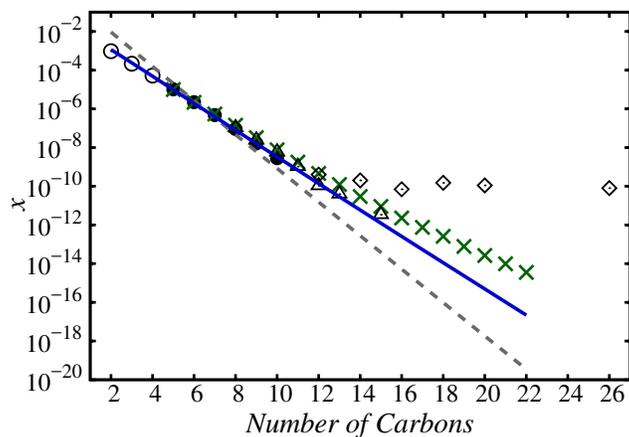
Figure 6: Solubilities of n-alkanes in the water-rich liquid phase at conditions of three-phase (orthobaric) equilibria at $T = 298$ K. The filled circles represent the experimental correlated data from the IUPAC-NIST solubility data series;[95–100] the circles,[101] triangles[102] and diamonds[103] represent the raw experimental data; the crosses[12] represent MD Simulation data; the continuous curve are the SAFT-γ Mie calculations using parameters reported in Tables 1 and 2, and the dash curve the SAFT-γ Mie calculations using previously reported parameters[73] .

predictions are very good in both the VLE and LLE regions for these systems. We emphasize that the simultaneous description of both the VLE and LLE of a system is considered a difficult challenge in with GC approaches. The ability of the method to describe accurately the high-temperature/pressure fluid-phase behaviour of these systems demonstrates the wide range of reliable applicability of SAFT-γ Mie EoS compared to other predictive approaches. It is especially noteworthy that no n-hexadecane data were included in the estimation of the group parameters. As can be seen in Fig. 7c), the SAFT-γ Mie method is shown to predict the fluid-phase behaviour of the highly assymetric n-hexadecane aqueous mixture[108,110] with good accuracy. It is apparent that the newly estimated unlike group interaction parameters allow for an accurate description of all types of fluid phase equilibria considered (VLE, LLE and VLLE) for the aqueous alaknes over a wide range of thermodynamic conditions. In the following section, these parameters are transferred to study the more complex n-alkan-1-ols + water mixtures.
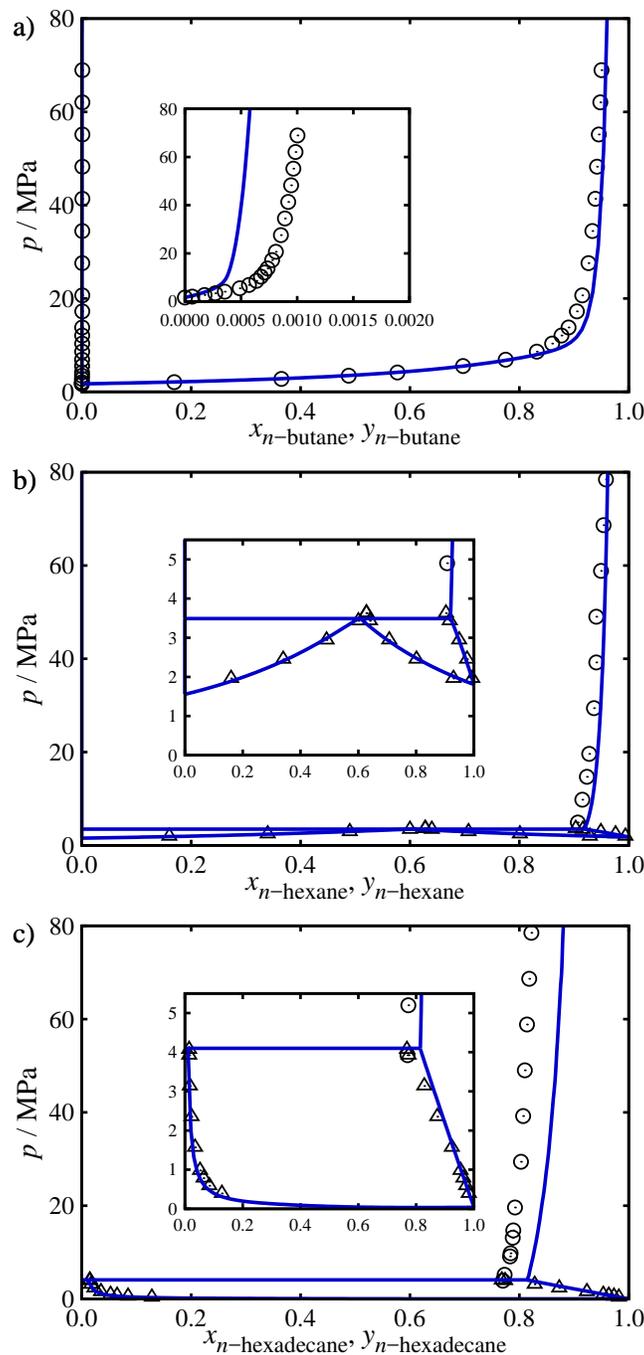
Figure 7: Isothermal pressure–mole fraction (P–x) phase diagrams of (a) n-butane + water at a temperature of $T = 477.59$ K[106] (which is above the critical point of n-butane, $T_{c,\mathrm{C_4H_{10}}}$ $= 425.12$ K[111]), (b) n-hexane + water at a temperature of $T = 473.15$ K (VLE[107] and LLE[108,109]), and (c) n-hexadecane + water at a temperature of $T = 523.15$ K (VLE[110] and LLE[108]). The symbols represent the experimental data, the continuous curves the prediction with the SAFT-γ Mie approach, and the horizontal lines the calculated three-phase line. The inset image in (a) corresponds to a magnified view of the water-rich phase, and the inset images in (b) and (c) correspond to a magnified view of the VLE region.

The alkanol-water CH_2OH - H_2O unlike interaction parameters

The addition of hydroxyl functional groups to the alkyl chain to form alkanols results in dramatic changes in the physical properties of the system. Weak attractive van der Waals interactions dominate the phase behaviour of the very hydrophobic n-alkanes and water, while in the case of aqueous mixtures of alkanols the strong attractive unlike interactions mediated by hydrogen bonding are also important. While all n-alkanes are markedly immiscible, a number of alcohols are fully miscible in water. In particular, aqueous mixtures of the shorter homologues of the alkan-1-ol series, *i.e.*, methanol, ethanol, and propan-1-ol[112] exhibit homogeneous liquid phases. regions of liquid-liquid immiscibility appear for aqueous solutions of longer chains, *i.e.*, in n-butan-1-ol and longer n-alkan-1-ols.[112] Modelling mixtures of alkanols and is challenging as the fluid-phase behaviour is determined by the relative magnitudes of the unlike dispersion energy and hydrogen-bonding interactions. The simultaneous description of the VLE and LLE phase equilibria with a set of transferable parameters constitutes a stringent test of any GC model.

To complete the development of a model for alkanol + water mixtures, the unlike interaction parameters between the CH_2OH and H_2O groups are estimated building on the CH_3 – H_2O and CH_2 – H_2O parameters established in the previous section. Both CH_2OH and H_2O are associating groups, modelled with 3B and 4C association schemes, respectively, following the notation of Huang and Radosz.[86] The unlike association interactions between the two groups are assumed to be asymmetric, *i.e.*, $\varepsilon_{CH_2OH-H_2O,e_1H}^{HB} \neq \varepsilon_{CH_2OH-H_2O,He_1}^{HB}$ and $K_{CH_2OH-H_2O,e_1H} \neq K_{CH_2OH-H_2O,He_1}$. Although in SAFT approaches it is common to model the alcohol-water hydrogen-bonding interaction as symmetric (i.e., with the same value for the hydrogen bonding energy and volume for the O-H interaction when it involves an oxygen atom of water or of the alcohol), from a physical point of view the interactions are expected to be different; see the work of Fileti *et al.*[113] in which differences can be seen in the calculated hydrogen-bonding distance and dipole moment between alcohol-water and water-alcohol heterodimers. Here we estimate the unlike dispersion energy ($\varepsilon_{CH_2OH-H_2O}$),

the unlike association energies ($\varepsilon^{\mathrm{HB}}_{\mathrm{CH_2OH-H_2O,e_1H}}$ and $\varepsilon^{\mathrm{HB}}_{\mathrm{CH_2OH-H_2O,He_1}}$), and the unlike bonding volumes ($K_{\mathrm{CH_2OH-H_2O,e_1H}}$ and $K_{\mathrm{CH_2OH-H_2O,He_1}}$) using experimental data for the LLE of n-octan-1-ol + water ($N_{x_{\mathrm{oct}}} = 8$ and $N_{y_{\mathrm{oct}}} = 8$) at $p = 0.101$ MPa[114] . The unlike segment diameter ($\sigma_{\mathrm{CH_2OH-H_2O}}$), the unlike repulsive range ($\lambda^{\mathrm{r}}_{\mathrm{CH_2OH-H_2O}}$), and the unlike attractive range ($\lambda^{\mathrm{a}}_{\mathrm{CH_2OH-H_2O}}$) are prescribed using combining rules.[71] The corresponding objective function is given in the Appendix. The optimal values for the unlike interaction parameters are summarized in Table 2 and the resulting description of the liquid-liquid equilibrium of the binary mixture of n-octan-1-ol + water is displayed in Fig. 8a). As can be seen in the figure the estimated values of the unlike interaction parameters result in an excellent description of the data[114,115] . In addition to the description of the fluid-phase behaviour of the n-octan-1-ol + water binary mixture, the SAFT-γ Mie predictions of the densities of the two coexisting liquid phases are found to be in excellent agreement with the experimental data[114] (cf. Fig. 8b). The objective function used in the parameter estimation is presented in the Appendix.

Two examples of the predictive capabilities of the model based on n-hexan-1-ol + water and n-butan-1-ol + water are shown in Fig. 9,. The SAFT-γ Mie predictions of both binary mixtures can be seen to be in good agreement with the experimental data[116–119] for the fluid-phase behaviour of these systems. Both the VLE and the LLE regions of the phase envelope are well represented, together with the location of the three-phase line. The accuracy of the description of the solubility of the alkanols in the water-rich phase (cf. the magnified regions in Figs. 8 and 9) is exceptional considering the predictive nature of the calculations. While at ambient conditions n-butan-1-ol + water exhibits liquid-liquid demixing, the shorter alcohols are completely miscible with water.[112] The complete miscibility of ethanol + water and propan-1-ol + water is also predicted with our model. This suggests that the right balance between the hydrophobic and hydrophilic interactions is captured in the model. We emphasize again that this is achieved while using a unique set of transferable interaction parameters for all of the alkanol-water systems. The transferability of the model is to a
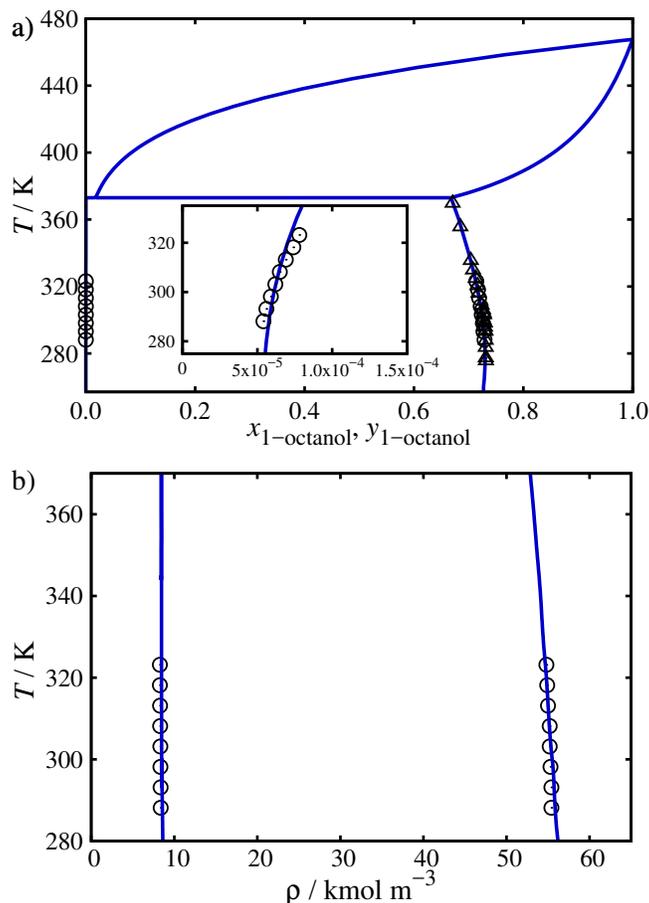
Figure 8: (a) Isobaric temperature–mole fraction (T–x) phase diagram of the vapour-liquid-liquid equilibria of n-octan-1-ol + water at a pressure of $p = 0.101$ MPa; (b) saturated densities of the water-rich (right) and octanol-rich (left) phases. The circles[114] and tri-angles[115] represent the experimental data, the continuous curves the description with the SAFT-γ Mie approach, and the horizontal line the calculated three-phase line. The inset image in figure (a) corresponds to a magnified view of the water-rich boundary.

large degree due to the reliability of the alkyl-water interactions, as well as to the use of the asymmetric association parameters in the description of the $CH_2OH - H_2O$ interactions.
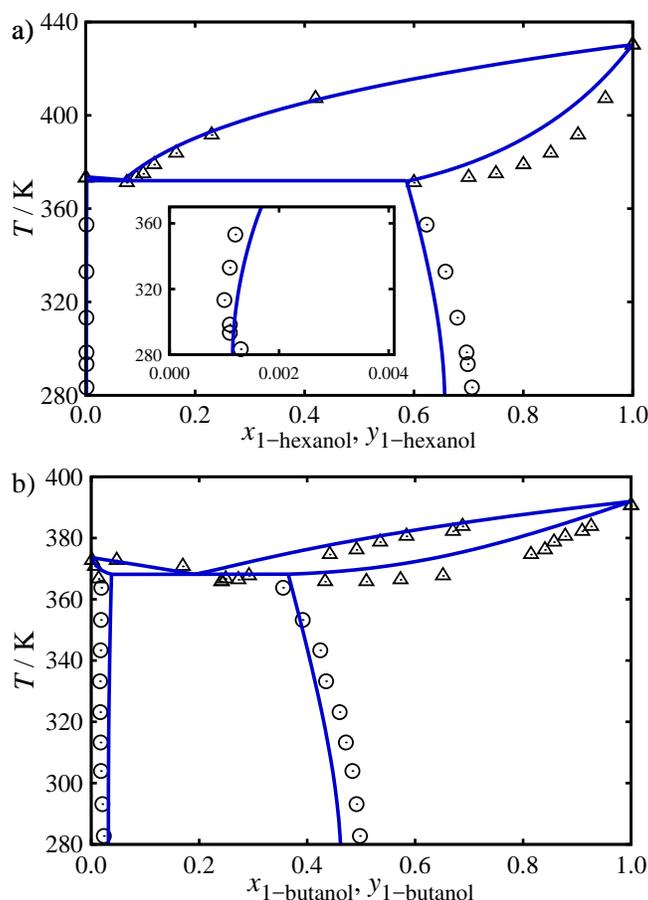

Figure 9: Isobaric temperature–mole fraction (T–x) phase diagrams of the vapour-liquid-liquid equilibria at $p = 0.101$ MPa of (a) n-hexan-1-ol + water (triangles[116] and circles[117]) and (b) n-butan-1-ol + water (triangles[118] and circles[119]). The symbols represent the experimental data, the continuous curves the predictions with the SAFT-γ Mie approach and the horizontal lines the predicted three phase lines. The inset image in figure (a) corresponds to a magnified view of the water-rich boundary.

Prediction of solvation properties

An important application of the models developed here is the study of the infinite-dilution (solvation) properties. To the best of our knowledge, reports of the prediction of different types of fluid-phase equilibria (VLE, LLE, and VLLE) over the entire range of composition including the infinite-dilution region with GC approaches are very scarce. Pereda *et*

al.[43,44,120,121] have used the GCA-EoS model to predict the VLE, LLE, and infinite-dilution activity coefficients of mixtures containing water, alcohols, and hydrocarbons, although different sets of alkyl-water unlike parameters (H_2O-CH_3^∞ and H_2O-CH_2^∞) were required to provide accurate predictions of the mutual solubility of water and hydrocarbons. Possani *et al.*[39] have used the F-SAC model to represent the mutual solubilities as well as the infinite-dilution activity coefficients of aqueous hydrocarbon, finding good agreement with the experimental data, although it should be noted that experimental $\gamma_{i,j}^\infty$ data were used in the calibration of the model parameters. A unique set of SAFT-γ Mie group parameters (*cf.* Tables 1 and 2) is employed in our work to predict the phase equilibria as well as the infinite-dilution properties for the hydrocarbon and alkanol aqueous systems without the need to include the experimental infinite-dilution data to calibrate the models or further adjust any parameters.

Numerous experimental studies[101,122–124] report the distribution of a solute between water and a gaseous phase in terms of Henry's law constants $K_{H_{i,j}}$. Experimental data of the $K_{H_{i,j}}$ for ethane, *n*-butane, *n*-hexane, and *n*-octane in water over a temperature range are compared with the corresponding SAFT-γ Mie predictions in Fig. 10. It is apparent from the figure our predictive approach is found to represent correctly the experimental measurements for the different hydrocarbons considered. In addition, the solvation Gibbs free energies $\Delta G_{i,j}^{\text{sol}}$, which are directly related to the Henry's law constants (*cf.* Eq. 10), are presented in Fig. 11 for a series of *n*-alkanes and *n*-alkan-1-ols in an aqueous environment at $T = 298.15$ K and $p = 0.100$ MPa. The calculations are performed as single-phase calculations, *i.e.*, using Eq. (9) with specified T, p, and $x_i = 10^{-10}$. In the case of the shorter alcohols, however, the coexisting composition of the alcohol in the aqueous environment is far from the dilute limit at the specified T and p, and therefore a single-phase calculation in which the composition of the solute is specified explicitly is employed. As expected from the hydrophobic nature of the *n*-alkanes, large positive values are seen in $\Delta G_{i,j}^{\text{sol}}$ for these molecules (*cf.* Fig. 11a); we note that the values increase as the chain length of the hydrocarbon increases. On the other hand,

the hydration of the alkanols is favourable, with corresponding negative values of $\Delta G_{i,j}^{\text{sol}}$ (cf. Fig. 11b). The values of $\Delta G_{i,j}^{\text{sol}}$ of the n-alkanes (Fig 11a) predicted with the SAFT-γ Mie GC approach are in excellent agreement with the experimental data[83], especially for ethane to n-undecane, with slightly larger deviations observed for the longer n-alkanes. The level of agreement seen here is consistent with the predictions of the solubility of the n-alkanes in water shown earlier in Fig. 6. The Gibbs free energy of hydration can be obtained from experimental aqueous solubility data (or from Henry's law constants or activity coefficients of a solute in aqueous solution at infinite dilution)[83]. The larger uncertainties reported in the experimental data[83] for $\Delta G_{i,j}^{\text{sol}}$ data for the longer alkanes are due to the higher uncertainty in the solubility measurements of these highly hydrophobic compounds. While our solvation Gibbs free energy predictions for the n-alkanes are in overall good agreement with the experimental data[83] (cf. Fig. 11a), a larger deviation is observed in the case of the shorter alkanols (cf. Fig. 11b). The assumption of the transferability of group parameters is less applicable in these small polar molecules due to proximity effects.
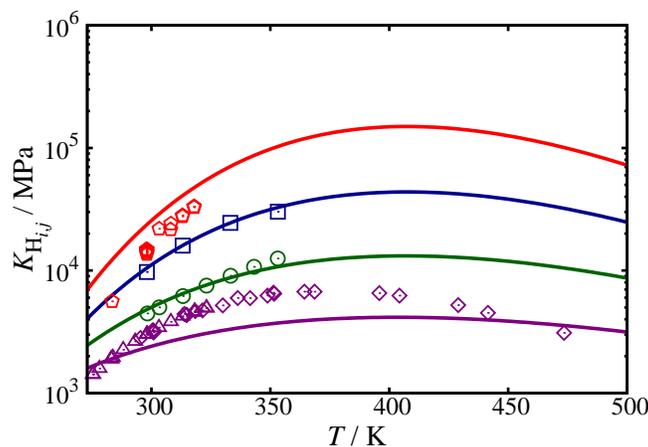

Figure 10: Henry's law constants $K_{\text{H}_{i,j}}$ for n-alkanes in water as a function of temperature at the corresponding vapour pressure of the mixture. The symbols represent the experimental data and the continuous curves the predictions with the SAFT-γ Mie approach: ethane (triangles[122] and diamonds[123]), n-butane (circles[101]), n-hexane (squares[101]) and n-octane (pentagons[124]).

The infinite-dilution activity coefficient $\gamma_{i,j}^{\infty}$ provides a measure of solvation with reference to the fugacity coefficient of the pure solute (cf. Eq. 11). It describes the behaviour of
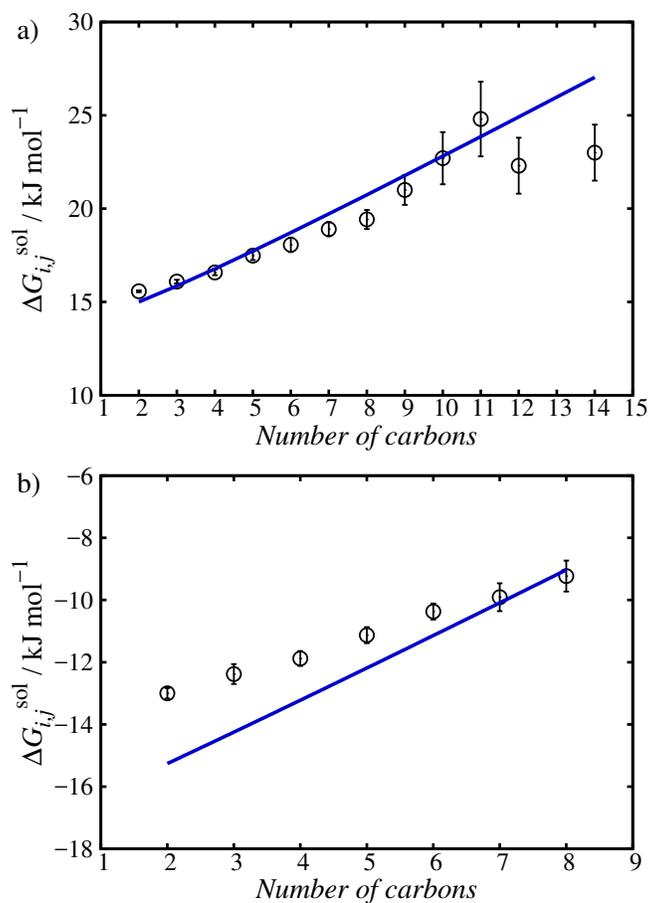
Figure 11: Gibbs free energies of solvation $\Delta G_{i,j}^{\text{sol}}$ for (a) n-alkanes and (b) n-alkan-1-ols in water at $T = 298.15$ K and $p = 0.100$ MPa. The circles[83] represent the experimental data, the error bars the standard deviations, and the continuous curves the predictions with the SAFT-γ Mie approach.

a solute i entirely surrounded by solvent j and reflects the maximum deviation from the ideal solution behaviour. Reliable information on $\gamma_{i,j}^{\infty}$ is of importance in predicting the fate of chemicals in the environment.[125] It is apparent from Figs. 12 and 13 that $\gamma_{i,j}^{\infty}$ of n-alkanes and n-alkan-1-ols in water predicted with the SAFT-γ Mie model are also in good agreement with the experimental data[102,126]. Some deterioration of the agreement is seen for the longer n-alkanes (carbon number > 11)[103], though we note there are discrepancies among the different sets of experimental data[103] and simulation results.[12] As in the case for $\Delta G_{i,j}^{\text{sol}}$, the $\gamma_{i,j}^{\infty}$ data for long n-alkanes are commonly obtained from aqueous solubility data[102,103] (cf. Fig. 6), which can be extremely difficult to measure. The same argument can therefore be used to explain the disagreement between the experimental and theoretical values as for the aqueous solubility and $\Delta G_{i,j}^{\text{sol}}$ of the longer n-alkanes. The quantitative description of most of the experimental data and with the simulation data suggests that our theoretical approach can be used to validate the experimental measurements for the longer n-alkanes. The predicted values of $\gamma_{i,j}^{\infty}$ of n-alkan-1-ols at different temperatures are compared with experimental data in Fig. 13. As can be seen very good agreement with the experimental data is observed for the entire temperature range available for compounds larger than n-butan-1-ol. We only consider n-alkan-1-ols larger than n-butan-1-ol since, as was pointed out earlier, our group-contribution method is not as accurate for the shorter alcohols. The level of agreement observed for the longer alaknols suggests that our SAFT-γ Mie model should also provide accurate predictions of $K_{i,\text{OW}}$ given the relation between the two thermodynamic quantities (cf. Eq. (12)).

The octanol-water partition coefficient $K_{i,\text{OW}}$ is an infinite-dilution property describing the distribution of a solute i across coexisting octanol-rich and water-rich phases which are of significantly different polarity. At ambient conditions, water and n-octan-1-ol are partially miscible, the mixture exhibits an octanol-rich liquid phase of composition $x_{\text{oct}} = 0.726$, and a water-rich liquid phase of composition $x_{\text{oct}} = 5.60 \times 10^{-5}$.[114] The activity coefficients contained in Eq. (12) strictly refer to a three-component mixture (water, n-octan-1-ol, and
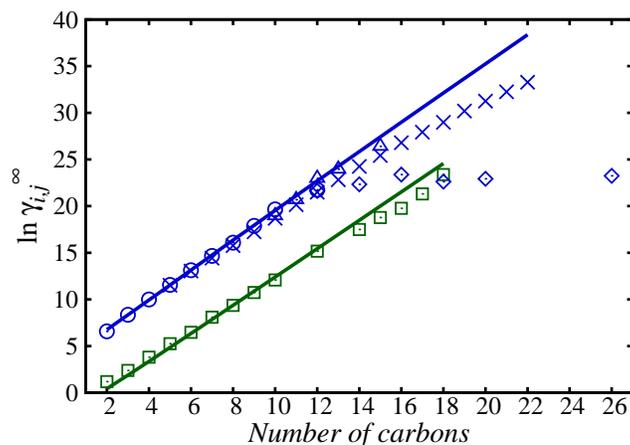
Figure 12: Infinite-dilution activity coefficients $\gamma_{i,j}^{\infty}$ of n-alkanes (top curve) and n-alkan-1-ols (bottom curve) in aqueous solution at $T = 298.15$ K and $p = 0.101$ MPa (except for ethane-butane, where the calculations are done at 5.00 MPa). The circles[126], triangles[102], diamonds[103], and squares[126] represent the experimental data, the crosses[12] the MD simulation data, and the continuous curves the predictions with the SAFT-γ Mie approach.
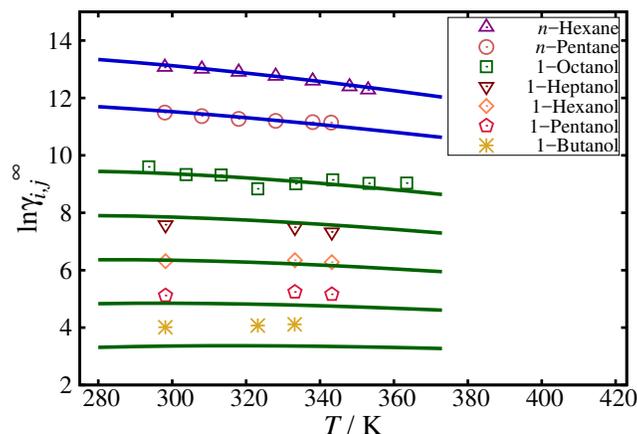

Figure 13: Infinite-dilution activity coefficients $\gamma_{i,j}^{\infty}$ of n-alkanes (n-pentane and n-hexane)[101] and n-alkan-1-ols (n-butan-1-ol[127], n-heptan-1-ol[127], n-octan-1-ol[128]) in aqueous solution at $p = 0.101$ MPa (except for n-pentane and n-hexane, where the calculations were carried out at 7.00 MPa). The symbols represent the experimental data and the continuous curves the predictions with the SAFT-γ Mie approach.

the relevant solute i), although it is useful to note that the water-rich phase is essentially pure water given the very small amount of octanol present and hence the $\gamma_{i,j}^{\infty}$ of a solute in pure water provides a direct link to the $K_{i,\text{OW}}$ calculation. Having acknowledged this, most solutes behave in a more non-ideal manner in the water-rich phase[6,129] (suggesting that the variation of $\gamma_{i,\text{WR}}$ is much greater than that of $\gamma_{i,\text{OR}}$) and we find that an accurate prediction of $\gamma_{i,\text{WR}}$ plays a vital role in the reliability of the description of $K_{i,\text{OW}}$. Knowledge of $K_{i,\text{OW}}$ is very useful in product and process design applications as it is often used to indicate the lipophilicity of compounds.[8] A key example is the application of $K_{i,\text{OW}}$ in predicting the pharmacokinetic properties and toxicity of organic chemicals, especially drug molecules.[10,130] The prediction of $K_{i,\text{OW}}$ using thermodynamic approaches is also an active subject of research.[131–135] The predictions of $K_{i,\text{OW}}$ for various n-alkanes and n-alkan-1-ols using the SAFT-γ Mie approach are presented in Fig. 14, where excellent agreement with the experimental data[136,137] is found. We emphasize that within our approach all the thermodynamic variables involved in the calculation of $K_{i,\text{OW}}$ (*cf.* Eq. (12)), including the molar volumes of the phases (*cf.* Fig. 8b) are determined with the SAFT-γ Mie approach; numerous alternative thermodynamic approaches[131–133,135] employ the experimental saturated volumes for the $K_{i,\text{OW}}$ prediction instead.
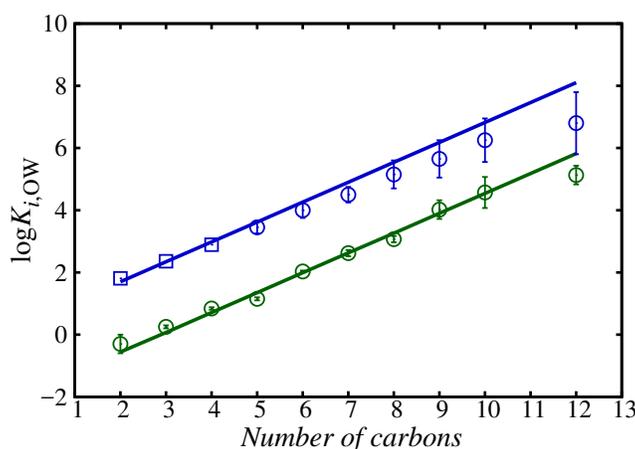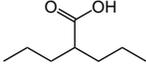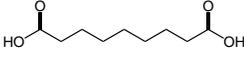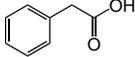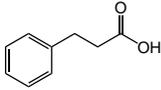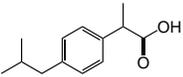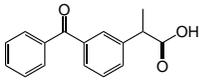

Figure 14: Octanol-water partition coefficients $K_{i,\text{OW}}$ of n-alkanes (top curve) and n-alkan-1-ols (bottom curve) at $T = 298.15$ K and $p = 0.101$ MPa. The symbols (squares[136] and circles[137]) represent the experimental data, the error bars the corresponding uncertainty,[137] and the continuous curves the predictions with the SAFT-γ Mie approach.

Prediction of partitioning and solubility of pharmaceuticals

The prediction the thermodynamic properties and fluid-phase behaviour of complex multifunctional molecules is often the primary goal of group-contribution approaches and constitutes a stringent test of the validity of the models. We investigate the prediction of $K_{i,\mathrm{OW}}$ for several multifunctional compounds, as well as the solubility of a number of active pharmaceutical ingredients (APIs) in different solvents. These properties are of particular relevance in the development of new drugs and the corresponding manufacturing processes. The APIs studied range from valproic acid (an anti-epileptic drug), azelaic acid (a topical anti-inflammatory treatment), phenylalkanoic acids (drug precursor) to ibuprofen and ketoprofen (nonsteroidal anti-inflammatory drugs). The octanol-water partition coefficients are calculated using Eq. (12) and are reported in Table 4. The relevant SAFT-γ Mie parameters are provided in the Appendix. Despite the presence of a larger number of functional groups present in these API compunds compared to the molecules considered in previous sections, it is apparent that the predicted values of $K_{i,\mathrm{OW}}$ agree well with the experimental data[137–139] . The level of agreement of our predictions is particularly encouraging considering that although experimental uncertainties of some of the APIs were not reported in the literature, an error of ± 0.5 log units is generally observed in different the determination of $K_{i,\mathrm{OW}}$ from different experiments[140] . In the case of ibuprofen, for instance, the reported experimental values for $K_{i,\mathrm{OW}}$ vary from 2.48[141] to 3.97.[142] Our predictions of the partition coefficient suggest that an appropriate balance of the various energetic contributions has been achieved with the current set of group interaction parameters, especially in terms of the like and unlike interactions and of the balance between the dispersion and hydrogen-bonding interactions.

The solubility of solid organic compounds is also a key physicochemical property in many industrial applications. In the pharmaceutical industry, the solubility is of particular importance because APIs are commonly recovered/purified as pure solids by crystallization

Table 4: Octanol-water partition coefficients ($K_{i,\text{OW}}$) of active pharmaceutical ingredients at $T = 298.15$ K and $p = 0.101$ MPa. The deviation is defined as $|\log K_{i,\text{OW}}^{\text{exp}} - \log K_{i,\text{OW}}^{\text{calc}}|$.

API	Structure	Exp. $\log K_{i,\text{OW}}$	Uncertainty	SAFT-γ Mie $\log K_{i,\text{OW}}$	Deviation
Valproic acid[137]		2.75	0.25	3.04	0.29
Azelaic acid[138]		1.57	n/a	1.69	0.12
Phenylacetic acid[137]		1.41	0.15	1.12	0.29
Hydrocinnamic acid[137]		1.84	0.15	1.77	0.07
Ibuprofen[139]		3.50	n/a	3.32	0.18
Ketoprofen[139]		3.12	n/a	3.06	0.06

and, as a result, the solubility of a pure crystalline API x_{API} in a given solvent as a function of temperature plays a key role in the development of downstream processes. A detailed knowledge of solubility is also of importance in the determination of the bioavailability of the compound. The solubility can be calculated by considering the solid-liquid equilibrium (SLE) between the pure crystalline solid and the liquid phase containing the solvent and the API using the following expression:[143]

$$\ln x_{\text{API}}(T, p) = \frac{\Delta H_m^{\text{fus}}}{R}\left(\frac{1}{T_m} - \frac{1}{T}\right) - \ln \gamma_{\text{API}}(T, p, \mathbf{x}), \tag{14}$$

where ΔH_m^{fus} and T_m are the molar enthalpy of fusion and the melting temperature of the solute, respectively, and where the difference Δc_p between the heat capacities of the solid and the liquid form of the solute has been neglected. The values for ΔH_m^{fus} and T_m are taken from experimental data[144–146]. The SAFT-γ Mie EoS is employed in the calculation of the activity coefficient of the API in the liquid phase $\gamma_{\text{API}}(T, p, \mathbf{x})$ at a pressure of $p = 0.101$

MPa. The calculation of the solubility through this route is predictive, as neither SLE data nor data relating to the APIs have been used to obtain the interaction parameters of our SAFT-γ Mie model. Information about the melting point and enthalpy of fusion of is of course still required for each API.

The solubilities of azelaic acid, ibuprofen, and ketoprofen in different solvents as predicted using the SAFT-γ Mie approach are compared with the corresponding experimental data in Fig. 15. Whereas azelaic acid is a relatively simple molecule, comprising only of a linear alkyl chain and two carboxylic acid groups, ibuprofen and ketoprofen are multifunctional compounds with at least one aromatic ring, alkyl chains and carboxylic acid groups. In Fig. 15a), the prediction of the solubility of azelaic acid in water is compared with three sets of experimental data[147–149] . It is apparent that there are inconsistencies between the sets of experimental data. One could use the theoretical SAFT-γ Mie approach to discriminate between the different data sets. Our predictions find remarkable agreement with the data of Chen *et al.*[148] . These authors attribute the deviation between the different sets of experimental data to the different measurement techniques used in each of the studies and to standard experimental error. Though it is worth mentioning that the differences observed appear larger than might be expected from the use of different techniques. Solubility predictions of ibuprofen and ketoprofen in water and *n*-butan-1-ol are shown Figs. 15b) and 15c), respectively. As can be been the SAFT-γ Mie predictions provide a very good representation of the experimental data,[150–156] accurately capturing the markedly different ranges of solubility in the different solvents. The solubility of these APIs in water is four orders of magnitude smaller than that in *n*-butan-1-ol and our GC approach captures this behaviour. We note that there is also some inconsistency in the aqueous solubility data,[150–152,157–159] possibly due to their very small magnitude, which makes the experimental determination challenging. In addition, a small deviation from the experimental measurements in the prediction of the solubility of these ionizable APIs[160,161] in aqueous solution is also likely due to the fact that our approach does not at this point account for the ionization.

In addition to the solubilities in alcohol and water, the predicted solubilities of ibuprofen and ketoprofen in acetone are presented in Fig. 15d). The excellent agreement between the SAFT-γ Mie predictions and the experimental data is found, demonstrating the applicability of the method to predict the solubility of a selection of APIs in different solvents.

Pharmaceutical compounds are composed of several functional groups, which often favour intramolecular association. This type of interaction gives rise to conformational changes within the molecules, which often also impacts the solubility of the molecules.[162] The predicted $K_{i,\mathrm{OW}}$ and solubilities of a number of compounds presented in this study have been obtained using a novel approach to treat intramolecular hydrogen bonding effectively within the SAFT-γ Mie framework. In brief, we "mimic" the intramolecular hydrogen bonds (IMHB) by switching off sites, both the electron donating (e site) and the electron accepting (H site), that are involved in the IMHB, thus preventing them to interact with other molecules. Details of this effective approach and the impact on the thermodynamic properties of the system is the subject of future work.

Conclusions

The description of the fluid-phase equilibria of aqueous mixtures of alkanes and alkanols using theoretically sound models, especially in a group-contribution framework, remains a challenge due to the non-ideality and complexity of interactions that the systems exhibit. For design purposes, a thermodynamic model should ideally be able to predict quantitatively with a single set of parameters the phase behaviour of the systems of interest over a range of thermodynamic conditions and the entire composition range, including the infinite-dilution region. In our work, the SAFT-γ Mie group-contribution approach is used to predict vapour-liquid and liquid-liquid equilibria of mixtures containing n-alkanes, n-alkan-1-ols, and water using a single set of group parameters. New interaction parameters for the description of the family of n-alkan-1-ols by means of a CH_2OH functional group are obtained, and revised
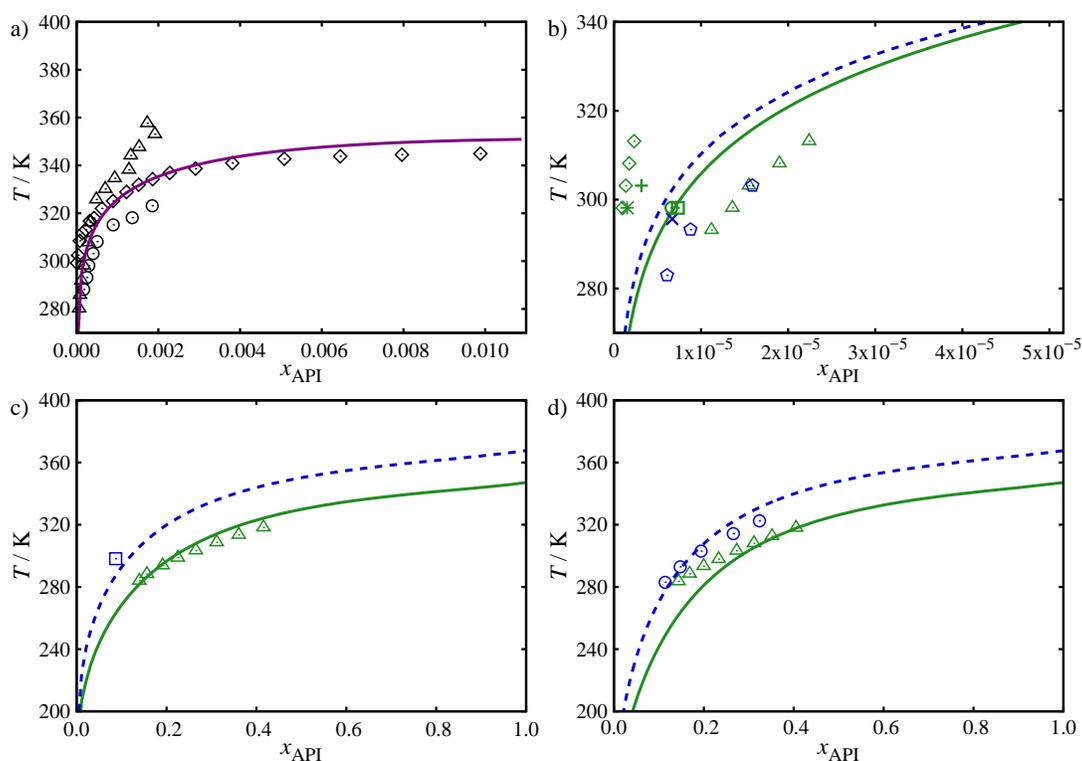
Figure 15: Prediction of the solid-liquid equilibria in binary systems of APIs in different solvents as a function of temperature at ambient pressure. The symbols represent the experimental data and the continuous and dashed curves the description with the SAFT-γ Mie approach: (a) azelaic acid (triangles[147], diamonds[148], and circles[149]) in water; (b) ibuprofen (diamonds[157], star[158], plus[159], circle[150], squares[151], triangles[152], and continuous curve) and ketoprofen (cross[153], pentagons[154], and dashed curve) in water; (c) ibuprofen (triangles[155] and continuous curve) and ketoprofen (squares[156] and dashed curve) in n-butan-1-ol; (d) ibuprofen (triangles[155] and continuous curve) and ketoprofen (circles[154] and dashed curve) in acetone.

unlike parameters between the alkyl methyl CH_3 and methylene CH_2 groups and the H_2O group for water are developed. The predictive capability of the models is confirmed for the fluid phase equilibria of n-alkane + n-alkan-1-ol and n-alkane + water mixtures by stringent comparison with experimental data. The solubilities of n-alkanes in the water-rich phase at conditions of three-phase coexistence are predicted for long n-alkanes in good agreement with experimental and simulation data for concentrations as low as 10^{-14} in mole fraction of the alkane for the larger chains (C_{22}). This interesting result highlights the difficulty in the experimental determination of these extremely low concentrations. New model parameters characterizing the CH_2OH - H_2O unlike interaction have also been determined and used to predict the phase behaviour of a number of alkanol + water binary mixtures that are not used in the development of the model. Our SAFT-γ Mie models provide good predictions for several aqueous mixtures of n-alkanes and n-alkan-1-ols over a wide range of thermodynamic conditions, including VLE, LLE, and VLLE. The robustness of the group parameters is further demonstrated in the prediction of several infinite-dilution properties. The Henry's law constants, solvation Gibbs free energies, and infinite-dilution activity coefficients of n-alkanes and n-alkan-1-ols in water are predicted and found to be in good agreement with the experimental data. The octanol-water partition coefficients ($K_{i,OW}$) of n-alkanes and n-1-alkanols are also obtained as an example of an infinite-dilution property of a ternary mixture. Obtaining an accurate prediction of $K_{i,OW}$ of these chemical families is very much reliant on an accurate description of the alkane solubility in the aqueous phase; this in turn relates to the group interaction parameters between the alkyl groups and water. In addition, we present the first example of the predictive capability of the SAFT-γ Mie approach for the solvation properties of more complex molecules such as active pharmaceutical ingredients (APIs). The predictions of $K_{i,OW}$ and solubilities of a number of representative APIs in water, n-butan-1-ol, and acetone are found to be in very good agreement with the corresponding experimental data. The SAFT-γ Mie approach accurately captures the marked differences in the solubilities of the API in different solvents. This is particularly gratifying considering

that no data related to the APIs are used to obtain the model parameters. The findings of our study validate the applicability and generic nature of the SAFT-γ Mie platform and confirm it as a promising tool in modelling complex molecules of relevance to pharmaceutical systems.

Data statement

Data underlying this article can be accessed on Zenodo at https://zenodo.org/record/xxx, and used under the Creative Commons Attribution licence.

Acknowledgements

Appendix

SAFT-γ Mie parameter tables

Table 5: Like group parameters for use in the SAFT-γ Mie group-contribution approach: ν_k^* is the number of segments constituting group k, S_k the shape factor, λ_{kk}^r the repulsive exponent, λ_{kk}^a the attractive exponent, σ_{kk} the segment diameter of group k, and ε_{kk} the dispersion energy of the Mie potential characterising the interaction of two k groups; $N_{\text{ST},k}$ represents the number of association site types on group k, with $n_{k,\text{H}}$, n_{k,e_1}, and n_{k,e_2} denoting the number of association sites of type H, e_1, and e_2 respectively.

k	Group k	ν_k^*	S_k	λ_{kk}^r	λ_{kk}^a	$\sigma_{kk}/\text{Å}$	$(\varepsilon_{kk}/k_\text{B})/\text{K}$	$N_{\text{ST},k}$	$n_{k,\text{H}}$	n_{k,e_1}	n_{k,e_2}	Ref.
1	CH_3	1	0.57255	15.050	6.0000	4.0773	256.77	–	–	–	–	71
2	CH_2	1	0.22932	19.871	6.0000	4.8801	473.39	–	–	–	–	71
3	CH	1	0.07210	8.0000	6.0000	5.2950	95.621	–	–	–	–	73
4	C	1	0.04072	8.0000	6.0000	5.6571	50.020	–	–	–	–	73
5	aCH	1	0.32184	14.756	6.0000	4.0578	371.53	1	0	1	0	73
6	aCCH	1	0.20650	8.0000	6.0000	4.3128	61.325	1	0	1	0	73
7	$aCCH_2$	1	0.20859	8.5433	6.0000	5.2648	591.56	1	0	1	0	73
8	$aCCH_3$	1	0.31655	23.627	6.0000	5.4874	651.41	1	0	1	0	163
9	COOH	1	0.55593	8.0000	6.0000	4.3331	405.78	3	1	2	2	164
10	CH_3COCH_3	3	0.72135	17.433	6.0000	3.5981	286.02	3	1	1	1	164
11	H_2O	1	1.0000	17.020	6.0000	3.0063	266.68	2	2	2	0	76
12	CH_2OH	2	0.58538	22.699	6.0000	3.4054	407.22	2	1	2	0	this work
13	aCCOaC	3	0.18086	9.8317	6.0000	4.0670	656.71	–	–	–	–	165

Table 6: Group dispersion interaction energies ε_{kl} and repulsive exponent $\lambda^{\mathrm{r}}_{kl}$ for use in the SAFT-γ Mie group-contribution approach. The unlike segment diameter σ_{kl} is obtained from the arithmetic combining rule and all unlike attractive exponents $\lambda^{\mathrm{a}}_{kl} = 6.0000$; these are not shown in the table. CR indicates that the unlike repulsive exponent $\lambda^{\mathrm{r}}_{kl}$ is obtained from a combining rule[71] and * indicates parameters obtained in this work.

k	l	Group k	Group l	$(\varepsilon_{kl}/k_{\mathrm{B}})/\mathrm{K}$	$\lambda^{\mathrm{r}}_{kl}$	Ref.	k	l	Group k	Group l	$(\varepsilon_{kl}/k_{\mathrm{B}})/\mathrm{K}$	$\lambda^{\mathrm{r}}_{kl}$	Ref.
1	1	CH_3	CH_3	256.77	15.050	71	4	12	C	CH_2OH	0.00000	CR	165
1	2	CH_3	CH_2	350.77	CR	71	5	5	aCH	aCH	371.53	14.756	163
1	3	CH_3	CH	387.48	CR	73	5	6	aCH	aCCH	429.16	CR	73
1	4	CH_3	C	339.91	CR	73	5	7	aCH	$aCCH_2$	416.69	CR	73
1	5	CH_3	aCH	305.81	CR	73	5	8	aCH	$aCCH_3$	471.23	CR	163
1	6	CH_3	aCCH	455.85	CR	73	5	9	aCH	COOH	331.61	9.0687	73
1	7	CH_3	$aCCH_2$	396.91	CR	73	5	10	aCH	CH_3COCH_3	333.11	CR	73
1	8	CH_3	$aCCH_3$	358.58	CR	163	5	11	aCH	H_2O	357.78	38.640	73
1	9	CH_3	COOH	255.99	CR	164	5	12	aCH	CH_2OH	386.05	CR	165
1	10	CH_3	CH_3COCH_3	233.48	14.449	164	6	6	aCCH	aCCH	61.325	8.0000	73
1	11	CH_3	H_2O	358.18	100.00	*	6	7	aCCH	$aCCH_2$	462.04	CR	73
1	12	CH_3	CH_2OH	333.20	CR	*	6	9	aCCH	COOH	599.28	CR	73
2	2	CH_2	CH_2	473.39	19.871	163	6	10	aCCH	CH_3COCH_3	459.22	CR	73
2	3	CH_2	CH	506.21	CR	73	6	11	aCCH	H_2O	314.03	CR	165
2	4	CH_2	C	300.07	CR	73	6	12	aCCH	CH_2OH	436.14	CR	165
2	5	CH_2	aCH	415.64	CR	73	7	7	$aCCH_2$	$aCCH_2$	591.56	8.5433	73
2	6	CH_2	aCCH	345.80	CR	73	7	9	$aCCH_2$	COOH	473.66	CR	73
2	7	CH_2	$aCCH_2$	454.16	CR	73	7	10	$aCCH_2$	CH_3COCH_3	394.83	CR	73
2	8	CH_2	$aCCH_3$	569.18	CR	163	7	11	$aCCH_2$	H_2O	329.03	CR	165
2	9	CH_2	COOH	413.74	CR	164	7	12	$aCCH_2$	CH_2OH	434.37	CR	165
2	10	CH_2	CH_3COCH_3	299.48	11.594	164	8	8	$aCCH_3$	$aCCH_3$	651.41	23.627	163
2	11	CH_2	H_2O	423.63	100.00	*	8	11	$aCCH_3$	H_2O	360.70	CR	165
2	12	CH_2	CH_2OH	423.17	CR	*	8	12	$aCCH_3$	CH_2OH	486.62	CR	165
3	3	CH	CH	95.621	8.0000	73	9	9	COOH	COOH	405.78	8.0000	164
3	4	CH	C	2.0000	CR	165	9	10	COOH	CH_3COCH_3	393.71	CR	73
3	5	CH	aCH	441.43	CR	73	9	11	COOH	H_2O	289.76	CR	165
3	6	CH	aCCH	67.510	CR	73	9	12	COOH	CH_2OH	656.80	CR	165
3	7	CH	$aCCH_2$	65.410	CR	73	10	10	CH_3COCH_3	CH_3COCH_3	286.02	17.433	164
3	9	CH	COOH	504.99	CR	73	10	11	CH_3COCH_3	H_2O	287.26	CR	164
3	10	CH	CH_3COCH_3	637.29	CR	164	11	11	H_2O	H_2O	226.68	17.020	76
3	11	CH	H_2O	275.75	CR	165	11	12	H_2O	CH_2OH	353.37	CR	*
3	12	CH	CH_2OH	329.22	CR	165	12	12	CH_2OH	CH_2OH	407.22	22.699	*
4	4	C	C	50.020	8.0000	73	13	13	aCCOaC	aCCOaC	656.71	9.8317	165
4	11	C	H_2O	420.82	CR	165	-	-	-	-	-	-	-

Table 7: Group association energies $\varepsilon^{\mathrm{HB}}_{kl,ab}$ and bonding volume parameters $K_{kl,ab}$ for use within the SAFT-γ Mie group-contribution approach.

k	l	Group k	Site a of group k	Group l	Site b of group l	$(\varepsilon^{\mathrm{HB}}_{kl,ab}/k_{\mathrm{B}})/\mathrm{K}$	$K_{kl,ab}/\text{Å}^3$
5	11	aCH	e_1	H_2O	H	563.56	339.61
6	11	aCCH	e_1	H_2O	H	563.56	339.61
7	11	$aCCH_2$	e_1	H_2O	H	563.56	339.61
8	11	$aCCH_3$	e_1	H_2O	H	563.56	339.61
9	9	COOH	H	COOH	H	6427.9	0.8062
9	11	COOH	e_1	H_2O	H	1451.8	280.89
9	11	COOH	e_2	H_2O	H	1252.6	150.98
9	11	COOH	H	H_2O	e_1	2567.7	270.09
9	12	COOH	e_1	CH_2OH	H	1015.5	21.827
9	12	COOH	e_2	CH_2OH	H	547.42	53.150
9	12	COOH	H	CH_2OH	e_1	524.04	14.017
10	10	CH_3COCH_3	e_1	CH_3COCH_3	H	980.20	2865.2
10	11	CH_3COCH_3	e_1	H_2O	H	1588.7	772.77
10	11	CH_3COCH_3	e_2	H_2O	H	417.24	1304.3
10	11	CH_3COCH_3	H	H_2O	e_1	1386.8	188.83
11	11	H_2O	e_1	H_2O	H	1985.4	101.69
11	12	H_2O	e_1	CH_2OH	H	621.68	425.00
11	12	H_2O	H	CH_2OH	e_1	2153.2	147.40
12	12	CH_2OH	e_1	CH_2OH	H	2097.9	62.309

Objective functions

The objective function function used in the parameter estimation for the CH_2OH group and its binary interaction parameters is given by:

$$
\begin{aligned}
\min_{\Omega} f_{\text{obj}} = {} & w_1 \sum_{i=1}^{N_C} \sum_{q=1}^{N_{p_{\text{vap}},i}} \left[\frac{p_{\text{vap},i}^{\exp}(T_q) - p_{\text{vap},i}^{\text{calc}}(T_q; \Omega)}{p_{\text{vap},i}^{\exp}(T_q)} \right]^2 \\
& + w_2 \sum_{i=1}^{N_C} \sum_{q=1}^{N_{\rho_{\text{sat}},i}} \left[\frac{\rho_{\text{sat},i}^{\exp}(T_q) - \rho_{\text{sat},i}^{\text{calc}}(T_q; \Omega)}{\rho_{\text{sat},i}^{\exp}(T_q)} \right]^2 \\
& + w_3 \sum_{q=1}^{N_{hE}} \left[\frac{h^{E,\exp}(T_q, p_q, \mathbf{x}_q) - h^{E,\text{calc}}(T_q, p_q, \mathbf{x}_u; \Omega)}{h^{E,\exp}(T_q, p_q, \mathbf{x}_q)} \right]^2 \qquad (15)\\
& + w_4 \sum_{q=1}^{N_{x_{C14}}} \left[\frac{x_{C14}^{\exp}(T_q, p_q) - x_{C14}^{\text{calc}}(T_q, p_q; \Omega)}{x_{C14}^{\exp}(T_q, p_q)} \right]^2 \\
& + w_5 \sum_{q=1}^{N_{y_{C14}}} \left[\frac{y_{C14}^{\exp}(T_q, p_q) - y_{C14}^{\text{calc}}(T_q, p_q; \Omega)}{y_{C14}^{\exp}(T_q, p_q)} \right]^2 ,
\end{aligned}
$$

where the first two sums are over the N_C pure components i included in the estimation over the number $N_{p_{\text{vap}}}$ of experimental vapour pressure points ($N_{p_{\text{vap}}} = 214$) or $N_{\rho_{\text{sat}}}$ saturated liquid density points ($N_{\rho_{\text{sat}}} = 336$), while the third term sums over the number N_{hE} of experimental molar excess enthalpy points ($N_{hE} = 25$). The last two terms sum over LLE data points of the ethanol+tetradecane mixture: $N_{x_{C14}}$ ethanol-rich equilibrium mole fraction data points ($N_{x_{C14}} = 7$) and $N_{y_{C14}}$ alkane-rich equilibrium mole fraction data points ($N_{y_{C14}} = 9$). Ω denotes the vector of parameters to be estimated. The desired level of accuracy for each calculated (calc) property can be adjusted by means of weighting factors: w_1 for $N_{p_{\text{vap}}}$, w_2 for $N_{\rho_{\text{sat}}}$, w_3 for N_{hE}, w_4 for N_{x_j}, and w_5 for N_{y_j}. In this case, $w_1 = w_2 = 5$, $w_3 = 1$, $w_4 = w_5 = 10$ are employed. The estimations are performed using the commercial software package gPROMS.[79]

The objective function function used in the parameter estimation for the CH_3 - H_2O and

CH_2 - H_2O binary interaction parameters is given by:

$$\min_{\Omega} f_{\text{obj}} = w_1 \sum_{C_5, C_8} \sum_{q=1}^{N_{x_i}} \left[\frac{x_i^{\exp}(T_q) - x_i^{\text{calc}}(T_q; \Omega)}{x_i^{\exp}(T_q)} \right]^2$$
$$+ w_2 \sum_{C_5}^{C_{10}} \left[\frac{x_i^{\exp}(298) - x_i^{\text{calc}}(298; \Omega)}{x_i^{\exp}(298)} \right]^2 \tag{16}$$
$$+ w_3 \sum_{C_5}^{C_{10}} \sum_{q=1}^{N_{y_j}} \left[\frac{y_j^{\exp}(T_q) - y_j^{\text{calc}}(T_q; \Omega)}{y_j^{\exp}(T_q)} \right]^2,$$

in which sums over the square of the relative residuals between the experimental (exp) and calculated (calc) equilibrium mole fraction of the alkane in the water-rich (liquid) phase $x_i(T)$ and $x_i(298)$ (water-rich phase compositions at $T = 298.00$ K) and alkane-rich (liquid) phase $y_j(T)$ of a given mixture at specified values of temperature (including $y_j(298)$) over all experimental points, where i denotes n-alkane and j denotes water. $N_{x_i} = 15$, $N_{y_j} = 21$, and weighting factors $w_1 = 5$, $w_2 = 10$ and $w_3 = 1$ are used. The minimisation is performed using the commercial software package gPROMS.[79] A multistart gradient-based optimisation algorithm (HELD algorithm)[80,81] is used as input to local optimisations.

The objective function used in the parameter estimation for the CH_2OH - H_2O binary interaction parameters is given by:

$$\min_{\Omega} f_{\text{obj}} = w_1 \sum_{q=1}^{N_{x_i}} \left[\frac{x_i^{\exp}(T_q, p_q) - x_i^{\text{calc}}(T_q, p_q; \Omega)}{x_i^{\exp}(T_q, p_q)} \right]^2$$
$$+ w_2 \sum_{z=1}^{N_{y_i}} \left[\frac{y_i^{\exp}(T_q, p_q) - y_i^{\text{calc}}(T_q, p_q; \Omega)}{y_i^{\exp}(T_q, p_q)} \right]^2, \tag{17}$$

where the first sum is over the square of the relative residuals between the experimental (exp) and calculated (calc) values of the equilibrium mole fractions of the water-rich (liquid) phase $x_i(T)$ and the second over the equilibrium mole fractions of the octanol-rich (liquid) phase $y_i(T)$ and where i denotes 1-octanol. Here $N_{x_i} = 8$, $N_{y_i} = 8$, $w_1 = 1$ and $w_2 = 20$.

References

(1) Noel T. Southall, K. A. D.; Haymet, A. D. J. A View of the Hydrophobic Effect. *J. Phys. Chem. B* **2002**, *106*, 521–533.

(2) Dill, K. A.; Truskett, T. M.; Vlachy, V.; Hribar-Lee, B. Modeling Water, the Hydrophobic Effect, and Ion Solvation. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 173–199.

(3) Chandler, D. Interfaces and the driving force of hydrophobic assembly. *Nature* **2005**, *437*, 640–647.

(4) Tanford, C. The hydrophobic effect and the organization of living matter. *Science* **1978**, *200*, 1012–1018.

(5) Mączyński, A.; Wiśniewska-Gocłowska, B.; Góral, M. Recommended Liquid-Liquid Equilibrium Data. Part 1. Binary alkane-water systems. *J. Phys. Chem. Ref. Data* **2004**, *33*, 549–577.

(6) Tsonopoulos, C. Thermodynamic analysis of the mutual solubilities of normal alkanes and water. *Fluid Phase Equilib.* **1999**, *156*, 21–33.

(7) Voutsas, E. C.; Boulougouris, G. C.; Economou, I. G.; Tassios, D. P. Water/Hydrocarbon Phase Equilibria Using the Thermodynamic Perturbation Theory. *Ind. Eng. Chem. Res.* **2000**, *39*, 797–804.

(8) Leo, A.; Hansch, C.; Elkins, D. Partition coefficients and their uses. *Chem. Rev.* **1971**, *71*, 525–616.

(9) Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235–249.

(10) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings1. *Adv. Drug Delivery Rev.* **2001**, *46*, 3–26.

(11) Jorgensen, W. L.; Gao, J.; Ravimohan, C. Monte Carlo simulations of alkanes in water: hydration numbers and the hydrophobic effect. *J. Phys. Chem.* **1985**, *89*, 3470–3473.

(12) Ferguson, A. L.; Debenedetti, P. G.; Panagiotopoulos, A. Z. Solubility and Molecular Conformations of n-Alkane Chains in Water. *J. Phys. Chem. B* **2009**, *113*, 6405–6414.

(13) Ashbaugh, H. S.; Liu, L.; Surampudi, L. N. Optimization of linear and branched alkane interactions with water to simulate hydrophobic hydration. *J. Chem. Phys.* **2011**, *135*, 054510.

(14) Voutsas, E. C.; Tassios, D. P. Prediction of Infinite-Dilution Activity Coefficients in Binary Mixtures with UNIFAC. A Critical Evaluation. *Ind. Eng. Chem. Res.* **1996**, *35*, 1438–1445.

(15) Zhang, S.; Hiaki, T.; Hongo, M.; Kojima, K. Prediction of infinite dilution activity coefficients in aqueous solutions by group contribution models. A critical evaluation . *Fluid Phase Equilib.* **1998**, *144*, 97–112.

(16) Oracz, P.; Góral, M. Application of the Unified Functional Activity Coefficient (UNIFAC) and Analytical Solution of Groups (ASOG) for the Calculation of Mutual Solubilities in Water Systems of Alkanes, Arenes, and Alkanols. *J. Chem. Eng. Data* **2011**, *56*, 4853–4861.

(17) Satyro, M. A.; Shaw, J. M.; Yarranton, H. W. A practical method for the estimation of oil and water mutual solubilities. *Fluid Phase Equilib.* **2013**, *355*, 12–25.

(18) Klamt, A. Prediction of the mutual solubilities of hydrocarbons and water with COSMO-RS. *Fluid Phase Equilib.* **2003**, *206*, 223–235.

(19) Folas, G. K.; Gabrielsen, J.; Michelsen, M. L.; Stenby, E. H.; Kontogeorgis, G. M. Application of the Cubic-Plus-Association (CPA) Equation of State to Cross-Associating Systems. *Ind. Eng. Chem. Res.* **2005**, *44*, 3823–3833.

(20) Vega, L.; Llovell, F. Review and new insights into the application of molecular-based equations of state to water and aqueous solutions. *Fluid Phase Equilib.* **2016**, *416*, 150 – 173.

(21) Peng, D.-Y.; Robinson, D. B. A New Two-Constant Equation of State. *Ind. Eng. Chem. Fundam.* **1976**, *15*, 59–64.

(22) Soave, G. S. Estimation of the critical constants of heavy hydrocarbons for their treatment by the Soave-Redlich-Kwong equation of state. *Fluid Phase Equilib.* **1998**, *143*, 29–39.

(23) Tsonopoulos, C.; Heidman, J. High-pressure vapor-liquid equilibria with cubic equations of state. *Fluid Phase Equilib.* **1986**, *29*, 39–414.

(24) Kabadi, V. N.; Danner, R. P. A modified Soave-Redlich-Kwong equation of state for water-hydrocarbon phase equilibria. *Ind. Eng. Chem. Process Des. Dev.* **1985**, *24*, 537–541.

(25) Michel, S.; Hooper, H. H.; Prausnitz, J. M. Mutual solubilities of water and hydrocarbons from an equation of state. Need for an unconventional mixing rule . *Fluid Phase Equilib.* **1989**, *45*, 173–189.

(26) Chapman, W.; Gubbins, K.; Jackson, G.; Radosz, M. SAFT: Equation-of-state solution model for associating fluids. *Fluid Phase Equilib.* **1989**, *52*, 31–38.

(27) Chapman, W. G.; Gubbins, K. E.; Jackson, G.; Radosz, M. New reference equation of state for associating liquids. *Ind. Eng. Chem. Res.* **1990**, *29*, 1709–1721.

(28) Kontogeorgis, G. M.; Voutsas, E. C.; Yakoumis, I. V.; Tassios, D. P. An Equation of State for Associating Fluids. *Ind. Eng. Chem. Res.* **1996**, *35*, 4310–4318.

(29) Wertheim, M. Fluids with highly directional attractive forces. I. Statistical thermodynamics. *J. Stat. Phys.* **1984**, *35*, 19–34.

(30) Wertheim, M. Fluids with highly directional attractive forces. II. Thermodynamic perturbation theory and integral equations. *J. Stat. Phys.* **1984**, *35*, 35–47.

(31) Wertheim, M. Fluids with highly directional attractive forces. III. Multiple attraction sites. *J. Stat. Phys.* **1986**, *42*, 459–476.

(32) Wertheim, M. Fluids with highly directional attractive forces. IV. Equilibrium polymerization. *J. Stat. Phys.* **1986**, *42*, 477–492.

(33) Papaioannou, V.; Adjiman, C. S.; Jackson, G.; Galindo, A. *Process Systems Engineering*; Wiley-VCH Verlag GmbH & Co. KGaA, 2011; pp 135–172.

(34) Pierotti, G. J.; Deal, C. H.; Derr, E. L. Activity Coefficients and Molecular Structure. *Ind. Eng. Chem.* **1959**, *51*, 95–102.

(35) Zhang, S.; Hiaki, T.; Kojima, K. Prediction of infinite dilution activity coefficients for systems including water based on the group contribution model with mixture-type groups: 1. Alkane-H_2O and alkanol-H_2O mixtures. *Fluid Phase Equilib.* **1998**, *149*, 27–40.

(36) Mengarelli, A. C.; Brignole, E. A.; Bottini, S. B. Activity coefficients of associating mixtures by group contribution. *Fluid Phase Equilib.* **1999**, *163*, 195–207.

(37) Soares, R. d. P.; Gerber, R. P. Functional-Segment Activity Coefficient Model. 1. Model Formulation. *Ind. Eng. Chem. Res.* **2013**, *52*, 11159–11171.

(38) Soares, R. d. P.; Gerber, R. P.; Possani, L. F. K.; Staudt, P. B. Functional-Segment Activity Coefficient Model. 2. Associating Mixtures. *Ind. Eng. Chem. Res.* **2013**, *52*, 11172–11181.

(39) Possani, L.; Oes, R. S.; Staudt, P.; de P. Soares, R. Mutual solubilities of hydrocarbon-water systems with F-SAC. *Fluid Phase Equilib.* **2014**, *384*, 122–133.

(40) Müller, E. A.; Jackson, G. Force-Field Parameters from the SAFT-γ Equation of State for Use in Coarse-Grained Molecular Simulations. *Annu. Rev. Chem. Biomol. Eng.* **2014**, *5*, 405–427.

(41) Hajiw, M.; Chapoy, A.; Coquelet, C. Hydrocarbons-water phase equilibria using the CPA equation of state with a group contribution method. *Can. J. Chem. Eng.* **2014**, 432–442.

(42) Gros, H.; Bottini, S.; Brignole, E. High pressure phase equilibrium modeling of mixtures containing associating compounds and gases. *Fluid Phase Equilib.* **1997**, *139*, 75–87.

(43) Soria, T. M.; Sánchez, F. A.; Pereda, S.; Bottini, S. B. Modeling alcohol+water+hydrocarbon mixtures with the group contribution with association equation of state GCA-EoS. *Fluid Phase Equilib.* **2010**, *296*, 116–124.

(44) Soria, T. M.; Andreatta, A. E.; Pereda, S.; Bottini, S. B. Thermodynamic modeling of phase equilibria in biorefineries. *Fluid Phase Equilib.* **2011**, *302*, 1–9.

(45) Nguyen-Huynh, D.; de Hemptinne, J.-C.; Lugo, R.; Passarello, J.-P.; Tobaly, P. Modeling Liquid-Liquid and Liquid-Vapor Equilibria of Binary Systems Containing Water with an Alkane, an Aromatic Hydrocarbon, an Alcohol or a Gas (Methane, Ethane, CO_2 or H_2S), Using Group Contribution Polar Perturbed-Chain Statistical Associating Fluid Theory. *Ind. Eng. Chem. Res.* **2011**, *50*, 7467–7483.

(46) Nguyen, T.-B.; de Hemptinne, J.-C.; Creton, B.; Kontogeorgis, G. M. GC-PPC-SAFT Equation of State for VLE and LLE of Hydrocarbons and Oxygenated Compounds. Sensitivity Analysis. *Ind. Eng. Chem. Res.* **2013**, *52*, 7014–7029.

(47) Nguyen, T.-B.; de Hemptinne, J.-C.; Creton, B.; Kontogeorgis, G. M. Improving GC-PPC-SAFT equation of state for LLE of hydrocarbons and oxygenated compounds with water. *Fluid Phase Equilib.* **2014**, *372*, 113–125.

(48) Ahmed, S.; Ferrando, N.; de Hemptinne, J.-C.; Simonin, J.-P.; Bernard, O.; Baudouin, O. A new PC-SAFT model for pure water, water–hydrocarbons, and water–Oxygenates systems and subsequent modeling of VLE, VLLE, and LLE. *J. Chem. Eng. Data* **2016**, *61*, 4178–4190.

(49) Gil-Villegas, A.; Galindo, A.; Whitehead, P. J.; Mills, S. J.; Jackson, G.; Burgess, A. N. Statistical associating fluid theory for chain molecules with attractive potentials of variable range. *J. Chem. Phys.* **1997**, *106*, 4168–4186.

(50) Galindo, A.; Davies, L. A.; Gil-Villegas, A.; Jackson, G. The thermodynamics of mixtures and the corresponding mixing rules in the SAFT-VR approach for potentials of variable range. *Mol. Phys.* **1998**, *93*, 241–252.

(51) Lymperiadis, A.; Adjiman, C. S.; Galindo, A.; Jackson, G. A group contribution method for associating chain molecules based on the statistical associating fluid theory (SAFT-γ). *J. Chem. Phys.* **2007**, *127*, 234903.

(52) Lymperiadis, A.; Adjiman, C. S.; Jackson, G.; Galindo, A. A generalisation of the SAFT-γ group contribution method for groups comprising multiple spherical segments. *Fluid Phase Equilib.* **2008**, *274*, 85–104.

(53) Papaioannou, V.; Adjiman, C. S.; Jackson, G.; Galindo, A. Simultaneous prediction of vapour-liquid and liquid-liquid equilibria (VLE and LLE) of aqueous mixtures with the SAFT-γ group contribution approach. *Fluid Phase Equilib.* **2011**, *306*, 82–96.

(54) Chremos, A.; Forte, E.; Papaioannou, V.; Galindo, A.; Jackson, G.; Adjiman, C. S. Modelling the phase and chemical equilibria of aqueous solutions of alkanolamines and carbon dioxide using the SAFT-γ SW group contribution approach. *Fluid Phase Equilib.* **2015**, *407*, 280–297.

(55) Huang, S. H.; Radosz, M. Equation of state for small, large, polydisperse, and associating molecules: Extension to fluid mixtures. *Ind. Eng. Chem. Res.* **1991**, *30*, 1994–2005.

(56) Economou, I. G.; Tsonopoulos, C. Associating models and mixing rules in equations of state for water/hydrocarbon mixtures. *Chem. Eng. Sci.* **1997**, *52*, 511–525.

(57) Yakoumis, I. V.; Kontogeorgis, G. M.; Voutsas, E. C.; Hendriks, E. M.; Tassios, D. P. Prediction of Phase Equilibria in Binary Aqueous Systems Containing Alkanes, Cycloalkanes, and Alkenes with the Cubic-plus-Association Equation of State. *Ind. Eng. Chem. Res.* **1998**, *37*, 4175–4182.

(58) Galindo, A.; Whitehead, P. J.; Jackson, G. Predicting the High-Pressure Phase Equilibria of Water + *n*-Alkanes Using a Simplified SAFT Theory with Transferable Intermolecular Interaction Parameters. *J. Phys. Chem.* **1996**, *100*, 6781–6792.

(59) Patel, B. H.; Paricaud, P.; Galindo, A.; Maitland, G. C. Prediction of the Salting-Out Effect of Strong Electrolytes on Water + Alkane Solutions. *Ind. Eng. Chem. Res.* **2003**, *42*, 3809–3823.

(60) Vega, L. F.; Llovell, F.; Blas, F. J. Capturing the Solubility Minima of *n*-Alkanes in Water by Soft-SAFT. *J. Phys. Chem. B* **2009**, *113*, 7621–7630.

(61) Grenner, A.; Schmelzer, J.; von Solms, N.; Kontogeorgis, G. M. Comparison of Two Association Models (Elliott-Suresh-Donohue and Simplified PC-SAFT) for Complex Phase Equilibria of Hydrocarbon-Water and Amine-Containing Mixtures. *Ind. Eng. Chem. Res.* **2006**, *45*, 8170–8179.

(62) Emborsky, C. P.; Cox, K. R.; Chapman, W. G. Correlation and Prediction of Water Content in Alkanes Using a Molecular Theory. *Ind. Eng. Chem. Res.* **2011**, *50*, 7791–7799.

(63) Fouad, W. A.; Ballal, D.; Cox, K. R.; Chapman, W. G. Examining the Consistency of Water Content Data in Alkanes Using the Perturbed-Chain Form of the Statistical Associating Fluid Theory Equation of State. *J. Chem. Eng. Data* **2014**, *59*, 1016–1023.

(64) McCabe, C.; Galindo, A.; Cummings, P. T. Anomalies in the Solubility of Alkanes in Near-Critical Water. *J. Phys. Chem. B* **2003**, *107*, 12307–12314.

(65) Kraska, T.; Gubbins, K. E. Phase Equilibria Calculations with a Modified SAFT Equation of State. 2. Binary Mixtures of *n*-Alkanes, 1-Alkanols, and Water. *Ind. Eng. Chem. Res.* **1996**, *35*, 4738–4746.

(66) Karakatsani, E. K.; Kontogeorgis, G. M.; Economou, I. G. Evaluation of the Truncated Perturbed Chain-Polar Statistical Associating Fluid Theory for Complex Mixture Fluid Phase Equilib. *Ind. Eng. Chem. Res.* **2006**, *45*, 6063–6074.

(67) Liang, X.; Tsivintzelis, I.; Kontogeorgis, G. M. Modeling Water Containing Systems with the Simplified PC-SAFT and CPA Equations of State. *Ind. Eng. Chem. Res.* **2014**, *53*, 14493–14507.

(68) Al-Saifi, N. M.; Hamad, E. Z.; Englezos, P. Prediction of vapor-liquid equilibrium in water-alcohol-hydrocarbon systems with the dipolar perturbed-chain {SAFT} equation of state. *Fluid Phase Equilib.* **2008**, *271*, 82–93.

(69) de Villiers, A. J.; Schwarz, C. E.; Chobanov, K. G.; Burger, A. J. Application of sPC-SAFT-JC and sPC-SAFT-GV to Phase Equilibria Predictions of Alkane/Alcohol, Alcohol/Alcohol, and Water/Alcohol Binary Systems. *Ind. Eng. Chem. Res.* **2014**, *53*, 6065–6075.

(70) Kontogeorgis, G. M.; Michelsen, M. L.; Folas, G. K.; Derawi, S.; von Solms, N.; Stenby, E. H. Ten years with the CPA (cubic-plus-association) equation of state. Part 2. Cross-associating and multicomponent systems. *Ind. Eng. Chem. Res.* **2006**, *45*, 4869–4878.

(71) Papaioannou, V.; Lafitte, T.; Avendaño, C.; Adjiman, C. S.; Jackson, G.; Müller, E. A.; Galindo, A. Group contribution methodology based on the statistical associating fluid theory for heteronuclear molecules formed from Mie segments. *J. Chem. Phys.* **2014**, *140*, 054107.

(72) Lafitte, T.; Apostolakou, A.; Avendaño, C.; Galindo, A.; Adjiman, C. S.; Müller, E. A.; Jackson, G. Accurate statistical associating fluid theory for chain molecules formed from Mie segments. *J. Chem. Phys.* **2013**, *139*, 154504.

(73) Dufal, S.; Papaioannou, V.; Sadeqzadeh, M.; Pogiatzis, T.; Chremos, A.; Adjiman, C. S.; Jackson, G.; Galindo, A. Prediction of Thermodynamic Properties and Phase Behavior of Fluids and Mixtures with the SAFT-γ Mie Group-Contribution Equation of State. *J. Chem. Eng. Data* **2014**, *59*, 3272–3288.

(74) Mie, G. Zur kinetischen Theorie der einatomigen Körper. *Ann. Phys. (Berlin, Ger.)* **1903**, *316*, 657–697.

(75) Jackson, G.; Chapman, W. G.; Gubbins, K. E. Phase equilibria of associating fluids. *Mol. Phys.* **1988**, *65*, 1–31.

(76) Dufal, S.; Lafitte, T.; Haslam, A. J.; Galindo, A.; Clark, G. N.; Vega, C.; Jackson, G. The A in SAFT: Developing the contribution of association to the Helmholtz free energy within a Wertheim TPT1 treatment of generic Mie fluids. *Mol. Phys.* **2015**, *113*, 948–984.

(77) Sandler, S. I. *Chemical, Biochemical, and Engineering Thermodynamics, 4th Edition*; John Wiley & Sons, Inc, 2006.

(78) Michelsen, M. L.; Mollerup, J. M.; Breil, M. P. *Thermodynamic Models: Fundamental & Computational Aspects, 2nd Edition*; Tie-Line Publications, 2008.

(79) gPROMS v. 3.7.1; PSE Ltd.: London, 2013; http://www. psenterprise.com.

(80) Pereira, F. E.; Jackson, G.; Galindo, A.; Adjiman, C. S. A duality-based optimisation approach for the reliable solution of (P, T) phase equilibrium in volume-composition space. *Fluid Phase Equilib.* **2010**, *299*, 1–23.

(81) Pereira, F. E.; Jackson, G.; Galindo, A.; Adjiman, C. S. The {HELD} algorithm for multicomponent, multiphase equilibrium calculations with generic equations of state. *Comput. Chem. Eng.* **2012**, *36*, 99 – 118.

(82) Plyasunov, A. V.; Shock, E. L. Standard state Gibbs energies of hydration of hydrocarbons at elevated temperatures as evaluated from experimental phase equilibria studies. *Geochim. Cosmochim. Acta* **2000**, *64*, 2811–2833.

(83) Plyasunov, A. V.; Shock, E. L. Thermodynamic functions of hydration of hydrocarbons at 298.15K and 0.1 MPa. *Geochim. Cosmochim. Acta* **2000**, *64*, 439–468.

(84) Rowlinson, J. S.; Swinton, F. L. *Liquids and liquid mixtures*, 3rd ed.; Butterworths, 1982.

(85) Wu, H. S.; Sandler, S. I. Use of ab initio quantum mechanics calculations in group contribution methods. 1. Theory and the basis for group identifications. *Ind. Eng. Chem. Res.* **1991**, *30*, 881–889.

(86) Huang, S. H.; Radosz, M. Equation of state for small, large, polydisperse, and associating molecules. *Ind. Eng. Chem. Res.* **1990**, *29*, 2284–2294.

(87) Smith, B. D.; Srivastava, R. *Thermodynamic Data for Pure Compounds. Part B: Halogenated Compounds and Alcohols*; Elsevier Amsterdam, 1986.

(88) Hamam, S. E.; Kumaran, M.; Benson, G. C. Excess enthalpies of some binary mixtures: (a C_5-alkanol $+$ n-heptane) at 298.15 K. *J. Chem. Thermodyn.* **1984**, *16*, 1013–1017.

(89) Matsuda, H.; Ochi, K. Liquid-liquid equilibrium data for binary alcohol $+$ n-alkane (C_{10}-C_{16}) systems: methanol $+$ decane, ethanol $+$ tetradecane, and ethanol $+$ hexadecane. *Fluid Phase Equilib.* **2004**, *224*, 31–37.

(90) Máchová, I.; Linek, J.; Wichterle, I. Vapour-liquid equilibria in the heptane - 1-pentanol and heptane - 3-methyl-1-butanol systems at 75, 85 and 95 °C. *Fluid Phase Equilib.* **1988**, *41*, 257–267.

(91) Schmelzer, J.; Lieberwirth, I.; Krug, M.; Pfestorf, R. Vapour-liquid equilibria and heats of mixing in alkane-alcohol(1) systems. I. Vapour-liquid equilibria in 1-alcohol-undecane systems. *Fluid Phase Equilib.* **1983**, *11*, 187–200.

(92) Soo, C.-B.; Ahmar, E. E.; Coquelet, C.; Ramjugernath, D.; Richon, D. Vapor-liquid equilibrium measurements and modeling of the n-butane $+$ ethanol system from 323 to 423 K. *Fluid Phase Equilib.* **2009**, *286*, 79–87.

(93) Belabbaci, A.; Villamañan, R. M.; Negadi, L.; Martin, C. M.; Ait Kaci, A.; Villamañan, M. a. Vapor-Liquid Equilibria of Binary Mixtures Containing 1-Butanol and Hydrocarbons at 313.15 K. *J. Chem. Eng. Data* **2012**, *57*, 114–119.

(94) Lin, S. T.; Sandler, S. I. Multipole Corrections To Account for Structure and Proximity Effects in Group Contribution Methods:Octanol-Water Partition Coefficients. *J. Phys. Chem. A,* **2000**, *104*, 7099–7105.

(95) Maczynski, A.; Shaw, D. G.; Goral, M.; Wisniewska-Goclowska, B.; Skrzecz, A.; Maczynska, Z.; Owczarek, I.; Blazej, K.; Haulait-Pirson, M.-C.; Kapuku, F.; Hefter, G. T.; Szafranski, A. IUPAC-NIST Solubility Data Series. 81. Hydrocarbons with water and seawater–Revised and Updated Part 1. C_5 hydrocarbons with water. *J. Phys. Chem. Ref. Data* **2005**, *34*, 441–476.

(96) Shaw, D. G. IUPAC-NIST Solubility Data Series. 81. Hydrocarbons with water and seawater–Revised and Updated. Part 7. C_8H_{12} –C_8H_{18} hydrocarbons with water. *J. Phys. Chem. Ref. Data* **2005**, *34*, 2261–2298.

(97) Maczynski, A.; Shaw, D. G.; Goral, M.; Wisniewska-Goclowska, B.; Skrzecz, A.; Owczarek, I.; Blazej, K.; Haulait-Pirson, M.-C.; Hefter, G. T.; Kapuku, F.; Maczynska, Z.; Young, C. L. IUPAC-NIST Solubility Data Series. 81. Hydrocarbons with water and seawater–Revised and Updated. Part 4. C_6H_{14} hydrocarbons with water. *J. Phys. Chem. Ref. Data* **2005**, *34*, 709–753.

(98) Maczynski, A.; Shaw, D. G.; Goral, M.; Wisniewska-Goclowska, B.; Skrzecz, A.; Owczarek, I.; Blazej, K.; Haulait-Pirson, M.-C.; Hefter, G. T.; Kapuku, F.; Maczynska, Z.; Szafranski, A.; Young, C. L. IUPAC-NIST Solubility Data Series. 81. Hydrocarbons with water and seawater–Revised and Updated. Part 5. C_7 Hydrocarbons with water and heavy water. *J. Phys. Chem. Ref. Data* **2005**, *34*, 1399–1487.

(99) Shaw, D. G. IUPAC-NIST Solubility Data Series. 81. Hydrocarbons with water and seawater–Revised and Updated. Part 8. C_9 hydrocarbons with water. *J. Phys. Chem. Ref. Data* **2005**, *34*, 2299–2345.

(100) Shaw, D. G. IUPAC-NIST Solubility Data Series. 81. Hydrocarbons with water and seawater–Revised and Updated. Part 9. C_{10} hydrocarbons with water. *J. Phys. Chem. Ref. Data* **2006**, *35*, 93–151.

(101) Mokraoui, S.; Coquelet, C.; Valtz, A.; Hegel, P. E.; Richon, D. New solubility data of hydrocarbons in water and modeling concerning vapor-liquid-liquid binary systems. *Ind. Eng. Chem. Res.* **2007**, *46*, 9257–9262.

(102) Tolls, J.; van Dijk, J.; Verbruggen, E. J. M.; Hermens, J. L. M.; Loeprecht, B.; Schüürmann, G. Aqueous Solubility-Molecular Size Relationships: A Mechanistic Case Study Using $C_1$0- to $C_1$9-Alkanes. *J. Phys. Chem. A* **2002**, *106*, 2760–2765.

(103) Sutton, C.; Calder, J. A. Solubility of higher-molecular-weight normal-paraffins in distilled water and sea water. *Environ. Sci. Technol.* **1974**, *8*, 654–657.

(104) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1998**, *102*, 2569–2577.

(105) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269âĂŞ–6271.

(106) Reamer, H. H.; Sage, B. H.; Lacey, W. N. Phase Equilibria in Hydrocarbon Systems. *n*-Butane-Water System in the Two-Phase Region. *Ind. Eng. Chem.* **1952**, *44*, 609–615.

(107) Namiot, A. Y.; Skripka, V. G.; Lotter, G. Y. Phase Equilibria in Hydrocarbon-Water Systems at High Temperatures. *Deposited Doc. VINITI* **1976**, *1213*, 1–13.

(108) Sultanov, R. G.; Skripka, V. G. Solubility of Water in *n*-Alkanes at Elevated Temperatures and Pressures. *Deposited Doc. VINITI* **1972**, *4386*, 1–13.

(109) Sultanov, R. G.; Skripka, V. G. The Solubility of Water in *n*-Hexane, Cyclohexane and Benzene at elevated Temperatures and Pressures. *Deposited Doc. VINITI* **1973**, *2347*, 1–4.

(110) Sultanov, R. G.; Skripka, V. G.; Namiot, A. Y. Phase Equilibria in the System WaterâĹŠ *n*-Hexadecane at elevated Temperatures. *Deposited Doc. VINITI* **1971**, *3139*, 1.

(111) Ambrose, D.; Tsonopoulos, C. Vapor-Liquid Critical Properties of Elements and Compounds. 2. Normal Alkanes. *J. Chem. Eng. Data* **1995**, *40*, 531–546.

(112) Góral, M.; Wiśniewska-Gocłowska, B.; Mączyński, A. Recommended Liquid-Liquid Equilibrium Data. Part 4. 1-Alkanol-Water Systems. *J. Phys. Chem. Ref. Data* **2006**, *35*, 1391–1414.

(113) Fileti, E. E.; Chaudhuri, P.; Canuto, S. Relative strength of hydrogen bond interaction in alcohol–water complexes. *Chem. Phys. Lett.* **2004**, *400*, 494–499.

(114) Dallos, A.; Liszi, J. (Liquid + liquid) equilibria of (octan-1-ol + water) at temperatures from 288.15 K to 323.15 K. *J. Chem. Thermodyn.* **1995**, *27*, 447–448.

(115) Lang, B. E. Solubility of Water in Octan-1-ol from (275 to 369)K. *J. Chem. Eng. Data* **2012**, *57*, 2221–2226.

(116) Tunik, S.; Lesteva, T.; Chernaya, V. Phase Equilibria in the Water-Alcohols-Formaldehyde Systems. I. The Liquid-Vapor Equilibria in the Systems Water-C6 and C7 Alcohols. *Russ. J. Phys. Chem* **1977**, *51*, 751.

(117) Schmitt, M.; Hasse, H. Phase Equlibria for Hexyl Acetate Reactive Distillation. *J. Chem. Eng. Data* **2005**, *50*, 1677–1683.

(118) Zong, Z.-L.; Yang, X.-H.; Zheng, X.-Y. Determination and correlation of vapor-liquid equilibria of alcohol solutions. *J. Chem. Eng. Jpn.* **1983**, *16*, 1–6.

(119) Stephenson, R.; Stuart, J. Mutual binary solubilities: Water-alcohols and water-esters. *J. Chem. Eng. Data* **1986**, *31*, 56–70.

(120) Sánchez, F.; Pereda, S.; Brignole, E. GCA-EoS: A SAFT group contribution model - Extension to mixtures containing aromatic hydrocarbons and associating compounds . *Fluid Phase Equilib.* **2011**, *306*, 112–123.

(121) Soria, T. M.; Sánchez, F. A.; Pereda, S.; Bottini, S. B. Modeling the phase behavior of cyclic compounds in mixtures of water, alcohols and hydrocarbons. *Fluid Phase Equilib.* **2014**, *361*, 143–154.

(122) Rettich, T. R.; Handa, Y. P.; Battino, R.; Wilhelm, E. Solubility of gases in liquids. 13. High-precision determination of Henry's constants for methane and ethane in liquid water at 275 to 328 K. *J. Phys. Chem.* **1981**, *85*, 3230–3237.

(123) Crovetto, R.; Fernandez-Prini, R.; Japas, M. L. The solubility of ethane in water up to 473 K. *Ber. Bunsen-Ges. Phys. Chem.* **1984**, *88*, 484–488.

(124) Sarraute, S.; Delepine, H.; Costa Gomes, M. F.; Majer, V. Aqueous solubility, Henry's law constants and air/water partition coefficients of *n*-octane and two halogenated octanes. *Chemosphere* **2004**, *57*, 1543–51.

(125) Gruber, D.; Langenheim, D.; Gmehling, J.; Moollan, W. Measurement of Activity Coefficients at Infinite Dilution Using Gas-Liquid Chromatography. 6. Results for Systems Exhibiting Gas-Liquid Interface Adsorption with 1-Octanol. *J. Chem. Eng. Data* **1997**, *42*, 882–885.

(126) Kojima, K.; Zhang, S.; Hiaki, T. Measuring methods of infinite dilution activity coefficients and a database for systems including water. *Fluid Phase Equilib.* **1997**, *131*, 145–179.

(127) Tochigi, K.; Uchiyama, M.; Kojima, K. Measurement of infinite-dilution activity coefficients of alcohols in water using relative gas-liquid chromatographic method. *Korean J. Chem. Eng.* **2000**, *17*, 502–505.

(128) Stephenson, R.; Stuart, J.; Tabak, M. Mutual solubility of water and aliphatic alcohols. *J. Chem. Eng. Data* **1984**, *29*, 287–290.

(129) Tsonopoulos, C. Thermodynamic analysis of the mutual solubilities of hydrocarbons and water. *Fluid Phase Equilib.* **2001**, *186*, 185–206.

(130) Waterbeemd, H. V. D.; Gifford, E. ADMET in silico modelling: Towards prediction paradise? *Nat. Rev. Drug Discovery* **2003**, *2*, 192–204.

(131) Wienke, G.; Gmehling, J. Prediction of octanol-water partition coefficients, Henry coefficients and water solubilities using UNIFAC. *Toxicol. Environ. Chem.* **1998**, *65*, 57–86.

(132) Lin, S. T.; Sandler, S. I. Prediction of Octanol-Water Partition Coefficients Using a Group Contribution Solvation Model. *Ind. Eng. Chem. Res.* **1999**, *38*, 4081–4091.

(133) Lin, S. T.; Sandler, S. I. A Priori Phase Equilibrium Prediction from a Segment Contribution Solvation Model. *Ind. Eng. Chem. Res.* **2002**, *41*, 899–913.

(134) Garrido, N. M.; Economou, I. G.; Queimada, A. J.; Jorge, M.; Macedo, E. A. Prediction of the n-hexane/water and 1-octanol/water partition coefficients for environmentally relevant compounds using molecular simulation. *AIChE J.* **2012**, *58*, 1929–1938.

(135) Mohsen-Nia, M.; Ebrahimabadi, A. H.; Niknahad, B. Partition coefficient *n*-octanol/water of propranolol and atenolol at different temperatures: Experimental and theoretical studies. *J. Chem. Thermodyn.* **2012**, *54*, 393–397.

(136) Suzuki, T.; Kudo, Y. Automatic log P estimation based on combined additive modeling methods. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 155–98.

(137) Sangster, J. Octanol-water partition coefficients of simple organic compounds. *J. Phys. Chem. Ref. Data* **1989**, *18*, 1111–1229.

(138) Collander, R. The partition of organic compounds between higher alcohols and water. *Acta Chem. Scand.* **1951**, *5*, 774–780.

(139) Barbato, F.; La Rotonda, M. I.; Quaglia, F. Interactions of nonsteroidal antiinflammatory drugs with phospholipids: comparison between octanol/buffer partition coefficients and chromatographic indexes on immobilized artificial membranes. *J. Pharm. Sci.* **1997**, *86*, 225–229.

(140) OECD, Test No. 117: Partition Coefficient (*n*-octanol/water), HPLC Method.

(141) Scheytt, T.; Mersmann, P.; Lindstädt, R.; Heberer, T. 1-Octanol/Water Partition Coefficients of 5 Pharmaceuticals from Human Medical Care: Carbamazepine, Clofibric

Acid, Diclofenac, Ibuprofen, and Propyphenazone. *Water, Air, Soil Pollut.* **2005**, *165*, 3–11.

(142) Avdeef, A.; Box, K.; Comer, J.; Hibbert, C.; Tam, K. pH-Metric logP 10. Determination of liposomal membrane-water partition coefficients of ionizable drugs. *Pharm. Res.* **1998**, *15*, 209–215.

(143) Prausnitz, J. M.; Lichtenthaler, R. N.; de Azevedo, E. G. *Molecular thermodynamics of fluid-phase equilibria*; Pearson Education, 1998.

(144) Gmehling, J.; Krafczyk, J.; Ahlers, J.; Nebig, S.; Hunecker, I.; Eisel, M.; Fischer, D.; Krentscher, B.; Beyer, K. Pure compound data from DDB. *Dortmund Data Bank* **1983-2014**,

(145) Gracin, S.; ; Rasmuson, A. C. Solubility of Phenylacetic Acid, p-Hydroxyphenylacetic Acid, p-Aminophenylacetic Acid, p-Hydroxybenzoic Acid, and Ibuprofen in Pure Solvents. *J. Chem. Eng. Data* **2002**, *47*, 1379–1383.

(146) Roux, M. V.; Temprado, M.; Chickos, J. S. Vaporization, fusion and sublimation enthalpies of the dicarboxylic acids from C_4 to C14 and C16. *J. Chem. Thermodyn.* **2005**, *37*, 941 – 953.

(147) Apelblat, A.; Manzurola, E. Solubility of suberic, azelaic, levulinic, glycolic, and diglycolic acids in water from 278.25 K to 361.35 K . *J. Chem. Thermodyn.* **1990**, *22*, 289 – 292.

(148) Chen, S.; Xia, Q.; Li, D.; Yuan, W.-G.; Zhang, F.-B.; Zhang, G.-L. Solid-Liquid Equilibria of Nonanedioic Acid in Binary Ethanol + Water Solvent Mixtures from (292.35 to 345.52) K. *J. Chem. Eng. Data* **2009**, *54*, 1395–1399.

(149) Zhang, H.; Xie, C.; Liu, Z.; Gong, J.; Bao, Y.; Zhang, M.; Hao, H.; Hou, B.; xi-

ang Yin, Q. Identification and Molecular Understanding of the Odd-Even Effect of Dicarboxylic Acids Aqueous Solubility. *Ind. Eng. Chem. Res.* **2013**, *52*, 18458–18465.

(150) Fini, A.; Fazio, G.; Feroci, G. Solubility and solubilization properties of non-steroidal anti-inflammatory drugs. *Int. J. Pharm.* **1995**, *126*, 95–102.

(151) Soltanpour, S.; Jouyban, A. Solubility of acetaminophen and ibuprofen in polyethylene glycol 600, propylene glycol and water mixtures at 25 °C. *J. Mol. Liq.* **2010**, *155*, 80 – 84.

(152) Manrique, J.; Martinez, F. Solubility of ibuprofen in some ethanol+ water cosolvent mixtures at several temperatures. *Lat. Am. J. Pharm.* **2007**, *26*, 344.

(153) Bergström, C. A.; Norinder, U.; Luthman, K.; Artursson, P. Experimental and Computational Screening Models for Prediction of Aqueous Drug Solubility. *Pharm. Res.* **2002**, *19*, 182–188.

(154) Espitalier, F.; Biscans, B.; Laguerie, C. Physicochemical Data on Ketoprofen in Solutions. *J. Chem. Eng. Data* **1995**, *40*, 1222–1224.

(155) Wang, S.; Song, Z.; Wang, J.; Dong, Y.; Wu, M. Solubilities of Ibuprofen in Different Pure Solvents. *J. Chem. Eng. Data* **2010**, *55*, 5283–5285.

(156) Perlovich, G. L.; Kurkov, S. V.; Kinchin, A. N.; Bauer-Brandl, A. Thermodynamics of solutions IV: Solvation of ketoprofen in comparison with other NSAIDs. *J. Pharm. Sci.* **2003**, *92*, 2502–2511.

(157) Garzón, L.; Martínez, F. Temperature dependence of solubility for ibuprofen in some organic and aqueous solvents. *J. of Solution Chem.* **2004**, *33*, 1379–1395.

(158) Bustamante, P.; Pena, M. A.; Barra, J. The modified extended Hansen method to determine partial solubility parameters of drugs containing a single hydrogen bonding

group and their sodium derivatives: benzoic acid/Na and ibuprofen/Na. *Int. J. Pharm.* **2000**, *194*, 117–124.

(159) Yalkowsky, S. H.; Valvani, S. C.; Roseman, T. J. Solubility and partitioning VI: octanol solubility and octanol–water partition coefficients. *J. Pharm. Sci.* **1983**, *72*, 866–870.

(160) Murdande, S. B.; Pikal, M. J.; Shanker, R. M.; Bogner, R. H. Solubility Advantage of Amorphous Pharmaceuticals: II. Application of Quantitative Thermodynamic Relationships for Prediction of Solubility Enhancement in Structurally Diverse Insoluble Pharmaceuticals. *Pharm. Res.* **2010**, *27*, 2704–2714.

(161) Murdande, S. B.; Pikal, M. J.; Shanker, R. M.; Bogner, R. H. Solubility advantage of amorphous pharmaceuticals: I. A thermodynamic analysis. *J. Pharm. Sci.* **2010**, *99*, 1254–1264.

(162) Takahashi, O.; Kohno, Y.; Nishio, M. Relevance of Weak Hydrogen Bonds in the Conformation of Organic Compounds and Bioconjugates: Evidence from Recent Experimental Data and High-Level ab Initio MO Calculations. *Chem. Rev.* **2010**, *110*, 6049–6076.

(163) Papaioannou, V.; Calado, F.; Lafitte, T.; Dufal, S.; Sadeqzadeh, M.; Jackson, G.; Adjiman, C. S.; Galindo, A. Application of the SAFT-γ Mie group contribution equation of state to fluids of relevance to the oil and gas industry. *Fluid Phase Equilib.* **2016**, *416*, 104–119.

(164) Sadeqzadeh, M.; Papaioannou, V.; Dufal, S.; Adjiman, C. S.; Jackson, G.; Galindo, A. The development of asymmetric unlike induced association-site models to study the phase behaviour of aqueous mixtures comprising acetone, alkanes and alkyl carboxylic acids with the SAFT-γ Mie group contribution methodology. *Fluid Phase Equilib.* **2015**, *407*, 39–57.

(165) Hutacharoen, P. Prediction of Partition Coefficients and Solubilities of Active Pharmaceutical Ingredients with the SAFT-γ Mie Group-contribution Approach. Ph.D. thesis, Imperial College London, 2017.

Graphical TOC Entry