Evaluation of the SAFT-γ Mie force field with solvation free energy calculations

Isabela Q. Matos, Charlles R. A. Abreu*

Chemical Engineering Department, Escola de Química, Universidade Federal do Rio de Janeiro, Av. Athos da Silveira Ramos 149, Rio de Janeiro, RJ 21941-909, Brazil

Abstract

We studied the solvation free energy differences of aromatic solutes in aqueous and non-aqueous solvents with the SAFT-γ Mie force field. We obtained solvation free energy differences by carrying out molecular dynamics simulations at the expanded ensemble. The output from these simulations was then used to estimate the differences with the MBAR method. The results with solvents other than water had low absolute deviations to the experimental data. Meanwhile, the hydration free energy calculations required a binary interaction parameter estimated with output data from molecular dynamics in order to obtain accurate free energy differences. These results indicated some problems on the SAFT-γ Mie model for water, but, generally, proved that this coarse-grained model could represent the free energy differences of the studied sets of solute-solvent.

Keywords: solvation free energies, SAFT-γ Mie

1. Introduction

Solvation free energy calculations with molecular dynamics (MD) have a variety of applications ranging from drug design in the pharmaceutical industry to the development of separation technologies in the chemical industry. Solvation free energy is, more specifically, the difference in free energy related to the process of transferring the solute from the ideal gas phase condition to the liquid solvent phase condition [1]. Through the study of the solvation

*Corresponding author

Email address: abreu@eq.ufrj.br (Charlles R. A. Abreu)

phenomenon, it is possible to obtain information about the behavior of the solvent in different chemical environments and the influence of the solute's molecular geometry. It is also possible to calculate other important properties with the solvation free energy, namely the activity coefficient at infinite dilution, Henry constant and partition coefficients. Additionally, solvation free energy calculations can be part of the process of calculating solubility with molecular dynamics [2].

The solvation phenomenon described above is intrinsically complex. There are many competing forces interfering in the behavior of the solute-solvent interaction, and free energy simulations are susceptible to sampling problems in low energy regions. Various simulation methodologies were developed to enable estimations of free energy differences such as the expanded ensemble [3], thermodynamic integration [4], free energy perturbation (FEP) [5, 6, 7] and umbrella sampling [8]. Utilizing FEP methodologies, recent papers [9, 10] made available a big database of hydration free energy of small molecules using the GAFF force field. Beckstein et al. [11] also calculated the hydration free energies for fifty-two compounds with the OPLS-AA force field. They obtained an overall root mean square deviation to the experimental data of 1.75 kcal/mol and concluded that the reproducibility of the Lennard-Jones parameters is the main limiter of the precision of their results. Izairi and Kamberaj [12] also studied hydration free energies but with the intention of comparing the polar and nonpolar contributions. Garrido

et al. [13, 14] calculated the free energy of solvation of large alkanes in 1-octanol and water with three different force fields (TraPPE, GROMOS, OPLS-AA/TraPPE) and the solvation free energy of propane and benzene in non aqueous solvents like n-hexadecane, n-hexane, ethylbenzene, and acetone with the force fields TraPPE-UA and TraPPE-AA. Roy et al. [15] addressed the choice of the Lennard-Jones parameters for predicting solvation free energy in 1-octanol. They calculated the solvation free energy of a set of 205 small organic molecules in 1-octanol and found that the force field parametrization of n-octanol proposed by Kobryn and Kovalenko [16] provided the best agreement. Gonçalves and Stassen [17] calculated the free energy of solvation using the polarizable continuum model coupled to molecular dynamics computer simulation with the GROMOS force field. These calculations were done with a representative set of solutes and with the solvents tetrachloride, chloroform, and benzene. Using the GAFF and the polarizable AMOEBA force fields, Mohamed et al. [18] evaluated the solvation free energy of small molecules in toluene, chloroform, and acetonitrile, and obtained

a mean unsigned error of 1.22 kcal/mol for AMOEBA and 0.66 kcal/mol for GAFF. To define the role of solvent water in the docking structure determination of proteins, Matubayasi [19] developed a method to compute the solvation free energy of proteins while using OPLS-AA force field for the solutes and TIP3P for water. Genheden [20] expanded the Elba force field to calculate solvation free energies of more than 150 solutes taken from the Minnesota solvation database in polar (water, hexanol, octanol and nonanol) and apolar (hexane, octane, and nonane) solvents. They obtained mean absolute deviations of 1 kcal/mol for water and 1.5 kcal/mol for hexane. In this model, three carbons are represented by a single bead and water is also represented by a single bead.

As can be seen in the previous paragraph, solvation free energy simulation is performed in the literature using a variety force field since the choice of force field can be another influencing factor in the output of these calculations. Hence, we, in this study, assess the efficiencies and shortcomings of the SAFT-γ Mie coarse-grained force field [21] with free energy calculations for a variety of pairs solute-solvent. We chose a coarse-grained force fields because they generally reproduce free energy differences since the effects of reducing degrees of freedom in the entropy are counterbalanced by the reduction of enthalpic terms [22]. Additionally, the success of a coarse-grained force field is essential to decrease the computational time of solvation free energy calculations and to reveal deficiencies in the description of small molecules by these models [23, 1]. The SAFT-γ Mie coarse-grained force field was specifically picked because it uses, unlike the majority of the force fields, the Mie potential and because its method of obtaining parameters is more straightforward than other models. It was initially parameterized with pure component equilibrium and interfacial tension data [21], and this strategy has provided satisfactory results. Examples include the prediction of phase equilibrium of aromatic compounds, alkanes, light gases, and water [24, 25, 26], thermodynamic properties of carbon dioxide and methane [27], multiphase equilibrium of mixtures of water, carbon dioxide, and n-alkanes [28], and water/oil interfacial tension [29].

The solvents and solutes in our free energy calculations were picked to test the force field with standard sets used as a benchmark in solvation free energy calculations and with aromatic substances used as a model to asphaltenes. Asphaltenes are complicated to characterize by determining their composition on a molecular basis, but the literature broadly accepts that they can be described as a fraction of crude oil soluble in toluene and insolu-

ble in n-alkenes (pentane, hexane, heptane) [30]. They have motivated many studies with interest in developing models for their structure and behavior due to all the problems they can cause during their transportation and refining such as precipitation during the oil processing [31]. This precipitation issue is a recurrent problem due to the growing market of the production of crude oil in deep waters, which have conditions favorable to precipitation, such as pressure depletion and acid stimulation [32]. As an example, asphaltene precipitation due to pressure drop can clog oil production equipment and cause an almost exponential increase in the cost of production [33]. All these factors make the understanding of the behavior of asphaltenes in different chemical and physical environments relevant to the oil industry. As we said, asphaltene characterization still faces some issues. Hence, we choose to use polycyclic aromatic hydrocarbons (PAH'S), which have well-defined characteristics, to initially test the efficiency of the SAFT-γ Mie force field in describing the solvation phenomenon. The ones utilized in this work were phenanthrene, anthracene, and pyrene since they have similarities with asphaltenes regarding their solubility. Meanwhile, we selected compounds that are used to characterize asphaltenes (toluene, hexane) as solvents in our free energy calculations. We also tested the anti-solvent/solvent effect of carbon dioxide due to its influence in asphaltene precipitation during the oil processing [34]. With this study of solvation free energies with the SAFT-γ Mie model, we then intend to improve this force field and provide accurate free energy calculations of PAH's with a coarse-grained model. The correct description of these smaller asphaltene-like compounds by this force field opens up the possibility of obtaining satisfactory results for more complex asphaltene models with a less computational expensive force filed.

2. Computational Methods

2.1. SAFT-γ Mie Force Field

The SAFT-γ Mie Force Field uses a top-down coarse-graining methodology in its parameterization. This methodology aims to obtain the intermolecular parameters from macroscopic experimental data such as fluid-phase equilibrium or interfacial tension data. The idea is that the force field parameters estimated with the SAFT-VR Mie EoS can be used in molecular simulations since both the equation of state and the force field use the same explicit

intermolecular potential model (Mie potential):

$$U_{Mie}(r) = \epsilon \frac{\lambda_r}{\lambda_r - \lambda_a} \left(\frac{\lambda_r}{\lambda_a}\right)^{\left(\frac{\lambda_a}{\lambda_r - \lambda_a}\right)} \left[\left(\frac{\sigma}{r}\right)^{\lambda_r} - \left(\frac{\sigma}{r}\right)^{\lambda_a}\right]. \quad (1)$$

The parameter ϵ is the potential well depth, σ is the segment diameter, r is the distance between the spherical segments, λ_r is the repulsive exponent and λ_a is the attractive exponent. This correspondence between models has been used to parametrize a variety of fluids [35]. This force field has the advantage of incorporating the degrees of freedom provided by the use of the Mie Potential [24]. This flexibility offers the exploration of a vast parameter space without using an iterative simulation scheme [21]. Despite these advantages, the force field can be restricted by the shortcomings of the equation of state. As an example, the lack of an association term in the equation can cause an inadequate representation of the properties of hydrogen bonding compounds.

Each substance has initially five parameters to be estimated (m_s, σ, ϵ, λ_r, and, λ_a) according to Eq. 1. The number of segments are usually fixed in an integer value since each segment represents one pseudo atom. The attractive parameter is generally fixed due to its high correlation with the repulsive parameter. Usually, the chosen value for this parameter is 6, corresponding to the London model, which is a good representation of the dispersion scale of most simple fluids that do not have strong polar interactions [36, 24]. There are two strategies to obtain these parameters: one is by fitting the SAFT-Vr Mie EoS to experimental data as vapor pressure and liquid density [37], and the other one is by using correspondent state parametrization [38]. In the present work, the first strategy was used to find the parameters for phenanthrene with vapor-liquid equilibrium data [39, 40] following the methodology proposed by Müller and Mejía [25]. The parameterization was carried out with the number of segments equal to five and with a geometry such as that in Figure 1, since this level of coarse-graining was also used for a similar molecule (anthracene) in the original paper.

The parameters for the other compounds were retrieved from the literature, and all these parameters are exposed in Table 1. For a mixture, the mixing rules used on can be seen on Eqs. 2 to 4 [41] .

$$\sigma_{ij} = \frac{\sigma ii + \sigma jj}{2}, \quad (2)$$

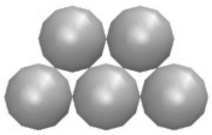

Figure 1: Representation of phenanthrene with the geometry proposed by Müller and Mejía [25].

$$\lambda_{k,ij} - 3 = \sqrt{(\lambda_{k,ii} - 3)(\lambda_{k,jj} - 3)}, \quad k = r, a, \tag{3}$$

$$\epsilon_{ij} = (1 - k_{ij}) \frac{\sqrt{\sigma_{ii}^3 \sigma_{jj}^3}}{\sigma_{ij}^3} \sqrt{\epsilon_{ii} \epsilon_{jj}}, \tag{4}$$

After the first estimations, we realized the need to estimate the binary interaction parameter of Eq. 4 for pairs with water as a solvent. Hence, we estimated k_{ij} for these pairs and, for all the other pairs, we set k_{ij} to zero. The estimation was done by performing trial expanded ensemble simulations in three values of the parameter, as suggested by Ervik et al [42]. With the ΔG_{solv} obtained with these simulations, we did a linear fit to acquire the refined value of the parameter. We used this strategy because the estimation with SAFT VR Mie EoS gave poor results for the solvation free energies.

Table 1: SAFT-γ Mie Force Field for each substance used in this work.

	m_s	ϵ/κ_b (K)	$\sigma(\dot{A})$	λ_r
Water [28]	1	305.21	2.902	8.0
Propane [24]	1	426.08	4.871	34.29
Carbon dioxide [24]	2	194.94	2.848	14.65
Hexane [24]	2	376.35	4.508	19.57
Octanol [35]	3	495.71	4.341	28.79
Toluene [25]	3	268.24	3.685	11.80
Benzene [25]	3	230.30	3.441	10.45
Pyrene [25]	4	459.04	4.134	15.79
Anthracene [25]	5	259.68	3.631	9.55
Phenanthrene	5	262.74	4.077	9.55

2.2. Expanded Ensemble

The strategy chosen in this work to calculate the solvation free energy differences was to use an alchemical method in which the solute molecule is gradually inserted in the solvent using a thermodynamic path [43]. Each insertion or alchemical state is represented by a coupling parameter, λ, that ranges from 0 to 1. When $\lambda = 0$, there is no interaction with the solvent and, when $\lambda = 1$, the interactions are fully activated. Since the force field used does not explicitly take in consideration the charges, the interactions are only due to the Mie potential. For the coupling of the Mie Potential, we propose generalized softcore Mie potential based on the softcore potential of Beutler et al. [44] :

$$\begin{aligned} U_{Mie}^{sc}(r) =& \lambda\epsilon \frac{\lambda_r}{\lambda_r - \lambda_a} \left(\frac{\lambda_r}{\lambda_a}\right)^{\left(\frac{\lambda_a}{\lambda_r - \lambda_a}\right)} \cdot \\ & \left\{ \frac{1}{[\alpha(1-\lambda) + (r/\sigma)^{\lambda_a}]^{\lambda_r/\lambda_a}} - \frac{1}{\alpha(1-\lambda) + (r/\sigma)^{\lambda_a}} \right\}. \end{aligned} \tag{5}$$

where α is a constant whose value is normally assumed to be 0.5. We decided to use the Expanded Ensemble method [3] in our solvation free energy simulations since it allows a non-Boltzmann sampling scheme of different states in a single simulation. In this scheme, the sampling is done by biasing the phase space exploration process with weights not related to the statistical ensemble. The partition function of the statistical expanded ensemble, Z^{EE}, is obtained from the probability distributions correspondent to each λ. Hence, Z^{EE} is defined as a sum of subensembles Z_i in different values of λ, that is,

$$Z = \sum_{i=1}^{N} Z_i exp(\eta_i), \tag{6}$$

where N is the number of alchemical states, η_i is the arbitrary weight of the subensemble at each state, and Z_i is the configurational partition function of state i. Here, we followed the flat-histogram approach [45, 46, 47] to calculate the weights. This strategy aims to obtain adequate sampling by ensuring that all the states have an equal number of visits, i.e. the ratio of the probability of sampling state i (π_i) to the probability of sampling state j (π_j) is equal to one. Using this relation, the following equation can be obtained:

$$(\eta_i - \eta_j)_{k+1} = \beta(G_i - G_j)_k. \tag{7}$$

Eq. 7 proposes that the choice of the new weights is dependent on the free energies that we are attempting to obtain. This equation is then solved iteratively with trial simulations. For the first simulation, the values of η are chosen or set to zero, and the histogram of the states visited is obtained. With this histogram, it is possible to estimate the free energy differences and, since the weights are related to the free energies by Eq. 7, the next values of η can be calculated. This iteration goes on until a uniform distribution is attained. The weights found are then used in a longer simulation to obtain the final solvation free energy differences. The choice of the λ set correspondent to overlapping alchemical states are crucial to acquire accurate free energy differences. In this work, the method chosen to obtain the optimal staging of the λ domain is the one developed by Escobedo and Martinez-Veracoechea [48] with a basis in the study of Katzgraber et al [49]. This method targets "bottlenecks" in the simulation. It does that by optimizing λ through the minimization of the number of round trips per CPU time between the lowest (0) and highest (1) values of λ. This is specifically done by maximizing the steady-state stream ϕ of the simulation, which "walks" among the values of λ. This flow is estimated from a Fick's diffusion type of law:

$$\phi = D(\Lambda)\Pi(\Lambda)\frac{dx(\Lambda)}{d\Lambda}. \tag{8}$$

In the equation above, Λ is the actual continuous value of the coupling parameter. This continuous function of $\lambda' s$ is obtained by interpolating the λ set linearly. $D(\Lambda)$ is the diffusivity at state Λ and $x(\Lambda)$ is the fraction of times that the trial simulation at state Λ_i has most recently visited the state $\lambda = 1$ as opposed to state $\lambda = 0$. The derivative $dx(\Lambda)/d\Lambda$ is approximated with the central finite differences method. Finally, $\Pi(\Lambda)$ is the probability of visiting Λ:

$$\Pi(\Lambda) = \frac{C'\bar{\Pi}(\lambda)}{\Lambda_{i+1} - \Lambda_i}. \tag{9}$$

The C' term in the equation above represents a constant and $\bar{\Pi}(\lambda)$ is the arithmetic average of the frequency of visits to the Λ state:

$$\bar{\Pi}_i(\lambda) = \frac{\pi_{i+1} - \pi_i}{2}. \tag{10}$$

The ϕ is maximum when the optimal probability $\Pi'(\Lambda_i)$ of visiting state Λ_i is proportional to $1/\sqrt{D(\Lambda)}$ [50]. With that information, it is possible to

estimate the diffusivity using one trial simulation with the following equation:

$$D(\Lambda) = \frac{\Lambda_{i+1} - \Lambda_i}{\bar{\Pi}(\lambda) dx(\Lambda)/d\Lambda}. \tag{11}$$

Hence, we can calculate $\bar{\Pi}$ and, consequently, the cumulative probability, which is used to obtain the new λ state, with

$$\Phi = \int_{\lambda=0}^{\lambda=1} \Pi^{'}(\Lambda_i) d\Lambda = \frac{i}{K}, \tag{12}$$

where K is the total number of λ states. We obtained these cumulative probabilities for every λ set we estimated in order to carry out our solvation free energy simulations.

3. Molecular Dynamics Simulations

Using the parameters of Table 1, we carried out molecular dynamics simulations to estimate solvation free energy differences. The chosen software package to perform the simulations was the LAMMPS [51]. In this package, the equations of motion were integrated with the velocity-Verlet algorithm [52] with a time step of 2 fs. As required by the coarse-grained model, molecules with more than one bead were treated as rigid bodies. The thermostat and the barostat were the Nosé Hoover chains as described initially in Hoover [53] and Martyna et al. [54] with damping factors of 100 and 1000 time steps, respectively. For the rigid bodies in our simulations, we used the rigid-body algorithm of Kamberaj

et al [55]. The potential cutoff was equal to 20 Å [25] with a neighbor list skin of 2 Å. The initial configurations of the solvated systems were generated using the Playmol package [56], which is integrated with the Packmol package [57]. For the binary mixtures, one molecule of solute and a varying number of solvent molecules- 700 molecules of toluene, 700 molecules of octanol, 1024 molecules of hexane, 3000 molecules of water - were randomly added to a cubic box. Besides the systems with pure substances acting as solvents, we performed simulations to study solvation free energy of phenanthrene in a mixture of toluene and carbon dioxide with different weight fractions (w_{CO_2}). The system consisted of one molecule of phenanthrene for all the cases and 123 molecules of CO_2 and 618 molecules of toluene ($w_{CO_2} = 0.087$); 166 molecules of CO_2 and 589 molecules of toluene ($w_{CO_2} = 0.119$); 232 molecules of CO_2 and 545 molecules of toluene ($w_{CO_2} = 0.169$); 380 molecules of CO_2

and 446 molecules of toluene ($w_{CO_2} = 0.289$). As we commented in the Introduction, the solvents and solutes used in this study were selected with the intention of testing the force field with standard sets used as a benchmark in solvation free energy calculations and with aromatic substances used as models to asphaltenes.

All simulations were performed with the constant temperature and pressure values of 298 K and 1 bar, except the ones containing carbon dioxide. These had the temperature of 298 K and the pressure of the experimental liquid-phase equilibrium correspondent to each composition of the system CO_2+toluene [58]. For all simulations, the initial box was equilibrated at the NPT ensemble for 2 ns, and the resulting configurations were used as the initial configuration of the expanded ensemble simulations. These were carried out with the LAMMPS user package for expanded ensemble simulations with the Mie Potential developed by our research group, available at https://github.com/atoms-ufrj/USER-ALCHEMICAL.

During these expanded ensemble simulations, the sampling of a new alchemical state was tried at every 10 MD steps. To define the optimal values of λ and η corresponding to each state, trial simulations, having around 9 ns of production time, were carried out. In the first simulation, we chose the group of λ values arbitrarily, and we either set all $\eta's$ to zero or assigned values previously found for similar solute-solvent pairs. The subsequent group of $\eta's$ were estimated with the flat histogram approach (Eq. 7). We then did another trial simulation with the new weights. The results of this simulation were used to optimize the group of $\lambda's$ by minimizing the number of round trips, as described in the previous section. The $\eta's$ corresponding to the newest group of $\lambda's$ were interpolated linearly from the free energy differences. With the final values of η and λ defined for each mixture, larger simulations with a production time of 20 ns were carried out.

Since the employed force field considers that the beads do not have charges, there are no Coulombic interactions and the only contribution to the total potential energy is due to the softcore potential of Eq. 5. The post-processing method used to effectively calculate free energy differences with the potential energies obtained from the expanded ensemble simulations was the Multistate Bennett Acceptance Ratio (MBAR) method [7]. The software alchemical-analysis [43] was utilized to obtain the ΔG_{solv} with MBAR and to assess the quality of the results. After the first estimations, we realized that the binary interaction parameter of Eq. 4 was necessary for systems containing water. Hence, we estimated k_{ij} for these pairs and, for all the

other pairs, we set k_{ij} to zero. The estimation was done by performing trial expanded ensemble simulations in three values of k_{ij}, as suggested by Ervik et al [42]. With the ΔG_{solv} obtained with these simulations, we did a linear fit to obtain the refined value of the parameter. We used this strategy because the estimation with SAFT VR Mie EoS gave poor results for the solvation free energies.

4. Results and discussion

4.1. Solvation free energies

Our primary intention with this study is to assess the capability of the SAFT-γ Mie force field to represent solvation free energies. Hence, we chose benchmark solutes used in the literature (benzene, propane) and polyaromatic solutes (benzene, pyrene, phenanthrene, anthracene), and, for the solvents, we picked non-polar (hexane), aromatic (toluene), and hydrogen bonding (1-octanol) substances. The solvation free energy simulations for the pairs chosen were carried with binary interaction parameters equal to zero since these parameters were not necessary according to our preliminary studies. Since the force field does not account for charges, we only calculated the Mie contribution (Eq. 5) to the solvation free energy. A total of 15 to 18 $\lambda's$, depending on the solute-solvent pairs, and their respective $\eta's$ were estimated as described in the previous sections. The final λ set for all the pairs was found using the cumulative probability distribution (Eq. 12). The distribution for the hexane(solvent)+benzene(solute) pair can be seen in Figure 2. The optimized values of λ and η for this pair are available in Table 2 and in Tables A.7 to A.10 for all the other pairs. Observing the coupling parameters found for all the pairs, we can see that they are concentrated on the region with a steeper slope as it is expected in this method.

After the expanded ensemble simulations with the optimized intermediate states and weights, we calculated the solvation free energy differences with MBAR. The results obtained and the absolute deviations to experimental data [59] are available in Table 3. The numerical values for solvation free energy differences in hexane had overall smaller absolute deviations to experimental data than the values in the other solvents. Additionally, this force field presented better results for the pair hexane+benzene than the TraPPE force field (- 4.35 $\pm$ 0.05 kcal/mol) [14] and the ELBA coarse-grained force field (-2.92 $\pm$ 0.01 kcal/mol) [20]. TraPPE is a force field parametrized with fluid-phase equilibria data that uses the Lennard-Jones potential to describe

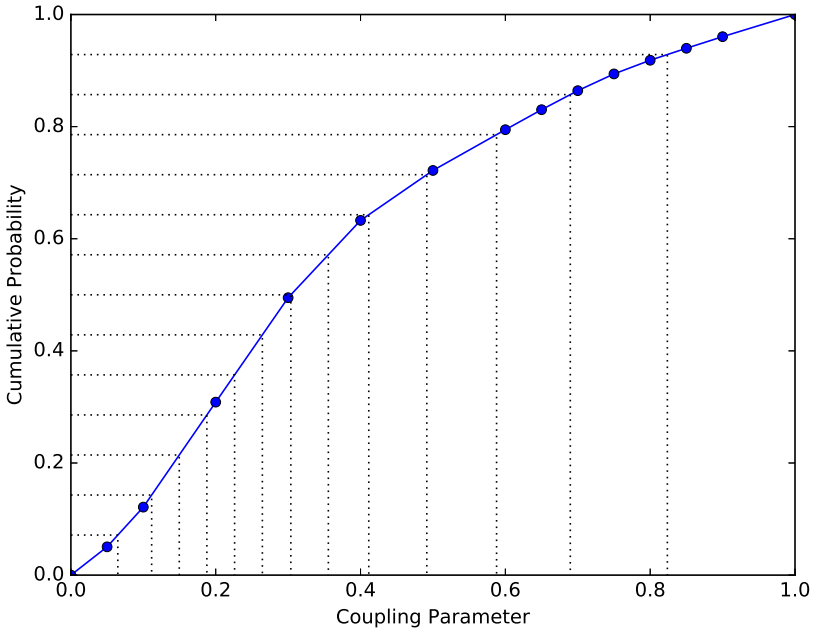

Figure 2: Cumulative probability used to obtain the optimized values of $\lambda's$ for the pair hexane+benzene.

the non-bonded interactions. In the cited paper, they used the united-atom description of the TraPPE force field for the alkyl group, the all-atom description for the polar groups and the explicit-hydrogen approach for the aromatic groups. In the explicit-hydrogen approach, the interaction sites for all hydrogen atoms, some lone pair electrons, and bond centers are accounted for [60]. In turn, the ELBA force field is a coarse-grained model that comprises six independent parameters. This force field models three carbons as one Lennard-Jones site and one water molecule as a single Lennard Jones site with a point dipole. The free energy profiles for all the pairs studied here are presented in Figures 3 to 5. Specifically observing the solvation free energy profiles in hexane (Figure 3), we can see the effect of the molecule's size on the entropic region of the free energy curve, that is, the region corresponding to the first values of λ where space in the solvent is being 'opened' for the insertion of the solute.

We expected that a force field based on an EoS that does not explicitly account for hydrogen bond would not perform well for 1-octanol in mixtures since the parameterization of this molecule did not explicitly account for the

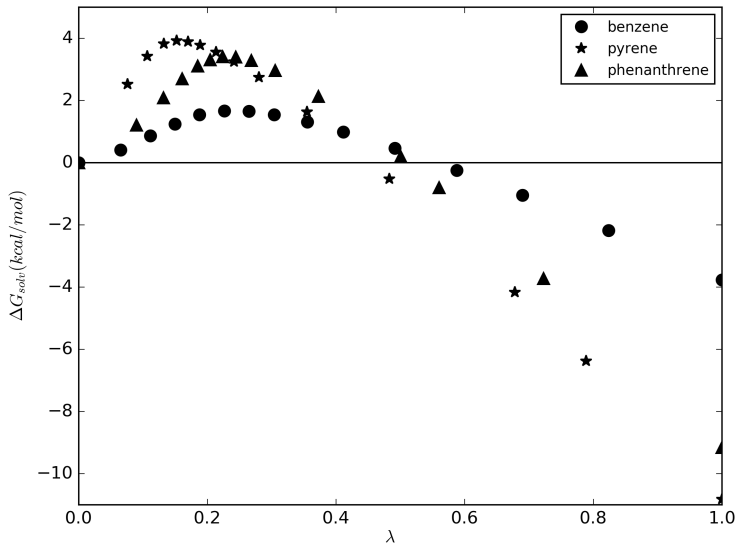

Figure 3: Representation of solvation free energy profiles obtained with MD simulations of different solutes in hexane.

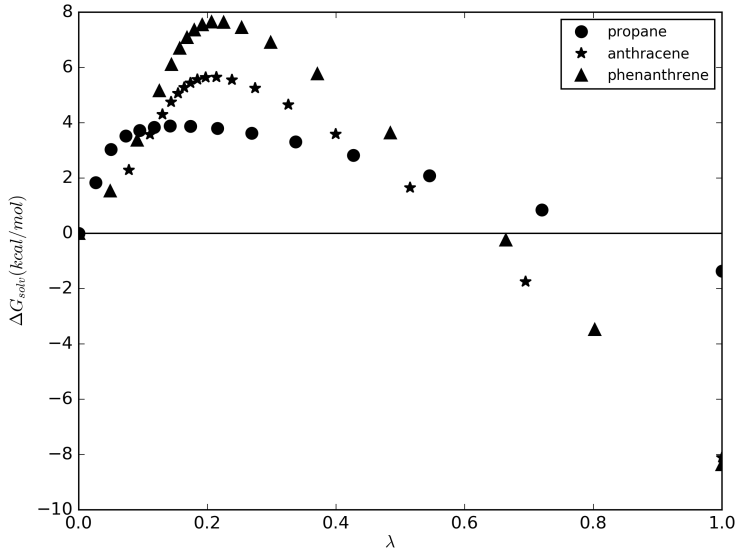

Figure 4: Representation of solvation free energy profiles obtained with MD simulations of different solutes in 1-octanol.

Table 2: Optimized values of λ and η for the pair hexane + benzene.

λ	η
0.000	0.000
0.065	0.708
0.112	1.385
0.15	1.892
0.188	2.399
0.226	2.519
0.264	2.457
0.304	2.367
0.356	1.921
0.411	1.411
0.492	0.524
0.588	-0.663
0.69	-2.016
0.824	-3.922
1.000	-6.583

association interactions. All the beads representing 1-octanol have the same intermolecular parameter, there is no distinction between the polar and apolar groups. Despite this, the solvation free energies of propane and phenanthrene in 1-octanol lied in the desired deviation range of 1-2 kcal/mol [61]. For propane, the observed deviation in solvation free energy differences was much smaller when compared to the other solutes, which can be attributed to the non-polarity of propane and smoother free energy curve (Figure 4). Such solvation free energy of propane in 1-octanol also had a smaller deviation than the prediction of the ELBA force field (-0.92 $\pm$ 0.01) [20]. The absolute deviation of the solvation free energy difference computed for anthracene in 1-octanol is much higher than the one calculated for phenanthrene in 1-octanol. The anthracene and phenanthrene molecules have the same geometry (Figure 1) in the SAFT-γ Mie model, although anthracene is a linear molecule and phenanthrene is not, and also similar physical properties. Hence, this high deviation of the solvation free energy of anthracene in 1-octanol may indicate a problem in the geometry chosen for anthracene in the SAFT-γ Mie force field and the importance of the geometry in modeling the molecules with this force field.

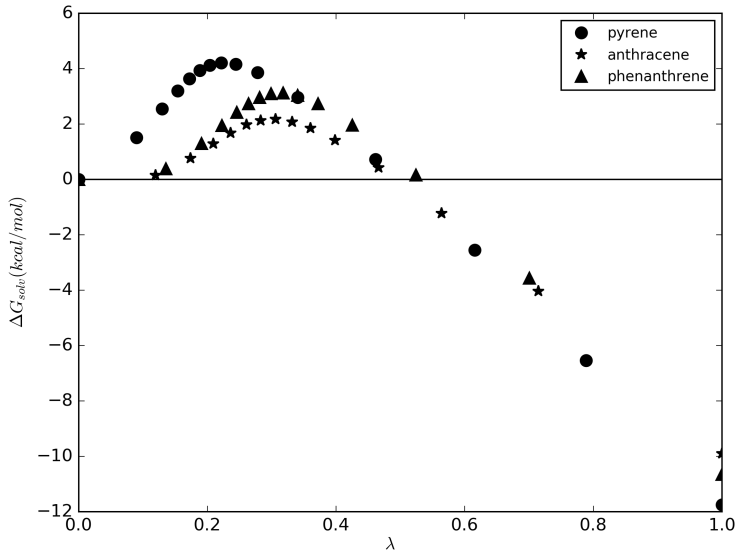

Figure 5: Representation of solvation free energy profiles obtained with MD simulations of different solutes in toluene.

The results also indicate the capability of the force field for predicting the solvation free energies of polyaromatic solutes in aromatic solvents. The influence of the molecular geometry on the solvation free energy curves was the same as the one observed for other solvents (Figure 5). ΔG_{solv} was also calculated for phenanthrene in pure toluene and in toluene+CO_2 mixtures. To the best of our knowledge, there were no available experimental data for these solvation free energies, but the previous results for phenanthrene in other solvents showed that the force field is adequate to describe the solvation phenomenon of phenanthrene in a pure aromatic solvent. The results for these sets are exposed in Table 4. The increase of the mass fraction of CO_2 in toluene caused a small effect on the solvation free energies in the range of fractions studied here. First, the ΔG_{solv} decreased with the increase of w_{CO_2} up to 0.119. After this, the effect was reversed, and carbon dioxide became an anti-solvent. Soroush et al. [34] reported that asphaltene precipitation occurs when carbon dioxide mass fractions became higher than 0.10 in the system asphaltene+toluene+carbon dioxide, which is in agreement with the anti-solvent effect of carbon dioxide observed in the values calculated here. However, since the differences in solvation free energy observed in Table 4 are

Table 3: Calculated and experimental values for the solvation free energy differences (kcal/mol) of solutes in non-aqueous solvents.

Solute	Solvent	ΔG^{exp}_{solv}	ΔG^{Mie}_{solv}	Absolute Deviation
benzene	hexane	-3.96	-3.76 $\pm$ 0.01	0.20
pyrene	hexane	-11.53	-10.82 $\pm$ 0.02	0.71
phenanthrene	hexane	-10.01	-9.16 $\pm$ 0.01	0.85
propane	1-octanol	-1.32	-1.36 $\pm$ 0.03	0.04
anthracene	1-octanol	-11.72	-8.12 $\pm$ 0.03	3.61
phenanthrene	1-octanol	-10.22	-8.34 $\pm$ 0.03	1.47
pyrene	toluene	-12.86	-11.74 $\pm$ 0.01	1.11
anthracene	toluene	-11.31	-9.90 $\pm$ 0.01	1.41

Table 4: Calculated values for the solvation free energy differences (kcal/mol) of phenanthrene in toluene+CO_2.

w_{CO_2}	ΔG^{Mie}_{solv}
0.0	-10.65 $\pm$ 0.02
0.087	-10.73 $\pm$ 0.02
0.119	-10.78 $\pm$ 0.02
0.169	-10.71 $\pm$ 0.02
0.289	-10.69 $\pm$ 0.02

slight, the effect of CO_2 may be insignificant in the solvation of phenanthrene in toluene when using the SAFT-γ Mie force field. Nevertheless, more studies need to be done to make a safe assertion about it. It is also worth remarking that this is a qualitative study due to the lack of experimental data. Overall, the methodology proposed by the SAFT-γ Mie force field was satisfactory in predicting the solvation free energies of the pairs solvent-solute studied here. For the pair 1-octanol+anthracene, the performance was not as good as it was for the other pairs. This highlights the importance of choosing a correct geometry for this coarse-grained force field.

4.2. Hydration free energies

Water is a solvent extensively used in experimental and computational studies. Because of this importance and the fact that water has unique properties, such as density maximum at 277 K and increased diffusivity upon

Table 5: Binary interaction parameters employed.

Pair	k_{ij}
water+propane	0.067
water+aromatic	0.154

compression, developing an accurate computational model for water is the subject of continuous research [62]. Hence, we also calculated the solvation free energies in water (hydration free energies) with the SAFT-γ Mie force field. With these calculations, we intend to verify if this coarse-grained model would represent the water molecule correctly and would be a good alternative to decrease the computational cost of solvation studies with asphaltene models. The simulations with water as a solvent were carried out using widely studied solutes (propane, benzene) and polyaromatic solutes (toluene, phenanthrene) with a set of fifteen intermediate states. We obtained these sets of λ and η with the same methodology used to acquire the sets for the solvation free energies with non-aqueous solvents. At first in our simulations, the binary interaction parameters of all aqueous mixtures were set to zero, but preliminary results for hydration free energies, displayed in Table 6, exhibited a high deviation from experimental data [63, 64]. With these results, the need for binary interaction parameters became clear. First, we estimated k_{ij} with the SAFT-VR Mie EoS and experimental vapor pressure data, but this strategy also provided results that had high absolute deviations to the experimental data. Hence, we used the approach of estimating the k_{ij} with the output from solvation free energy calculations with molecular dynamics, as described in the molecular dynamics simulations section. We initially found individual values for the interaction parameter of each pair, but, since the parameters for aromatic solutes were very similar (0.148, 0.162, 0.152), we averaged these values. By doing that, we obtained a general parameter for the water+aromatic pairs, which is exposed in Table 5. Also in this table, we display the binary interaction parameter for the pair water+propane.

The relatively large k_{ij} value of the interaction between aromatic solutes and water can be related to the lack of an explicit association term in the equation of state used to obtain the parameters for water. Actually, the SAFT-VR Mie has an association term [41], but it was not incorporated in the force field. The SAFT-γ Mie model for water [28] has two different temperature-dependent sets of parameters. The parameters utilized in this

work were those estimated with experimental interfacial tension data. Hence, we tested the only binary interaction parameter for water+toluene estimated with MD interfacial data available in the literature [29]. Nevertheless, the result was not satisfactory, and this parameter could not be transferred to the calculation of the solvation free energy of toluene in water.

These issues faced by SAFT-γ Mie model can also be related to the problems of modeling water with a coarse-grained force field. One of the main difficulties is the choice of which water molecules are going to be represented by which specific beads since water molecules move independently and are only bound by non-bonded interactions [65, 62]. The SAFT-γ Mie water considers that one water molecule corresponds to one bead. This strategy only saves a small amount of simulation time, but it can predict properties at physiological temperatures unlike other more aggressive models such as the MARTINI, which consider that one bead represents various water molecules. In light of all these facts, the SAFT-γ Mie force field appears to be a good alternative when working close to room temperatures, but the necessity of additional parameters estimated with molecular simulation indicates severe flaws in the methodology. This estimation of the binary parameter increased significantly the simulation time required to calculate the hydration free energies, since we had to carry out three additional simulations for every pair water-solute and then more three additional simulations for the three water+polyaromatic solutes in order to test the averaged binary interaction parameter. If these simulations are necessary for every time a new mixture with water is going to be studied with the SAFT-γ Mie force field, the use of this model can become impractical. With this idea in mind, a useful investigation to be made is to check how much other pairs of water+aromatic solute can be modeled using the binary interaction parameter estimated here. Using these binary interaction parameters calculated with data from molecular dynamics, we then obtained the final hydration free energy differences presented in Table 6.

Hydration free energy differences calculated using the SAFT-γ Mie force field with $k_{ij} \neq 0$ had low absolute deviations to the experimental data, as expected since the parameters were adjusted to fit these experimental data. In the table above, we also show the results obtained by Genheden [20] with the ELBA force field and by Mobley and Guthrie [9] with the GAFF force field for the solutes and with the TIP3P model for water. The GAFF (General Amber Force Field) force field is an all-atom model that consists of bonded and non-bonded parameters and is suitable for the study

Table 6: Calculated and experimental hydration free energy differences (kcal/mol) of solutes in water.

Solute	ΔG_{solv}^{exp}	ΔG_{solv}^{Mie} $k_{ij} = 0$	Absolute Deviation	ΔG_{solv}^{Mie} $k_{ij} \neq 0$	Absolute Deviation
propane	2.00 ± 0.20	1.10 ± 0.01	0.90	2.01 ± 0.01	0.01
benzene	-0.86 ± 0.20	-4.45 ± 0.03	3.59	-1.12 ± 0.01	0.26
toluene	-0.83 ± 0.20	-10.98 ± 0.30	10.15	-0.84 ± 0.01	0.01
phenanthrene	-3.88 ± 0.60	-10.90 ± 0.04	7.12	-3.47 ± 0.02	0.41

of a significant number of molecules. In turn, the TIP3P model considers that water is a rigid monomer represented by three interacting sites with non-bonded interactions and Coulombic interactions [66]. Both the GAFF and the TIP3P models use the Lennard-Jones potential to calculate the non-bonded interactions.

Comparing the three mentioned force fields, the root mean square error (RMSE) of all the pairs tested with the SAFT-γ Mie model was 0.24, the RMSE for hydration free energy differences obtained with the GAFF force field was 0.73, and that for the ELBA coarse-grained force field was 0.44. The difference in absolute deviations between the GAFF and SAFT-γ Mie force fields is significantly high for phenanthrene, hence the coarse-grained force field with a binary parameter is preferred if the application requires a high level of accuracy. The results also indicated that the SAFT-γ Mie model with the binary interaction parameter performed better than the ELBA force field in modeling the solvation phenomenon of the pairs studied in this work, but performed worse with the binary parameter set to zero. This occurred despite the fact that both the SAFT-γ Mie and ELBA models have the same level of coarse-graining for the solvent (one bead represents one water molecule). Hence, the choice between the two coarse-grained models is dependent on the availability and transferability of binary interaction parameters of the Mie Model. We also present, for the SAFT-γ Mie force field, the hydration free energy profiles in Figure 6. Bigger molecules had steeper free energy profiles, as it was for the solvation free energy study in other solvents. We also observe that the hydration free energy for the first non-zero λ is negative for benzene and toluene when a positive value is expected since free energy is required to 'open space' in the solvent for the solute's insertion. This anomaly can be caused by numerical errors during the estimation of the solvation free energy or by the fact that the attractive term in the Mie potential compensates the

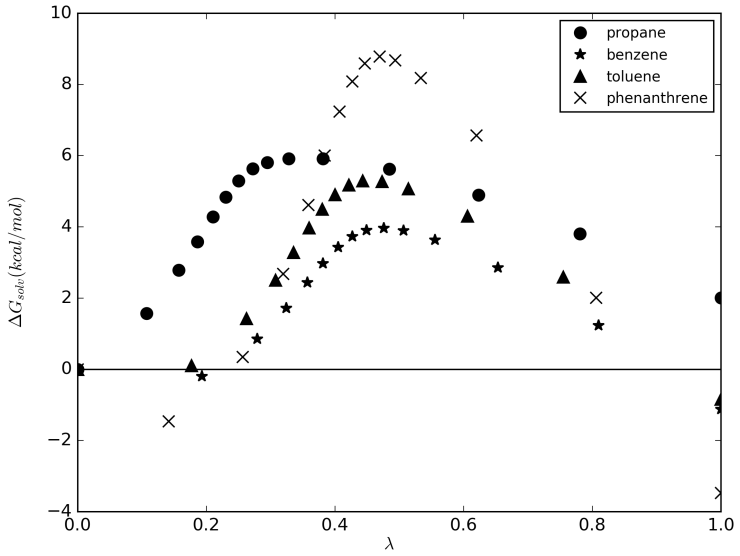

Figure 6: Hydration free energy profiles for different solutes.

need to open space.

The results found here for both the solvation free energies and hydration free energies fulfilled the intentions of this paper. We assessed the prediction capability of the SAFT-γ Mie force field and provided satisfactory solvation free energy estimates of PAH's using a coarse-grained force field. In addition to that, we found flaws in the methodology used by the SAFT-γ Mie force field to model the water molecule. Hence, these shortcomings of this model can now be addressed, and the force field can even be improved by using other mixing rules to avoid the use of a binary parameter or, even, using hydration free energy estimates in the parameterization of water. These results also encourage us to calculate solvation free energies of more complex molecules mimicking asphaltenes in non-aqueous solvents in future studies.

5. Conclusions

This study consisted of the calculation of solvation free energy differences of aromatic solutes in different solvents with the SAFT-γ Mie coarse-grained force field. Solvation free energy studies are mostly done using water as a

solvent and with all-atom force fields based on the Lennard-Jones Potential. Therefore, we were able to provide data that were lacking in the literature about the performance of a coarse-grained force field based on the Mie Potential for calculating solvation free energy differences. Additionally, the solvation free energy estimations carried out here can help improve the SAFT-γ Mie force field since these calculations are helpful in identifying errors in the modeling process. The SAFT-γ Mie uses the SAFT-VR Mie EoS in its parameterization, which results in a more straightforward method of obtaining parameters. Following this strategy, the phenanthrene parameters, which were not available in this force field database, were obtained using vapor-liquid equilibrium data.

To obtain accurate solvation free energies, we carefully selected and optimized the coupling parameter and their respective simulation weights used in our Expanded Ensemble simulations. The resulting potential energies from these simulations were then served as input to estimate solvation free energy differences with the MBAR method. The results for solvation free energy differences with non-aqueous solvents had absolute deviations to the experimental data of less than 2.0 kcal/mol, except for the pair 1-octanol+anthracene. We also observed the geometry effect on the free energy curves - larger molecules had steeper curves and more substantial absolute deviations. The influence of carbon dioxide on the solvation free energy of phenanthrene in toluene was found to be minimum according to the SAFT-γ Mie force field. Hydration free energy differences calculations with the SAFT-γ Mie model required the use of relatively large values of k_{ij} to obtain satisfactory results. We chose to estimate the parameter with the output from molecular dynamics data since the strategy of using the SAFT-VR Mie EoS provided high absolute deviations. This necessity of one additional parameter probably happens due to the lack of a term to account for the hydrogen bond on the EoS that this force field is based and due to the problems associated with the coarse-graining of water molecules. The results with k_{ij} estimated with MD output were satisfactory, the absolute deviations to the experimental data found were smaller than the ones for the GAFF and ELBA force field.

Overall, the SAFT-γ Mie force field proved to be an excellent model to represent the solvation phenomenon of non-aqueous solvents. It correctly described solvation free energy differences of solutes mimicking asphaltenes in hexane, toluene, 1-octanol. However, the calculation of hydration free energies required the use of a binary interaction parameter estimated with MD output, what significantly increased the simulation time. This fact evi-

denced flaws in the methodology used by the SAFT-γ force field and made us question the feasibility of this model when studying hydration free energies.

6. Acknowledgement

The authors thank the financial support provided by Petrobras (project code:).

References

[1] M. R. Shirts, J. W. Pitera, W. C. Swope, V. S. Pande, Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins, J. Chem. Phys. 119 (2003) 5740. doi:10.1063/1.1587119.

[2] M. J. Schnieders, J. Baltrusaitis, Y. Shi, G. Chattree, L. Zheng, W. Yang, P. Ren, The structure, thermodynamics, and solubility of organic crystals from simulation with a polarizable force field, J. Chem. Theory Comput. 8 (5) (2012) 1721–1736. doi:10.1021/ct300035u.

[3] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, P. N. VorontsovVelyaminov, New approach to monte carlo calculation of the free energy: Method of expanded ensembles, J. Chem. Phys. 96 (1992) 1776–1783. doi:10.1063/1.462133.

[4] J. Kirkwood, Statistical mechanics of fluid mixtures, J. Chem. Phys. 3 (1935) 300313. doi:10.1063/1.1749657.

[5] R. W. Zwanzig, Hightemperature equation of state by a perturbation method. i. nonpolar gases, J. Chem. Phys. 22 (1954) 1420. doi:10.1063/1.1740409.

[6] C. Bennett, Efficient estimation of free energy differences from monte carlo data, J. Comput. Phys. 22 (1976) 245268. doi:10.1016/0021-9991(76)90078-4.

[7] M. R. Shirts, J. D. Chodera, Statistically optimal analysis of samples from multiple equilibrium states, J. Chem. Phys. 129 (2008) 124105. doi:10.1063/1.2978177.

[8] G. Torrie, J. Valleau, Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling, J. Comput. Phys. 23 (2) (1977) 187 – 199. doi:10.1016/0021-9991(77)90121-8.

[9] D. L. Mobley, J. P. Guthrie, Freesolv: a database of experimental and calculated hydration free energies, with input files, J. Comput. Aided Mol. Des. 28 (7) (2014) 711720. doi:10.1007/s10822-014-9747-x.

[10] G. D. R. Matos, D. Y. Kyu, H. H. Loeffler, J. D. Chodera, M. R. Shirts, , D. L. Mobley, Approaches for calculating solvation free energies and enthalpies demonstrated with an update of the freesolv database, J. Chem. Eng. Data 62 (2017) 1559–1569. doi:10.1021/acs.jced.7b00104.

[11] O. Beckstein, A. Fourrier, B. I. Iorga, Prediction of hydration free energies for the sampl4 diverse set of compounds using molecular dynamics simulations with the opls-aa force field, J. Comput. Aided Mol. Des. 28 (3) (2014) 265–276. doi:10.1007/s10822-014-9727-1.

[12] R. Izairi, H. Kamberaj, Comparison study of polar and nonpolar contributions to solvation free energy, J. Chem. Inf. Model. 57 (2017) 25392553. doi:10.1021/acs.jcim.7b00368.

[13] N. M. Garrido, A. J. Queimada, M. Jorge, E. A. Macedo, I. G. Economou, 1-octanol/water partition coefficients of n-alkanes from molecular simulations of absolute solvation free energies, J. Chem. Theory Comput. 5 (2009) 2436–2446. doi:10.1021/ct900214y.

[14] N. M. Garrido, M. Jorge, A. J. Queimada, E. A. Macedo, I. G. Economou, Using molecular simulation to predict solute solvation and partition coefficients in solvents of different polarity, Phys. Chem. Chem. Phys 20 (2011) 9155–9164. doi:10.1039/C1CP20110G.

[15] D. Roy, N. Blinov, A. Kovalenko, Predicting accurate solvation free energy in n-octanol using 3d-rism-kh molecular theory of solvation: Making right choices, J. Phys. Chem. B, 121 (2017) 92689273. doi:10.1021/acs.jpcb.7b06375.

[16] A. E. Kobryn, A. Kovalenko, Molecular theory of hydrodynamic boundary conditions in nanofluidics, J. Chem. Phys. 129 (13) (2008) 134701. doi:10.1063/1.2972978.

[17] P. F. B. Gonçalves, H. Stassen, Free energy of solvation from molecular dynamics simulation applying voronoi-delaunay triangulation to the cavity creation, J. Chem. Phys. 123 (2005) 214109. doi:10.1063/1.2132282.

[18] N. A. Mohamed, R. T. Bradshaw, J. W. Essex, Evaluation of solvation free energies for small molecules with the amoeba polarizable force field, J. Comput. Chem. 37 (2016) 2749–2758. doi:10.1002/jcc.24500.

[19] N. Matubayasi, Free-energy analysis of protein solvation with all-atom molecular dynamics simulation combined with a theory of solutions, Curr. Opin. Struct. Biol. 43 (2017) 45 – 54. doi:https://doi.org/10.1016/j.sbi.2016.10.005.

[20] S. Genheden, Predicting partition coefficients with a simple all-atom/coarse-grained hybrid model, J. Chem. Theory Comput. 12 (1) (2016) 297–304. doi:10.1021/acs.jctc.5b00963.

[21] C. Avendaño, T. Lafitte, A. Galindo, C. S. Adjiman, G. Jackson, E. A. Muller, Saft-γ force field for the simulation of molecular fluids.1. a single-site coarse grained model of carbon dioxide, J. Phys. Chem. B 115 (2011) 11154–11169. doi:10.1021/jp204908d.

[22] S. Kmiecik, D. Gront, M. Kolinski, L. Wieteska, A. E. Dawid, A. Kolinski, Coarse-grained protein models and their applications, Chem. Rev. 116 (2016) 78987936. doi:10.1021/acs.chemrev.6b00163.

[23] D. L. Mobley, E. Dumont, J. D. Chodera, K. A. Dill, Comparison of charge models for fixed charge force fields: small-molecule hydration free energies in explicit solvent, J. Chem. Phys. B 111 (9) (2007) 2242–2254. doi:10.1021/jp0667442.

[24] C. Herdes, T. S. Totton, E. A. Müller, Coarse grained force field for the molecular simulation of natural gases and condensates, Fluid Phase Equilib. 406 (2015) 91–100. doi:10.1016/j.fluid.2015.07.014.

[25] E. A. Müller, A. Mejía, Extension of the saft-vr mie eos to model homonuclear rings and its parametrization based on the principle of corresponding states, Langmuir - (2017) AL. doi:10.1021/acs.langmuir.7b00976.

[26] O. Lobanova, C. Avendaño, T. Lafitte, E. A. Müller, G. Jackson, Saft-γ force field for the simulation of molecular fluids: 4. a single-site coarse-grained model of water applicable over a wide temperature range, Mol. Phys. 113 (2015) 12281249. doi:10.1080/00268976.2015.1004804.

[27] C. G. Aimoli, E. J. Maginn, C. R. Abreu, Force field comparison and thermodynamic property calculation of supercritical co_2 and ch_4 using molecular dynamics simulations, Fluid Phase Equilib. 368 (2014) 80–90. doi:10.1016/j.fluid.2014.02.001.

[28] O. Lobanova, A. Mejía, G. Jackson, E. A. Müller, Saft-γ force field for the simulation of molecular fluids 6: Binaryand ternary mixtures comprising water, carbon dioxide, and n-alkanes, J. Chem. Thermodyn. 93 (2016) 320–336. doi:10.1016/j.jct.2015.10.011.

[29] C. Herdes, A. Ervik, A. Mejía, E. A. Müller, Prediction of the water/oil interfacial tension from molecular simulations using the coarse-grained saft-γ mie force field, Fluid Phase Equilib.doi:10.1016/j.fluid.2017.06.016.

[30] J. Sjöblom, N. Aske, I. H. Auflem, ystein Brandal, T. E. Havre, ystein Sther, A. Westvik, E. E. Johnsen, H. Kallevik, Our current understanding of water-in-crude oil emulsions.: Recent characterization techniques and high pressure performance, Adv. Colloid Interface Sci. 100-102 (Supplement C) (2003) 399 – 473. doi:https://doi.org/10.1016/S0001-8686(02)00066-0.

[31] J. Sjöblom, S. Simon, Z. Xu, Model molecules mimicking asphaltenes, Adv. Colloid Interface Sci. 218 (Supplement C) (2015) 1 – 16. doi:https://doi.org/10.1016/j.cis.2015.01.002.

[32] E. Buenrostro-Gonzalez, C. Lira-Galeana, A. Gil-Villegas, J. Wu, Asphaltene precipitation in crude oils: Theory and experiments, AIChE J. 50 (10) (2004) 2552–2570. doi:10.1002/aic.10243.

[33] N. B. Joshi, O. C. Mullins, A. Jamaluddin, J. Creek, J. McFadden, Asphaltene precipitation from live crude oil, Energy Fuels 15 (4) (2001) 979–986. doi:10.1021/ef0100471.

[34] S. Soroush, E. J. Straver, E. S. J. Rudolph, C. J. Peters, T. W. de Loos, P. L. Zitha, M. Vafaie-Sefti, Phase behavior of the ternary system carbon

dioxide+toluene+asphaltene, Fuel 137 (Supplement C) (2014) 405 – 411. doi:https://doi.org/10.1016/j.fuel.2014.05.043.

[35] A. Ervik, A. Mejía, E. A. Müller, Bottled saft: A web app providing saft-γ mie force field parameters for thousands of molecular fluids, J. Chem. Inf. Model. 56 (2016) 1609–1614. doi:10.1021/acs.jcim.6b00149.

[36] N. Ramrattan, C. Avendaño, E. Müller, A. Galindo, A corresponding-states framework for the description of the mie family of intermolecular potentials, Mol. Phys. 113 (2015) 1–16. doi:10.1080/00268976.2015.1025112.

[37] C. Avendaño, T. Lafitte, C. S. Adjiman, A. Galindo, E. A. Muller, G. Jackson, Saft-γ force field for the simulation of molecular fluids.3. coarsegrained models of greenhouse gases, refrigerants, and long alkanes, J. Phys. Chem. B 117 (2013) 2717–2733. doi:10.1021/jp306442b.

[38] A. Mejía, C. Herdes, E. A. Müller, Force fields for coarse-grained molecular simulations from a corresponding states correlation, Ind. Eng. Chem. Res. 53 (2014) 4131–4141. doi:10.1021/ie404247e.

[39] S. Mortimer, R. Murphy, The vapor pressures of some substances found in coal tar, Ind. Eng. Chem. Res. 14 (1923) 1140–1142. doi:10.1021/ie50167a012.

[40] A. G. Osborn, D. R. Douslin, Vapor pressures and derived enthalpies of vaporization for some condensed-ring hydrocarbons, J. Chem. Eng. Data 20 (1975) 229–231. doi:10.1021/je60066a022.

[41] T. Lafitte, A. Apostolakou, C. Avendano, A. Galindo, C. S. Adjiman, E. A. Muller, G. Jackson, Accurate statistical associating fluid theory for chain molecules formed from mie segments, J. Chem. Phys. 139 (2013) 154504. doi:10.1063/1.4819786.

[42] A. Ervik, M. O. Lysgaard, C. Herdes, G. Jiménez-Serratos, E. A. MÏller, S. T. Munkejord, B. Müller, A multiscale method for simulating fluid interfaces covered with large molecules such as asphaltenes, J. Comput. Phys 327 (2016) 576–611. doi:10.1016/j.jcp.2016.09.039.

[43] P. V. Klimovich, M. R. Shirts, D. L. Mobley, Guidelines for the analysis of free energy calculations, J. Comput. Aided Mol. Des. 29 (2015) 397–411. doi:10.1007/s10822-015-9840-9.

[44] T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber, W. F. van Gunsteren, Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations, Chem. Phys. Lett 6 (1994) 529–539. doi:10.1016/0009-2614(94)00397-1.

[45] B. A. Berg, T. Neuhaus, Multicanonical ensemble: A new approach to simulate first-order phase transitions, Phys. Rev. Lett. 68 (1992) 9–12. doi:10.1103/PhysRevLett.68.9.

[46] J. Lee, New monte carlo algorithm: Entropic sampling, Phys. Rev. Lett. 71 (1993) 211–214. doi:10.1103/PhysRevLett.71.211.

[47] P. Dayal, S. Trebst, S. Wessel, D. Würtz, M. Troyer, S. Sabhapandit, S. N. Coppersmith, Performance limitations of flat-histogram methods, Phys. Rev. Lett. 92 (2004) 097201. doi:10.1103/PhysRevLett.92.097201.

[48] F. A. Escobedo, F. J. Martinez-Veracoechea, Optimized expanded ensembles for simulations involving molecular insertions and deletions. i. closed systems, J. Chem. Phys 127 (2007) 174103. doi:10.1063/1.2800320.

[49] H. G. Katzgraber, S. Trebst, D. A. Huse, M. Troyer, Feedback-optimized parallel tempering monte carlo, J. Stat. Mech. Theory Exp. 2006 (2006) P03018. doi:10.1088/1742-5468/2006/03/P03018.

[50] S. Trebst, D. A. Huse, M. Troyer, Optimizing the ensemble for equilibration in broad-histogram monte carlo simulations, Phys. Rev. E 70 (2004) 046701. doi:10.1103/PhysRevE.70.046701.

[51] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, J. Comp. Phys. 117 (1995) 1–19. doi:10.1006/jcph.1995.1039.

[52] L. Verlet, Computer "experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules, Phys. Rev. 159 (1967) 98–103. doi:10.1006/jcph.1995.1039.

[53] W. G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, Phys. Rev. A 31 (1985) 1695–1697. doi:10.1103/PhysRevA.31.1695.

[54] G. J. Martyna, M. L. Klein, M. Tuckerman, Nosé - hoover chains: The canonical ensemble via continuous dynamics, J. Chem. Phys. 97 (1992) 2635. doi:10.1063/1.463940.

[55] H. Kamberaj, R. Low, M. Neal, Time reversible and symplectic integrators for molecular dynamics simulations of rigid molecules, J. Chem. Phys. 122 (2005) 224114. doi:10.1063/1.1906216.

[56] C. R. A. ABREU, Playmol, http://atoms.peq.coppe.ufrj.br/playmol/index.html, accessed: 2017-03-10 (2017).

[57] L. Martínez, R. Andrade, E. G. Birgin, J. M. Martínez, Packmol: a package for building initial configurations for molecular dynamics simulations, J. Comput. Chem. 30 (2009) 2157–2164. doi:10.1002/jcc.21224.

[58] C. J.Chang, The solubility of carbon dioxide in organic solvents at elevated pressures, Fluid Phase Equilib. 15 (1992) 235–242. doi:10.1016/0378-3812(92)85064-F.

[59] A. R. Katritzky, A. A. Oliferenko, P. V. Oliferenko, R. Petrukhin, D. B. Tatham, U. Maran, A. Lomaka, W. E. Acree, A general treatment of solubility. 1. the qspr correlation of solvation free energies of single solutes in series of solvents, J. Chem. Inf. Comput. Sci. 43 (6) (2003) 1794–1805. doi:10.1021/ci034120c.

[60] N. Rai, J. I. Siepmann, Transferable potentials for phase equilibria. 9. explicit hydrogen description of benzene and five-membered and six-membered heterocyclic aromatic compounds, The Journal of Physical Chemistry B 111 (2007) 10790–10799. doi:10.1021/jp073586l.

[61] D. L. Mobley, M. K. Gilson, Predicting binding free energies: Frontiers and benchmarks, Annu. Rev. Biophys. 46 (1) (2017) 531–558. doi:10.1146/annurev-biophys-070816-033654.

[62] K. R. Hadley, C. McCabe, Coarse-grained molecular models of water: a review, Mol. Simul. 38 (2012) 671681. doi:10.1080/08927022.2012.671942.

[63] M. H. Abraham, G. S. Whiting, R. Fuchs, E. J. Chambers, Thermodynamics of solute transfer from water to hexadecane, J. Chem. Soc., Perkin Trans. 2 (1990) 291–300doi:10.1039/P29900000291.

[64] R. C. Rizzo, T. Aynechi, D. A. Case, I. D. Kuntz, Estimation of absolute free energies of hydration using continuum methods accuracy of partial charge models and optimization of nonpolar contributions, J. Chem. Theory Comput. 2 (1) (2006) 128–139. doi:10.1021/ct0500971.

[65] K. R. Hadley, C. McCabe, On the investigation of coarse-grained models for water: Balancing computational efficiency and the retention of structural properties, J. Phys. Chem. B 114 (13) (2010) 4590–4599. doi:10.1021/jp911894a.

[66] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water, The Journal of Chemical Physics 79 (1983) 926–935. doi:10.1063/1.445869.

Appendix A.

Table A.7: Optimized values of λ and η for the hexane+solute pairs.

pyrene		phenanthrene	
λ	η	λ	η
0.000	0.000	0.000	0.000
0.076	4.234	0.090	1.981
0.107	5.620	0.132	3.461
0.132	6.499	0.161	4.494
0.152	6.690	0.185	5.185
0.170	6.643	0.205	5.552
0.189	6.461	0.224	5.725
0.213	6.091	0.244	5.722
0.242	5.566	0.268	5.523
0.280	4.729	0.305	4.975
0.355	2.853	0.372	3.576
0.483	-0.778	0.500	0.297
0.678	-6.947	0.560	-1.390
0.788	-10.631	0.722	-6.309
1.000	-18.141	1.000	-15.448

Table A.8: Optimized values of λ and η for the 1-octanol+solute pairs.

propane		anthracene		phenanthrene	
λ	η	λ	η	λ	η
0.000	0.000	0.000	0.000	0.000	0.000
0.027	3.126	0.078	3.932	0.049	2.578
0.050	5.109	0.111	6.178	0.091	5.663
0.073	6.093	0.130	7.426	0.125	8.575
0.095	6.570	0.143	8.201	0.144	10.069
0.117	6.826	0.154	8.717	0.157	10.978
0.142	6.956	0.164	9.085	0.169	11.599
0.174	6.969	0.174	9.357	0.180	12.040
0.215	6.847	0.184	9.556	0.192	12.340
0.269	6.554	0.197	9.676	0.206	12.499
0.337	6.050	0.214	9.681	0.225	12.478
0.427	5.228	0.238	9.490	0.253	12.161
0.545	3.955	0.274	8.958	0.298	11.280
0.720	1.843	0.326	7.906	0.371	9.406
1.000	-1.903	0.399	6.088	0.484	5.891
		0.515	2.777	0.664	-0.516
		0.695	-2.960	0.802	-5.908
		1.000	-13.768	1.000	-14.073

Table A.9: Optimized values of λ and η for the toluene+solute pairs.

pyrene		anthracene		phenanthrene	
λ	η	λ	η	λ	η
0.000	0.000	0.000	0.000	0.000	0.000
0.090	2.563	0.119	0.218	0.136	0.726
0.130	4.338	0.174	1.210	0.191	2.307
0.154	5.439	0.209	2.052	0.223	3.430
0.172	6.181	0.236	2.664	0.246	4.233
0.188	6.670	0.261	3.122	0.264	4.780
0.204	6.986	0.283	3.378	0.281	5.149
0.222	7.121	0.306	3.449	0.299	5.354
0.244	7.025	0.332	3.311	0.318	5.389
0.278	6.520	0.360	2.936	0.340	5.222
0.340	5.010	0.399	2.209	0.372	4.717
0.462	1.247	0.466	0.567	0.425	3.440
0.616	-4.283	0.564	-2.211	0.524	0.444
0.788	-11.032	0.715	-6.983	0.701	-5.814
1.000	-19.814	1.000	-16.923	1.000	-17.803

Table A.10: Optimized values of λ and η for the phenanthrene+CO_2+solute pairs with different values of w_{CO_2}.

$w_{CO_2} = 0.087$		$w_{CO_2} = 0.119$		$w_{CO_2} = 0.169$		$w_{CO_2} = 0.289$	
λ	η	λ	η	λ	η	λ	η
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.128	0.604	0.128	0.732	0.064	0.883	0.066	0.806
0.184	2.067	0.186	2.223	0.108	0.764	0.111	0.760
0.217	3.164	0.219	3.319	0.175	1.969	0.172	1.983
0.240	3.940	0.244	4.098	0.214	3.156	0.204	2.967
0.260	4.472	0.267	4.704	0.240	3.974	0.227	3.627
0.277	4.823	0.289	5.031	0.258	4.457	0.245	4.082
0.295	5.035	0.313	5.084	0.273	4.750	0.262	4.395
0.318	5.059	0.339	4.950	0.287	4.921	0.279	4.583
0.347	4.762	0.373	4.371	0.305	4.962	0.299	4.621
0.397	3.753	0.425	3.055	0.326	4.885	0.325	4.423
0.491	1.031	0.488	1.196	0.361	4.401	0.365	3.739
0.670	-5.148	0.525	-0.027	0.419	2.990	0.428	2.198
0.791	-9.713	0.730	-7.185	0.527	-0.299	0.530	-0.842
1.000	-18.098	1.000	-17.769	0.697	-6.180	0.701	-6.763
				1.000	-17.998	1.000	-18.163