



Strathprints Institutional Repository

Garrido, Nuno M. and Jorge, Miguel and Queimada, Antonio J. and Macedo, Eugenia A. and Economou, Ioannis G. (2011) Using molecular simulation to predict solute solvation and partition coefficients in solvents of different polarity. Physical Chemistry Chemical Physics, 13 (20). pp. 9155-9164. ISSN 1463-9076 , <http://dx.doi.org/10.1039/c1cp20110g>

This version is available at <http://strathprints.strath.ac.uk/42567/>

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<http://strathprints.strath.ac.uk/>) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to Strathprints administrator: strathprints@strath.ac.uk

**Using Molecular Simulation to Predict Solute Solvation and Partition Coefficients in
Solvents of Different Polarity**

Nuno M. Garrido^{1,2}, Miguel Jorge¹, António J. Queimada¹,

Eugénia A. Macedo¹ and Ioannis G. Economou^{2*}

1. LSRE Laboratory of Separation and Reaction Engineering, Departamento de Engenharia Química, Faculdade de Engenharia, Universidade do Porto, Rua do Dr. Roberto Frias, 4200 - 465 Porto, Portugal

2. The Petroleum Institute, Department of Chemical Engineering, PO Box 2533, Abu Dhabi, United Arab Emirates

*Author to whom all correspondence should be addressed at: ieconomou@pi.ac.ae

Revised manuscript for publication to *Physical Chemistry Chemical Physics*

March 2011

Abstract

A methodology is proposed for the prediction of the Gibbs energy of solvation ($\Delta_{Solv}G$) based on MD simulations. The methodology is then used to predict $\Delta_{Solv}G$ of four solutes (namely propane, benzene, ethanol and acetone) in several solvents of different polarities, including *n*-hexane, *n*-hexadecane, ethylbenzene, 1-octanol, acetone and water, while testing the validity of the TraPPE force field parameters. Excellent agreement with experimental data is obtained, with average deviations of 0.2, 1.1, 0.75 and 1.2 kJ/mol, for the four solutes respectively. Subsequently, partition coefficients for forty different solute/solvent systems are predicted. The *a priori* knowledge of partition coefficient values is of high importance in chemical and pharmaceutical separation process design or as a measure of the increasingly important environmental fate. Here again, the agreement between experimental data and simulation predictions was excellent. An absolute average deviation of 0.28 log P units was obtained. However, this deviation can be decreased down to 0.14 log P units, just by optimizing partial atomic charges of acetone in the water phase, which proves molecular simulation to be a promising tool with strong physical basis to predict log P with competitive accuracy when compared to the popular statistics based methods with weak physical basis.

1. Introduction

The Gibbs energy of solvation is a vital thermodynamic quantity used to describe and fully understand the behavior of complex systems. Free energy calculations in molecular systems appeared as a predictive tool to express how species may associate or react¹. Over the last decade, significant progress has been made regarding the prediction of the Gibbs energy of hydration of small molecules^{2,3}, amino acid side-chains⁴⁻⁶ or even drug-like compounds^{7,8}, as well as the prediction of ligand-protein binding affinities⁹, the prediction of Gibbs energies for the characterization of reaction pathways and chemical equilibrium¹⁰ or to larger databases using continuum solvent approaches¹¹.

The overwhelming majority of experimental and theoretical studies concerning Gibbs energies of solvation deals with aqueous systems – i.e., by reporting the Gibbs energy of hydration¹². However, many important biological processes involve the interaction of solutes with both hydrophobic and hydrophilic environments (e.g., cell-membrane transfer, protein-ligand binding, etc.). In this context, it becomes increasingly important to understand how compounds partition between different fluid phases^{13,14}. Partition coefficients (henceforth referred to as $\log P$) reflect the solute hydrophobicity, partitioning in different solvents and pharmacokinetic characteristics. They are widely used for several applications in the pharmaceutical (e.g. drug design) or chemical industries (e.g. separation technology), and as a measure of the increasingly important environmental fate¹⁵. The prediction of partition coefficients is done mainly based on Quantitative Structure–Property Relationship (QSPR) algorithms¹⁶. QSPR methods are, to a large extent, statistically-based techniques. They are considerably fast and applicable to large databases of molecular structures but they also require the input of large multi-parameter tables. More importantly, they often lose predictive ability for compounds that are significantly different from their training set, and have thus been the object of criticism¹⁷. Alternative approaches may involve linear response theory based simulations¹⁸ or continuum free energy perturbation based simulations¹⁹⁻²².

A different approach to the previously mentioned tools is to use the absolute Gibbs energies of solvation calculated in the different phases to estimate the partition coefficient. From the Gibbs energies of solvation ($\Delta_{solv}G$) in two different phases (here generically named a and b) at constant temperature (T) and pressure (p), one can readily calculate the corresponding partition coefficient, $\log P^{a/b}$, according to the following expression:

$$\log P^{a/b} = \frac{\Delta_{solv}G^a - \Delta_{solv}G^b}{2.303RT} \quad (1)$$

In principle, both free energies of solvation that appear in equation (1) can be computed directly using molecular simulation methods. Initially, this approach was used to predict the water/CCl₄ partition coefficients of small alcohols²³ and it has previously been applied only to a few specific cases, e.g., 1-octanol/water partition of *n*-alkanes¹⁴, 1-ethylpropylamine²⁴ and 3-methyl-1-pentanol²⁴, chloroform/water²⁵ and *n*-octane/water²⁶ partition coefficients.

Equation (1) requires knowledge of the Gibbs energy of solvation of the solute in both solvents of interest (normally water and an organic solvent). As mentioned previously, there are very few studies that calculate solvation energies by molecular simulation in solvents other than water. Gonçalves and Stassen²⁷ computed $\Delta_{solv}G$ in low-dielectric constant solvents, including carbon tetrachloride, chloroform and benzene, using the GROMOS force field. Average absolute deviations (AAD) to experiments of 2.1 kJ/mol and 3.3 kJ/mol were reported for the first two solvents (23 solutes) and for benzene (17 solutes), respectively. Oostenbrink et al.²⁸ proposed a new parameterization of GROMOS based on simulations of the Gibbs energy of solvation in both water and cyclohexane. The authors concluded that it was not possible to adequately reproduce both data sets using a single set of parameters, and proposed the use of two different sets of point charges for polar and non-polar solvents. Chang et al.²⁹ also carried out calculations of solvation energies of 15 neutral amino acids and side-chain analogues in cyclohexane, but used the OPLS-AA model to represent the solutes and the solvent. Deviations of 2.3 kJ/mol were reported. Other studies reported Gibbs energies of solvation in 1-octanol, either pure¹⁴ or saturated with water^{30,31}. We are only aware of one previous study where the authors calculated solvation Gibbs energies in a range of solvents of different polarity. Geerke and van Gunsteren³² computed solvation energies of several solutes (most of them were amino-acid side chains) in methanol (14 solutes), dimethyl sulfoxide (12 solutes), acetonitrile (7 solutes) and acetone (11 solutes), using version 53A6 of the GROMOS force field²⁸ and relatively short simulation runs. They reported AAD ranging from 3.3 to 7.5 kJ/mol, depending on the solvent. Solvation in acetone was found to be particularly problematic and deviations to experiments were explained in terms of the properties of the two solvent models tested.

Although most experimental studies deal with aqueous systems, a large database of experimental measurements of $\Delta_{solv}G$ in several non-aqueous solvents appeared recently in the literature³³. Taking advantage of the availability of these data for comparison purposes, we report here the Gibbs energy of solvation of different compounds in diverse environments, calculated using molecular dynamics simulations (MD) with the TraPPE³⁴ force field. In order to cover a wide spectrum of polarity, we have decided to study the solvation of four chemically different solutes: propane (non-polar solute), benzene (aromatic solute), ethanol (protic polar solute) and acetone (aprotic dipolar solute), in different media: *n*-hexane or *n*-hexadecane (non-polar solvents), acetone (aprotic dipolar solvent), ethylbenzene (quadrupolar and aromatic

solvent), 1-octanol and water (hydrogen bonding solvents). We will then use these results to estimate the partition coefficients in different solvent combinations, comparing them to available experimental data. This will contribute to validate our methodology for general use in chemical and pharmaceutical applications. Apart from the practical interest of such calculations, our results are also useful for force field development purposes. Typically, force fields are developed based on pure component data, hence leaving out the behavior of molecules in heterogeneous/multicomponent environments. In this paper, we will test for the first time the Gibbs energy predictive capability of the TraPPE force field in a wide range of environments, identifying success and failure cases, and pointing out particular classes of compounds that could benefit from a refinement of the force field parameters. Contrary to the most recent versions of the GROMOS force field²⁸, no solvation data was included during the parameterization of TraPPE³⁴. Moreover, the accuracy of the force field for the prediction of Gibbs energies has not yet been thoroughly evaluated, and the calculations presented here represent an important step in this direction.

The remainder of this paper is organized as follows: in Section 2 we describe the computational methods used, provide MD simulation details, and briefly present the force field parameters used in this work. In Section 3.a, we evaluate the performance of the TraPPE force field for pure component physical properties, while the capability of the force field to predict absolute solvation energies and partition coefficients is discussed in sections 3.b and 3.c, respectively. The main conclusions are summarized in Section 4.

2. Computational Methods

a. Thermodynamic Integration

The solvation process consists of the transfer of a solute molecule from a well-defined gas (or vacuum) state into solution. The Gibbs energy of solvation can be experimentally measured from equilibrium concentration measurements in different phases (see reference³⁵ for further details). In statistical mechanics terms, it can be expressed as the total reversible work associated with changing the Hamiltonian of the system from the gas to the liquid state. Computationally, $\Delta_{solv}G$, being a state function, can be calculated from a thermodynamic cycle in which solute-solvent interactions are progressively turned off³⁶. The Gibbs energy associated with this mutation is estimated here using MD simulations and the Thermodynamic Integration (TI) method.

The TI algorithm, free of hysteresis, has been described in detail in previous publications³⁷⁻³⁹. **In short, if we consider any two well-defined states, the classical Hamiltonian**

of the system (the sum of kinetic and potential energies of the system) can be made a function of a coupling parameter, λ , and used to describe the transition between those two states; in this case, an initial fully interacting system and a final “dummy” state where all solute-solvent interactions have been turned off. The derivative of the Hamiltonian with respect to the coupling parameter is derived analytically, and equilibrium averages are used to evaluate these derivatives at several discrete and independent λ values. One then integrates these derivatives along a continuous path connecting the initial and final states to obtain the free energy between them.

In our protocol, in order to avoid charge-fusion effects^{5,40} we have first turned off the electrostatic interactions and then the Lennard-Jones (LJ) interactions. During the decoupling process, the electrostatic interactions, that are a function of λ , were linearly interpolated between neighboring states while the LJ interactions were interpolated via soft-core interactions⁴¹, with parameters given in our previous publication⁷. The electrostatic term represents the free energy cost of charging/discharging the solute, while the LJ component is the sum of an unfavorable cavity formation term and a favorable dispersion contribution. The total solvation Gibbs energy can then be estimated from:

$$\Delta_{Solv} G = \Delta_{Solv} G^{Elec} + \Delta_{Solv} G^{LJ} = - \left[\int_0^1 \left\langle \frac{\partial \mathcal{H}(\lambda)}{\partial \lambda} \right\rangle_{\lambda}^{Elec} d\lambda + \int_0^1 \left\langle \frac{\partial \mathcal{H}(\lambda)}{\partial \lambda} \right\rangle_{\lambda}^{LJ} d\lambda \right] \quad (2)$$

In this work, a MD simulation was performed for each discrete λ value, ranging from $\lambda = 0$ (fully interacting solute) to $\lambda = 1$ (non-interacting solute). We have used 5 λ values for the electrostatic decoupling and 15 λ points to evaluate the LJ contribution. As demonstrated in a previous work, this choice of intermediate states ensures sufficient accuracy in the free energy calculations⁴². The integration of the LJ Hamiltonian derivatives was carried out by fitting the data to a physically-based approximation to the cavity formation and dispersion interaction terms, and then integrating the curve analytically⁴². The fitting function is:

$$\left(\frac{\partial \mathcal{H}(\lambda)}{\partial \lambda} \right)_{LJ} = A_0 \lambda^2 + A_1 \lambda - \frac{A_2}{\lambda^2 - A_3 \lambda + A_4} + \frac{A_2}{A_4} \quad (3)$$

where A_i are fitting parameters. In all cases, we have employed non-linear weighted least square fittings of the simulation data. For the integration of the electrostatic component, Simpson’s rule was used.

b. Molecular Dynamics Simulations

MD simulations were conducted both for the calculation of pure solvent physical properties (Section 3.a) and for computing Gibbs energies of solvation (Section 3.b). All simulations were carried out using the GROMACS 4.0.7 suite⁴³, using the leap-frog Verlet integration algorithm⁴⁴ with a time step of 2 fs to integrate Newton's equations of motion. A new feature in this version of GROMACS⁴³ allows for the intermolecular solute-solvent interactions to be turned off independently from the intramolecular interactions. This means that the solvation energy can be quantified by running only a liquid phase simulation, avoiding the need for a vacuum step, as discussed in our previous publication⁴⁵.

Simulations were performed using periodic boundary conditions in all directions. In order to improve efficiency (see reference⁷ for details), the reaction-field method⁴⁶, which approximates the medium beyond a cut-off distance of 1 nm by a dielectric continuum of uniform permittivity, was used to handle long-range electrostatics. The dielectric constant was adjusted to the permittivity of each solvent. The remaining cut-off radii were 1 nm for the short-range neighbor list and 0.8-0.9 nm switched cut-off for the LJ interactions. We have also applied long range dispersion corrections for energy and pressure⁵.

For the estimation of the pure component properties we have employed the Nosé-Hoover thermostat^{47,48} for temperature coupling using a time constant of 1 ps, and the Parrinello-Rahman barostat⁴⁹ with a time constant of 2 ps to enforce pressure coupling. The reference pressure was set to 1 bar and the compressibility was set according to the fluid under evaluation (data from Cibulka and Takagi⁵⁰). Each simulation box contained 512 molecules. Liquid simulations were conducted for 10 ns of production time, while vacuum simulations (necessary for calculating the enthalpies of vaporization) were conducted without cutoffs or periodic boundary conditions for 50 ns.

In the case of Gibbs energy calculations, solvated systems consisted of one solute molecule and 265 molecules in the case of 1-octanol, 286 for *n*-hexane, 296 for acetone, 181 for ethylbenzene, 214 for *n*-hexadecane, and 512 for water. The TraPPE^{34,51-55} force field was used to represent all the molecules except water, which was represented by the Modified Extended Simple Point Charge (MSPC/E)⁵⁶ model. This model was developed as an improvement over the popular SPC/E⁵⁷ water potential and is an accurate force field for pure water and aqueous phase equilibria thermodynamic properties. It also includes a polarization correction expected to improve solvation energy predictions. A few comparisons of the Gibbs energy results predicted with the MSPC/E against other water models were reported previously^{45,58}. All calculations were performed at 298 K and 1 bar. Langevin stochastic dynamics⁵⁹ was used to control the temperature, with a frictional constant of 1 ps⁻¹ while for constant pressure runs the Berendsen barostat⁶⁰, with a time constant of 0.5 ps and an isothermal compressibility of 4.5×10^{-5} bar⁻¹,

was used to enforce pressure coupling. For each simulation, we first run an energy minimization (using the Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm⁶¹ for 5000 steps followed by a steepest descent minimization for another 500 steps) and afterwards a quick constant volume equilibration (100 ps), a constant pressure equilibration (1 ns), enough to obtain complete equilibration of the box volume, and finally a 5 ns NpT production stage. All simulations were performed in a threefold replica in order to obtain statistically meaningful results.

c. The TraPPE Force-Field

The first development of Transferrable Potentials for Phase Equilibria (TraPPE) force field appeared in the late 1990's with the purpose of creating an accurate, yet simple and transferable model to reproduce several phase equilibrium properties. In the general framework of the TraPPE force field, non-bonded interactions are described by pairwise-additive LJ 12-6 potentials and Coulombic interactions of partial charges. In the present work we have used a united-atom (UA) description of alkyl CH_x groups^{34,52} and an all-atom (AA) description for the polar groups: OH group in 1-octanol⁵¹ and the CO group in acetone⁵⁴. For aromatic groups, we have tested both a UA model⁵⁵ and an AA or explicit-hydrogen (EH) approach⁵³. For the EH model of benzene, additionally to the original TraPPE parameters we have also tested a recent reparameterization of its partial atomic charges, designed to yield better agreement with experimental Gibbs energies of hydration⁴⁵. Lorentz-Berthelot combining rules were used to determine the parameters for unlike LJ interactions. All bond lengths were treated as rigid and, with the exception of the fully rigid aromatic rings, all molecular models included harmonic angle bending and an appropriate torsional potential.

3. Results and Discussion

a. Pure Component Physical Properties

The accuracy of the TraPPE force field for the prediction of pure fluid properties of all compounds was initially evaluated. The calculated densities and heats of vaporization from NpT MD at 298 K are shown in Table 1. Densities were directly obtained from the GROMACS suite using the `g_energy` tool, while heats of vaporization were estimated by taking the difference of enthalpy in the vapor and liquid phases:

$$\Delta_{\text{vap}}H = E_g - E_l + RT \quad (4)$$

where, E_g and E_l are the total energy in the gas phase and in the liquid phase, respectively.

Based on the data reported in Table 1, one can conclude that the TraPPE force field provides excellent accuracy for liquid densities but generally underestimates the enthalpy of

vaporization, especially in the case of the *n*-alkanes. However, it should be noted that enthalpy data was not included in the model parameterization for such non-polar species. In the case of ethylbenzene, the enthalpy of vaporization can be considerably improved if an explicit hydrogen description of the aromatic ring is considered. Our results for the density and heat of vaporization of alcohols and for the density of acetone are in very good agreement with previously published data using Configurational Bias Monte Carlo simulations and the TraPPE force field.^{51,54} This agreement provides an appropriate validation of our simulation protocol.

b. Gibbs Energies of Solvation

Initially, the Gibbs energies of solvation of propane in *n*-hexadecane, ethylbenzene, 1-octanol, acetone and water were examined. Since propane has no partial atomic charges³⁴ the Coulombic contribution to the total solvation energy is zero. Consequently, the Gibbs energy of solvation becomes equal to the LJ contribution (see equation 2). A typical plot of the Hamiltonian derivative as a function of time during the production stage of MD simulations for the case of benzene in *n*-hexane for different λ values is shown in Figure 1. As we can observe, for all values of the coupling parameter, $\langle \partial\mathcal{H}/\partial\lambda \rangle$ converges to its equilibrium value after a short period of time. Then, one can take the average value of the Hamiltonian derivative and plot it as a function of the coupling parameter, as represented in Figure 2, for the case of propane in different solvents (complementary plots are given as supporting information for the remaining solutes in Figures S1 to S3). Integration of this curve according to equation (2) yields the Gibbs energy of solvation.

The shape of the curves in Figure 2 reflects the opposite contributions of the cavity formation energy (at higher λ values) and the solute-solvent dispersion interactions (at lower λ values). It is also quite interesting to see that the water curve is considerably different from the other solvents of lower polarity, which will be reflected in the substantially different values predicted for the Gibbs energy of solvation. The unfavorable cavity contribution is much larger than in the other solvents, due to the difficulty in breaking the water hydrogen-bonded network to make room for the solute, while the favorable dispersion contribution is smaller in water than in the non-aqueous solvents⁶². This observation is also verified for the other solutes (Figures S1-S3).

Results for the predicted solvation energies using equation (2) together with experimental values are presented in Table 2. As it can be seen, an excellent agreement (AAD equal to 0.23 kJ/mol) between simulations and experiments is obtained. It is interesting, although not necessarily unexpected, that the best results are obtained in the solvent that is most chemically similar to propane (i.e., *n*-hexadecane). The model performs well for all other solvents, even for water, where the solvation of propane is highly unfavorable. This indicates that the TraPPE LJ parameters are able to adequately describe the Gibbs energy of solvation in a wide range of

environments. Since propane has no partial charges and consequently the electrostatic term is zero, we have also performed a simulation in which ethylbenzene is represented by the UA TraPPE force field⁵⁵, without partial atomic charges, where consequently the electrostatic term becomes zero. As we can see from Table 2, the prediction is also quite satisfactory and at a much lower computational cost. This suggests that UA models can be adequate representations of aromatic molecules in systems that are dominated by repulsion and dispersion interactions.

The next solute examined, following a scale of increasing polarity, was benzene. We have concluded in a previous publication⁴⁵ that in order to predict correct Gibbs energies of hydration, an EH description of benzene is required, and partial atomic charges need to be assigned to the different aromatic atoms. This is because for aromatic molecules the electrostatic component cannot be neglected in the calculation of the Gibbs energy of hydration. For the description of benzene carbon and hydrogen atoms, we have decided to keep the original LJ TraPPE parameters and tested 2 different charge sets: 1) the original TraPPE parameters ($q_C = -q_H = -0.095$); 2) a modified set of charges designed to accurately reproduce the benzene Gibbs energy of hydration⁴⁵ ($q_C = -q_H = -0.1225$).

Results for the averaged LJ term, the different electrostatic contributions and the respective total solvation energies for the two studied charge sets are presented in Table 3. In general, for both models tested the electrostatic component of the free energy increases in magnitude, becoming more negative as the polarity of the solvent increases. Nevertheless, the Gibbs energy of solvation of benzene in the non-aqueous solvents is still dominated by the LJ term, which means that the influence of the solute atomic partial charges on the Gibbs energy is relatively small. This helps explain why we are able to obtain a very good prediction for the Gibbs energy in ethylbenzene using a computationally much cheaper UA model for the solvent. In fact, the UA prediction is even closer to the experimental value than the EH approach. This may seem somewhat surprising, but it can be explained by the fact that the original TraPPE parameterization of the EH model focused only on benzene and heterocyclic aromatic compounds, and was not optimized for pure-component substituted alkylbenzenes. In our recent extension of the TraPPE-EH model to alkyl-aromatic molecules⁴⁵, we have simply adopted the LJ parameters from the UA model of alkanes³⁴ for the alkyl substituent groups, and focused only on optimizing the point charges to accurately describe the Gibbs energy of hydration. It is possible that a reparameterization of the LJ parameters for alkyl-aromatic molecules, following the TraPPE protocol, will yield improved results for the Gibbs energy of solvation in these solvents.

The electrostatic component of the Gibbs energy of solvation of benzene in water is much larger than for the other solvents – it is at least of the same magnitude as the LJ component. This means that an adequate parameterization of the solute point charges becomes extremely

important in aqueous media. From Table 3, one can see that the original TraPPE charges give good predictions for all the non-aqueous solvents, but fail for water. These charges were parameterized in a non-polar environment (pure benzene), and thus should be expected to perform well in similar environments. It is therefore not surprising that the best results are obtained for benzene solvation in ethylbenzene. Conversely, the modified TraPPE charges, optimized for accurate Gibbs energies of hydration of alkyl-aromatics, clearly improve the prediction for the water phase, but at the cost of a slightly poorer prediction in the remaining media (except of course in the cases where the electrostatic contribution is zero). It was shown previously⁴⁵ that it was necessary to rescale the point charges of alkyl-aromatic solutes when transferring them from a non-polar to a highly polar (i.e. water) solvent, so as to account for the different polarization environments surrounding the molecule. It is thus to be expected that the consequent improvement in the description of hydration causes a worse behavior in less polar solvents.

Generally, solvation Gibbs energies are in very good agreement with experimental values and an AAD of 1.1 kJ/mol was obtained with the previously optimized parameters for benzene⁴⁵. A slightly larger deviation is observed when using the original charge values (AAD of 1.6 kJ/mol), but this is caused mostly by the poor performance of the benzene model in water, as explained above. These results strongly suggest that the assignment of point charges, even for weakly polar molecules, should depend on the environment in which they are placed. A similar conclusion was reached by Oostenbrink et al.²⁸, leading to the proposal of two different GROMOS charge sets, one for polar and one for non-polar solvents.

We have also studied the solvation of a highly polar solute, ethanol, in the different solvents. Results for both the LJ and electrostatic contributions of ethanol solvation are given in Table 4. Contrary to the two above mentioned solutes, in highly polar solutes, Coulombic interactions play a very important role⁶³. Indeed, for all solvents except *n*-hexane, the electrostatic component is higher in magnitude than the LJ component. Our results for the solvation of ethanol in ethylbenzene confirm this analysis – in order to correctly capture the electrostatic contribution, an explicit hydrogen solvent description must be considered. Thus, for ethylbenzene-EH we have used the parameters given in Garrido et. al⁴⁵, which give very good results for the Gibbs energy of solvation of ethanol.

As in the previous cases, the best prediction is obtained for the solvent that is most chemically similar to the solute – 1-octanol – but the results are also quite good for ethanol in water. Because the TraPPE force field was parameterized in pure ethanol⁵¹, it is able to capture correctly the hydrogen bond interactions that are dominant in those systems. For ethanol in acetone, the results are not as good, even though this solvent is of intermediate polarity between 1-octanol and water. This may arise from an inability of the acetone parameters in describing hydrogen bond interactions, as discussed in more detail below. The failure of the model in the

case of ethanol in *n*-hexane was quite unexpected, as this is a completely non-polar solvent where only the LJ component is important. One possible explanation would be an inadequacy of the Lorentz-Berthelot combining rules to describe the cross-interactions between alcohols and alkanes, but this is unlikely, on its own, to justify such a large discrepancy. A more plausible justification is an inaccurate experimental value. This possibility will be discussed in more detail in section 3c. Even with the large error for the *n*-hexane solvent, a promising 1.5 kJ/mol overall average deviation was observed for the full data set.

Finally, we have studied a dipolar solute, acetone, in different media. Table 5 reports the results obtained, showing very good predictions for the different solvents, except water. The good results obtained for acetone solvation in *n*-hexane, where only LJ interactions are present, suggest a correct parameterization of the LJ term for acetone. In a previous solvation study³², acetone was identified as a problematic solvent, with relatively large errors (AAD of more than 5 kJ/mol) obtained for the Gibbs energies of solvation of apolar solutes in acetone using both the OPLS⁶⁴ and the Weerasinghe and Smith⁶⁵ models. Geerke and van Gunsteren³² suggested that this could be solved by the use of more attractive van der Waals interactions to reproduce acetone bulk properties in simulations. In the present work, the acetone LJ parameters can be assessed by directly comparing simulated and experimental Gibbs energies of solvation for the cases of propane in acetone (Table 2) and acetone in *n*-hexane (Table 5), where only LJ interactions between solute and solvent are present. In both cases, good agreement is obtained, which may indicate that TraPPE parameters give a better description of such LJ interactions.

Results for the water phase using the original TraPPE parameters are considerably off the experimental values. Given that the TraPPE LJ parameters for acetone seem to be sufficiently accurate (they correctly describe pure acetone properties (ref.⁵⁴ and Table 1) as well as Gibbs energies of solvation in non-polar environments), it is likely that the discrepancy is caused by an incorrect parameterization of the point charges. Following the same rationale as in our previous work⁴⁵, we have attempted to modify the atomic partial charges assigned to the carbonyl carbon and oxygen atoms in acetone to reproduce the Gibbs energy of hydration. An optimal parameterization was found when $q_C = -q_O = -0.545$, which exactly fits the experimental solvation Gibbs energy. This represents a relatively small increment in magnitude over the original TraPPE force field partial charges of $q_C = -q_O = -0.424$. This change in the point charges reflects the change of polarization environment when moving from pure acetone to water, analogously to what was shown in our previous study on alkyl-aromatic solutes⁴⁵.

However, there is a further aspect that is worth considering in the case of acetone. Apart from water, the worst results for the solvation of acetone are obtained in 1-octanol. Similarly, ethanol solvation in acetone is also poorly predicted by the TraPPE model (Table 4). This points to a deficiency in the description of interactions between acetone (or perhaps ketones in general)

and molecules that are hydrogen-bond donors, such as water or alcohols. Indeed, hydrogen bonds are not formed in pure acetone, and thus a correct description of these interactions is not required to model pure-component properties or interactions with other molecules that are not hydrogen donors (e.g., alkanes or aromatics). However, when acetone is in contact with water or alcohols, association becomes important and the force field fails (dramatically in the case of water). In contrast, when molecules that form hydrogen bonds in the pure fluid are mixed (e.g., different alcohols or water and alcohols), the force field performs well, because association is already accounted for in the pure-component parameterization. Similar effects are observed when mixtures of protic and aprotic molecules are described by molecular based models, such as excess Gibbs energy models and equations of state, that do not explicitly account for association.

Finally, an overall comparison of the predicted Gibbs energies of solvation against experimental data is summarized in Figure 3. The data shown in this figure should be seen as our “best predictions”, i.e. using the EH model for benzene and ethylbenzene, and the optimized point charges for benzene and acetone hydration. As we can see, the agreement is excellent for the entire range of solutes and solvents. These results suggest that good predictions of free energies of solvation can be achieved using simple fixed-charge molecular models, provided that the point charges are parameterized accounting for changes in the polarization environment as the solute is moved from non-polar to polar solvents. **This approach is fundamentally different from others presented in literature where liquid polarization effects can be derived from e.g. ab initio⁶⁶ or quantum mechanics/molecular mechanics calculations⁶⁷.** Although our results are encouraging, a more systematic examination of these effects is necessary.

c. Prediction of Partition Coefficients

The partition coefficients for propane, benzene, ethanol and acetone at 298 K can be readily estimated from equation (1) using the Gibbs energies of solvation calculated from the MD simulations. Results are shown for the different solutes in Table 6. Partition coefficients can also be experimentally measured via the shake-flask technique for generating saturated liquid phases followed by sampling and quantitative solute analysis (details are given in the review of Valkó⁶⁸). However, the availability of such experimental data is scarce and limited to the environmentally important 1-octanol/water⁶⁹ and *n*-hexane/water^{70,71} partition coefficients. Such values were included in Table 6 for comparison. Very good agreement is obtained for all solutes except for acetone, but the latter are improved by applying the modified set of point charges for aqueous media. It is also important to notice that although the simulation prediction of the ethanol Gibbs energy of solvation in *n*-hexane is quite far from the experimental value (see Table 4), the corresponding *n*-hexane/water partition coefficient is correctly predicted (see Table 6). Furthermore, the TraPPE force field gives good predictions of ethanol Gibbs energy of

solvation in water. Taken together, these results may indicate a possible error in the experimental value for the Gibbs energy of solvation of ethanol in *n*-hexane. Further experiments are needed to clarify this point.

One way to estimate the experimental partition coefficients when direct measurements are not available is to use experimental solvation energies in both phases as inputs to equation (1). We call these “estimated” experimental partition coefficients. A general overview of predicted simulation data against “estimated” experimental values is given in Figure 4. Small deviations, of 0.05, 0.16, 0.30 and 0.13 log P units were obtained for propane, benzene, ethanol and acetone, respectively. Moreover, the important 1-octanol/water partition coefficient is predicted very accurately for all the solutes.

In terms of accuracy comparison, and although applied to larger solute databases, deviations of 0.46 and 0.43 log P units for 1-octanol/water and *n*-hexane/water partition, respectively, can be found in the literature using an empirical thermodynamic approach⁷². Multiple regression techniques applied on descriptors based on molecular properties and molecular partial charges can produce deviations of 0.65 log P units for 1-octanol/water partition coefficient⁷³. For the same property, a combination of a group contribution approach together with several correction factors produces deviations of 0.4 log P units⁷⁴. Neural network techniques applied to the prediction of log P from non-linear regression of structure analysis using 10 to 18 descriptors produced deviations ranging from 0.3 to 0.4 log P units⁷⁵. Finally, a generic test of quantum mechanical and QSPR methodologies produced average deviations of 0.6 and a maximum of 2.15 log P units⁷⁶. This comparison suggests that molecular simulation methods can be competitive in terms of accuracy to other techniques that are more commonly employed to predict partition coefficients.

4. Conclusions

The focus of this work has been to predict partition coefficients of 4 solutes in several combinations of 6 different solvents, through the calculation of the absolute Gibbs energies of solvation using molecular dynamics simulations and thermodynamic integration. The TraPPE force field was used to account for the intra- and intermolecular interactions. The diversity of solutes and solvents considered in this work allowed us to assess the performance of the force field for Gibbs energy predictions in a wide range of environments. By comparing simulated and experimental data, we were thus able to identify several strong and weak points of the TraPPE force field, and suggest possible avenues for future model refinement. The promising results obtained for the solvation of propane (a non-polar solute) in all solvents and of different solutes in *n*-hexane (a non-polar solvent) indicate that the Lennard-Jones parameters of TraPPE are generally able to describe the dispersive and repulsive (cavity formation) contributions to the Gibbs energy of solvation. This is an important conclusion, because a correct description of

the free energy of solvation relies on a delicate balance between enthalpic and entropic contributions, and this is difficult to accurately capture using a simple molecular model. The only relevant exception was the Gibbs energy of solvation of ethanol in *n*-hexane, where a large deviation from the experimental value was observed. This may be due to a shortcoming of the molecular model, but also to inaccuracies in the experimental measurements. The fact that both the water/*n*-hexane partition coefficient and the Gibbs energy of solvation of ethanol in water obtained from the simulations are in excellent agreement with experiment suggests that the latter may be the most likely explanation. Further experimental measurements and simulations are necessary to clarify this issue.

Our results for solvation in ethylbenzene reinforce an earlier conclusion⁴⁵ that an explicit hydrogen description, together with an appropriate assignment of point charges, is necessary for an adequate description of Gibbs energies of solvation involving aromatic solutes and/or solvents. Nevertheless, in the particular cases of solvation of non-polar solutes in aromatic solvents or of aromatic solutes in non-polar solutes, a united-atom description of the aromatic molecules is usually sufficient to provide satisfactory free energy predictions, at a much lower computational cost. This is because in these cases the electrostatic component is zero, and a carefully parameterized united-atom model is capable of describing the LJ contribution to the Gibbs energy of solvation.

In general, the most accurate results were obtained for the solute/solvent pairs with chemically similar molecules (e.g., propane in *n*-hexadecane, benzene in ethylbenzene, ethanol in 1-octanol, and even acetone in acetone). This is to be expected, considering that force fields (and TraPPE is no exception) are normally parameterized based on pure-component data, and thus perform better for systems that involve the same type of interactions as in the pure fluid. Conversely, the worst performances were observed for systems that include interactions that were not present in the pure-fluid parameterization. A particularly problematic case for model development occurs for mixtures of protic and aprotic polar molecules – e.g., acetone in water or 1-octanol, and ethanol in acetone. In pure acetone, intermolecular hydrogen bonds do not play an important role, and thus the parameters are not necessarily designed to properly account for these interactions. However, when acetone is mixed with protic molecules (water or alcohols), hydrogen bonds are formed, and the model may be unable to correctly describe them. In the present paper, we demonstrate that a simple rescaling of the point charges on acetone is able to account for this discrepancy and accurately reproduce its free energy of hydration. Further tests are necessary to assess the general applicability of this approach. Nevertheless, our calculations confirm previous observations, by ourselves⁴⁵ and others²⁸, that the point charges assigned to a given molecule should reflect the environment in which it is placed. Clearly, a more physically realistic approach to address this issue is the use of polarizable force fields as discussed by others^{21,77,78}. However, they are significantly more computationally demanding and

thus, less attractive for the application presented here. Our results suggest that good predictions can be obtained using fixed-charge models, provided environment effects are taken into account in the parameterization of the point charges.

Our simulations are able to predict Gibbs energies of solvation that are in very good agreement with experimental data, deviating on average only 0.23 kJ/mol for propane, 1.10 kJ/mol for benzene, 0.29 kJ/mol in the case of acetone using solute partial charges rescaled to fit the solvation in water (original force field parameters predict deviations of 1.2 kJ/mol to experimental data), and 0.75 kJ/mol for ethanol. They should thus be considered as excellent predictions. Using the different simulation values of solvation free energies, we are able to readily estimate the corresponding partition coefficients. For the cases where independent experimental values are available (for water/1-octanol and water/*n*-hexane partition), excellent agreement was achieved. In total, 40 partition coefficients were predicted, deviating 0.14 log P units from experimental values measured directly or estimated from experimental solvation energy data using equation (1). It is important to notice that this value was obtained using optimized partial atomic charges for acetone in the water phase, as mentioned in Sections 3b. and 3c. The molecular simulation predictions reported here are, at least, comparable in accuracy to commonly used QSPR-based tools or group contribution methods. Contrary to other methodologies presented in the literature, we do not need the knowledge of the experimental solvation data in advance, which is a distinct advantage. Furthermore, our methodology can be generically extended to any two-solvent partitioning, which is a clear advantage when compared to the common log P estimation methods.

5. Acknowledgments

NMG acknowledges Petroleum Institute for a visiting PhD scholarship and his FCT PhD scholarship SFRH/BD/47822/2007 for financial support.

6. Cited References

- (1) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395.
- (2) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. *J. Phys. Chem. B* **2007**, *111*, 2242.
- (3) Shivakumar, D.; Deng, Y.; Roux, B. *J. Chem. Theory Comput.* **2009**, *5*, 919.
- (4) Wang, J.; Cieplak, P.; Kollman, P. *J. Comp. Chem.* **2000**, *21*, 1049.
- (5) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. *J. Chem. Phys.* **2003**, *119*, 5740.
- (6) Hess, B.; van der Vegt, N. F. A. *J. Phys. Chem. B* **2006**, *110*, 17616.
- (7) Garrido, N. M.; Jorge, M.; Queimada, A. J.; Economou, I. G.; Macedo, E. A. *Fluid Phase Equilib.* **2010**, *289*, 148.
- (8) Garrido, N. M.; Queimada, A. J.; Jorge, M.; Economou, I. G.; Macedo, E. A. *Fluid Phase Equilib.* **2010**, *296*, 110.
- (9) Gilson, M.; Zhou, H. *Ann. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21.
- (10) Jorgensen, W. L.; Thomas, L. L. *J. Chem. Theory Comput.* **2008**, *4*, 869.
- (11) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 1133.
- (12) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.
- (13) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Del. Rev.* **1997**, *23*, 3.
- (14) Garrido, N. M.; Queimada, A. J.; Jorge, M.; Macedo, E. A.; Economou, I. G. *J. Chem. Theory Comput.* **2009**, *5*, 2436.
- (15) Leo, A.; Hansch, C.; Elkins, D. *Chem. Rev.* **1971**, *71*, 525.
- (16) Karelson, M. *Molecular Descriptors in QSAR/QSPR*; Wiley - Interscience: New York, 2000.
- (17) Johnson, S. *J. Chem. Inf. Model.* **2008**, *48*, 25.
- (18) Duffy, E. M.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2000**, *122*, 2878.
- (19) Best, S. A.; Merz, K. M.; Reynolds, C. H. *J. Phys. Chem. B* **1999**, *103*, 714.
- (20) Orozco, M.; Luque, F. J. *Chem. Rev.* **2000**, *100*, 4187.
- (21) Giesen, D. J.; Chambers, C. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1997**, *101*, 5084.
- (22) Buggert, M.; Cadena, C.; Mokrushina, L.; Smirnova, I.; Maginn, E. J.; Arlt, W. *Chem. Eng. Techn.* **2009**, *32*, 977.
- (23) Essex, J. W.; Reynolds, C. A.; Richards, W. G. *J. Am. Chem. Soc.* **1992**, *114*, 3634.
- (24) Liu, Y.; Li, X. F.; Wang, L.; Sun, H. *Fluid Phase Equilib.* **2009**, *285*, 19.

- (25) Eksterowicz, J. E.; Miller, J. L.; Kollman, P. A. *J. Phys. Chem. B* **1997**, *101*, 10971.
- (26) Michel, J.; Orsi, M.; Essex, J. W. *J. Phys. Chem. B* **2008**, *112*, 657.
- (27) Goncalves, P. F. B.; Stassen, H. *J. Comput. Chem.* **2003**, *24*, 1758.
- (28) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. *J. Comp. Chem.* **2004**, *25*, 1656.
- (29) Chang, J.; Lenhoff, A. M.; Sandler, S. I. *J. Phys. Chem. B* **2007**, *111*, 2098.
- (30) Chen, B.; Siepmann, J. *J. Phys. Chem. B* **2006**, *110*, 3555.
- (31) Economou, I. G.; Garrido, N. M.; Makrodimitri, Z. A. *Fluid Phase Equilib.* **2010**, *296*, 125.
- (32) Geerke, D. P.; van Gunsteren, W. F. *ChemPhysChem* **2006**, *7*, 671.
- (33) Katritzky, A. R.; Oliferenko, A. A.; Oliferenko, P. V.; Petrukhin, R.; Tatham, D. B.; Maran, U.; Lomaka, A.; Acree Jr, A. E. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1794.
- (34) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1998**, *102*, 2569.
- (35) Abraham, M. *Chem. Soc. Rev.* **1993**, 73.
- (36) Leach, A. *Molecular Modeling: Principles and Applications*; Prentice-Hall, 2001.
- (37) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300.
- (38) C. Chipot; Pohorille, A. *Free Energy Calculations - Theory and Applications in Chemistry and Biology*; Springer: Berlin, 2007.
- (39) Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235.
- (40) Deng, Y.; Roux, B. *J. Phys. Chem. B* **2004**, *108*, 16567.
- (41) Beuler, T. M., R.; van Schaik, RC; Gerber, PR; van Gunsteren, WF *Chem. Phys. Lett.* **1994**, *222*, 529.
- (42) Jorge, M.; Garrido, N. M.; Queimada, A. J.; Economou, I. G.; Macedo, E. A. *J. Chem. Theory Comput.* **2010**, *6*, 1018.
- (43) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435.
- (44) van Gunsteren, W.; Berendsen, H. *Mol. Sim.* **1988**, *1*, 173.
- (45) Garrido, N. M.; Jorge, M.; Queimada, A. J.; Gomes, J.; Economou, I. G.; Macedo, E. *J. Phys. Chem. B* **2010**, *in Press*.
- (46) Lee, F. S.; Warshel, A. *J. Chem. Phys.* **1992**, *97*, 3100.
- (47) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.
- (48) Nose, S. *Molecular Physics* **1984**, *52*, 255.
- (49) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182.
- (50) Cibulka, I.; Takagi, T. *J. Chem. Eng. Data* **1999**, *44*, 411.
- (51) Chen, B.; Potoff, J. J.; Siepmann, J. I. *J. Phys. Chem. B* **2001**, *105*, 3093.

- (52) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1999**, *103*, 4508.
- (53) Rai, N.; Siepmann, J. I. *J. Phys. Chem. B* **2007**, *111*, 10790.
- (54) Stubbs, J. M.; Potoff, J. J.; Siepmann, J. I. *J. Phys. Chem. B* **2004**, *103*, 17596.
- (55) Wick, C. D.; Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **2000**, *104*, 8008.
- (56) Boulougouris, G. C.; Economou, I. G.; Theodorou, D. N. *J. Phys. Chem. B* **1998**, *102*, 1029.
- (57) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
- (58) Garrido, N.; Jorge, M.; Queimada, A.; Economou, I.; Macedo, E. *Fluid Phase Equilib.* **2010**, *289*, 148.
- (59) Van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1982**, *45*, 637.
- (60) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
- (61) Liu, D. C.; Nocedal, J. *Math. Programm.* **1989**, *45*, 503.
- (62) Chandler *Nature* **2005**, *437*, 640.
- (63) van Gunsteren, W. F.; Bakowies, D.; Baron, R.; Chandrasekhar, I.; Christen, M.; Daura, X.; Gee, P.; Geerke, D. P.; Glättli, A.; Hünenberger, P. H.; Kastenholz, M. A.; Oostenbrink, C.; Schenk, M.; Trzesniak, D.; van der Vegt, N. F. A.; Yu, H. B. *Angew. Chem. Int. Ed.* **2006**, *45*, 4064.
- (64) Jorgensen, W. L.; Briggs, J. M.; Contreras, M. L. *J. Phys. Chem.* **1980**, *94*, 1683.
- (65) Weerasinghe, S.; Smith, P. E. *J. Chem. Phys.* **2003**, *118*, 10663.
- (66) Bühl, M.; Chaumont, A.; Schurhammer, R.; Wipff, G. *J. Phys. Chem. B* **2005**, *109*, 18591.
- (67) Klahn, M.; Seduraman, A.; Wu, P. *J. Phys. Chem. B* **2008**, *112*, 10989.
- (68) Valko, K. *J. Chromat. A* **2004**, *1037*, 299.
- (69) Sangster, J. *Octanol-Water Partitioning Coefficients: Fundamentals and Physical Chemistry*; John Willey & Sons: Chichester, U.K., 1997.
- (70) Seiler, P. *Eur. J. Med. Chem.* **1999**, *34*, 473.
- (71) Schulte, J.; Durr, J.; Ritter, S.; Hauthal, W. G.; Quittsch, K.; Maurer, G. *J. Chem. Eng. Data* **1998**, *43*, 69.
- (72) Ruelle, P. *Chemosphere* **2000**, *40*, 457.
- (73) Xing, L.; Glen, R. C. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 796.
- (74) Klopman, G.; Li, J.-Y.; Wang, S.; Dimayuga, M. *J. Chem. Inf. Comput. Sci.* **2004**, *34*, 752.
- (75) Duprat, A. F.; Huynh, T.; Dreyfus, G. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 586.
- (76) Beck, B.; Breindl, A.; Clark, T. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1046.

- (77) Yu, H. B.; van Gunsteren, W. F. *Comput. Phys. Commun.* **2005**, *172*, 69.
- (78) Swope, W. C.; Horn, H. W.; Rice, J. E. *J. Phys. Chem. B* **2010**, *114*, 8631.

Table 1: Pure component densities and heats of vaporization at 1 bar and 298 K from experimental data and MD simulation.

Solvent	ρ (kg/m ³)			$\Delta_{vap}H$ (kJ/mol)		
	Experiment	Simulation	Absolute Deviation (%)	Experiment	Simulation	Absolute Deviation (%)
<i>n</i> -propane	491.3	485.9	1.1	14.8	11.7	20.9
<i>n</i> -hexane	656.2	657.8	0.2	31.6	26.5	16.1
<i>n</i> -hexadecane	769.0	779.0	1.3	79.7	72.0	9.7
ethanol	786.0	780.9	0.6	42.5	40.5	4.7
1-octanol	826.1	823.7	0.3	72.3	67.0	7.3
acetone	786.7	775.0	1.5	31.2	26.4	5.3
benzene ^{a,c}	857.9	877.6	2.3	33.0	33.0	0.0
ethylbenzene	863.7	861.7 ^b ; 857.5 ^c	0.2 ^b ; 0.7 ^c	42.2	35.5 ^b ; 40.9 ^c	16 ^b ; 3.1 ^c

^a data from ref⁴⁵ at 313 K^b using UA description of the aromatic ring^c using EH description of the aromatic ring

Table 2: Experimental and predicted values for the Gibbs energy of solvation of propane in different solvents. The subscripts give the statistical accuracy of the last decimal point shown.

Solvent	$\Delta_{Solv} G_{Exp}$ (kJ/mol)	$\Delta_{Solv} G_{Calc} = \Delta_{Solv} G^{LJ}$ (kJ/mol)	Absolute deviation (kJ/mol)
<i>n</i> -hexadecane	-6.0	-6.0 ₃	0.0
ethylbenzene - UA	-7.07	-6.82 ₉	0.25
ethylbenzene - EH		-7.19 ₁	0.12
1-octanol	-5.52	-5.12 ₇	0.40
acetone	-6.7	-6.12 ₅	0.58
water	8.5	8.6 ₄	0.10
<i>Overall</i>			0.23

Table 3: Experimental and predicted values for the Gibbs energy of solvation (in kJ/mol) of benzene in different solvents. The subscripts give the statistical accuracy of the last decimal point shown.

Solvent	$\Delta_{Solv} G_{Exp}$	$\Delta_{Solv} G_{calc}^{LJ}$	Original set of charges			Modified set of charges		
			$\Delta_{Solv} G_{calc}^{Coul}$	$\Delta_{Solv} G_{calc}$	Absolute Deviation	$\Delta_{Solv} G_{calc}^{Coul}$	$\Delta_{Solv} G_{calc}$	Absolute Deviation
<i>n</i> -hexane	-16.6	-18.2 ₂	0	-18.2 ₂	1.6	0	-18.2 ₂	1.6
ethylbenzene - UA	-18.1	-18.3 ₃	0	-18.3 ₃	0.2	0	-18.3 ₃	0.2
ethylbenzene - EH		-16.7 ₁	-0.64 ₆	-17.3 ₁	0.8	-0.53 ₆	-17.2 ₁	0.9
1-octanol	-15.9	-16.1 ₁	-0.69 ₂	-16.8 ₁	0.9	-1.11 ₂	-17.2 ₁	1.3
acetone	-18.3	-17.1 ₁	-2.15 ₂	-19.3 ₁	1.0	-3.46 ₇	-20.6 ₁	2.3
water ^a	-3.6	5.5 ₅	-5.04 ₉	0.5 ₅	4.1	-9.13 ₆	-3.6 ₅	0.0
<i>Overall</i>					0.9 ^b 1.6 ^c	<i>Overall</i>		1.4 ^b 1.1 ^c

^a From ref.⁴⁶

^b Evaluated for all the solvents except water

^c Considering the full data set

Table 4: Experimental and predicted Gibbs energy of solvation (kJ/mol) for ethanol in different solvents. The subscripts give the statistical accuracy of the last decimal point shown.

Solvent	$\Delta_{Solv} G_{Exp}$	$\Delta_{Solv} G_{calc}^{LJ}$	$\Delta_{Solv} G_{calc}^{Coul}$	$\Delta_{Solv} G_{calc}$	Absolute deviation (kJ/mol)
<i>n</i> -hexane	-10.1	-6.5 ₁	-0.06 ₁	-6.6 ₁	3.5
ethylbenzene - UA	-12.6	-5.6 ₂	0	-5.6 ₂	7.0
ethylbenzene - EH		-5.3 ₃	-6.33 ₁	-11.6 ₃	1.0
1-octanol	-18.4	-4.4 ₂	-14.1 ₂	-18.4 ₂	0.0
acetone	-18.4	-5.3 ₁	-11.10 ₂	-16.4 ₁	2.0
Water	-21.1	8.0 ₂	-28.33 ₁	-20.1 ₂	1.0
<i>Overall</i>					0.75

Table 5: Experimental and predicted Gibbs energy of solvation values (in kJ/mol) for acetone in different solvents using analytical integration for the LJ term. The subscripts give the statistical accuracy of the last decimal point shown.

Solvent	$\Delta_{Solv} G_{Exp}$	$\Delta_{Solv} G_{calc}^{LJ}$	$\Delta_{Solv} G_{calc}^{Coul}$	$\Delta_{Solv} G_{calc}$	Absolute deviation
<i>n</i> -hexane	-10.9	-11.46 ₈	0	-11.46 ₈	0.6
ethylbenzene - UA	-14.3	-11.4 ₁	0	-11.4 ₁	2.9
ethylbenzene-EH		-11.3 ₁	-2.38 ₂	-13.7 ₁	0.6
1-octanol	-13.2	-9.9 ₂	-2.29 ₃	-12.2 ₂	1.0
acetone	-17.3	-10.8 ₁	-5.90 ₅	-16.7 ₁	0.6
water	-15.9	6.45 ₆	-12.93 ₃ ^a	-6.48 ₆ ^a	9.4 ^a
			-22.40 ₅ ^b	-15.95 ₆ ^b	0.05 ^b
<i>Overall</i>					1.2 ^a , 0.3 ^b

^a Using the original acetone TraPPE parameters

^b Using optimized atomic partial charges for acetone in the water phase

Table 6: Predicted partition coefficient data from MD simulations. Available experimental data are given in parentheses.

Solvent pair	Solute			
	log P			
	propane	benzene	ethanol	acetone
1-octanol / water	2.40 (2.4)	2.38 (2.1)	-0.30 (-0.3)	1.00 ^b ; -0.66 ^c (-0.24)
<i>n</i> -alkane ^a / water	2.56	2.56 (2.1)	-2.37 (-2.3)	0.87 ^b ; -0.79 ^c (-0.9)
acetone / water	2.58	2.89	-0.65	1.79 ^b ; 0.13 ^c
ethylbenzene / water	2.70	2.58	-1.49	1.27 ^b ; -0.39 ^c
<i>n</i> -alkane ^a / 1-octanol	0.15	0.18	-2.07	-0.13
acetone / 1-octanol	0.18	0.51	-0.35	0.79
ethylbenzene / 1-octanol	0.30	0.19	-1.19	0.26
acetone / <i>n</i> -alkane ^a	0.02	0.33	1.72	0.92
ethylbenzene / <i>n</i> -alkane ^a	0.14	0.02	0.88	0.39
ethylbenzene / acetone	0.12	-0.32	-0.84	-0.53

^aalkane solvent is *n*-hexadecane for solvation of propane and *n*-hexane for all other solutes.

^bUsing the original acetone TraPPE parameters

^cUsing optimized atomic partial charges for acetone solvated in water

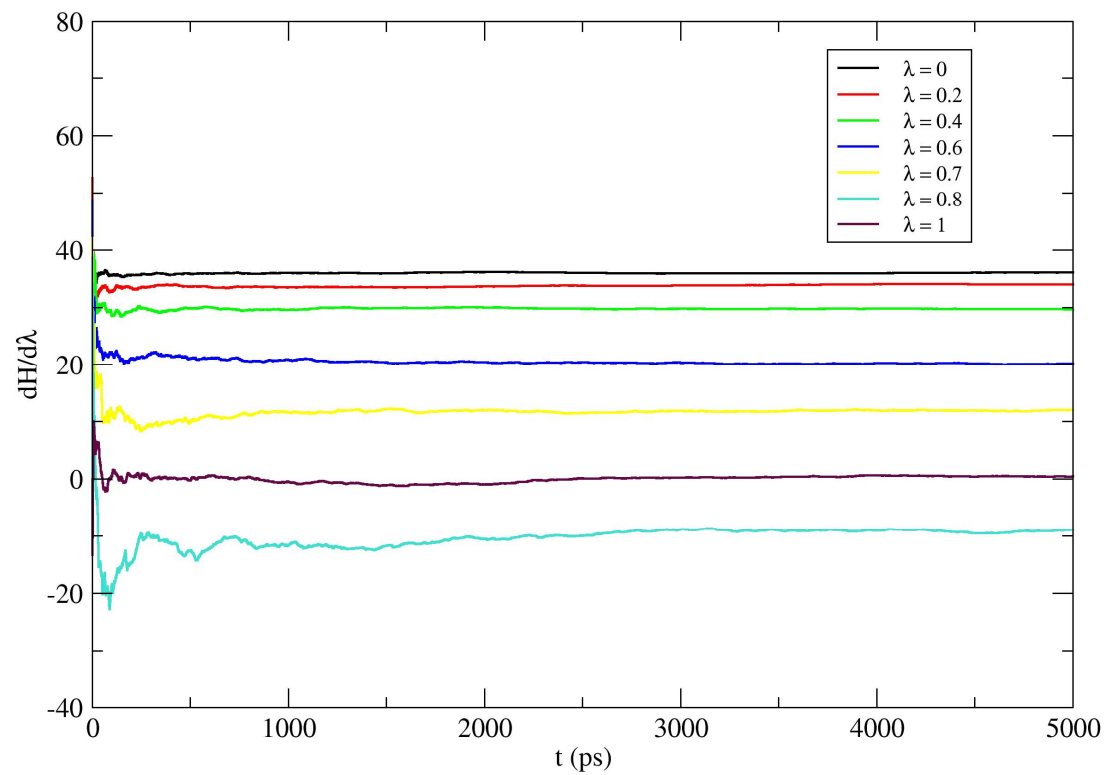


Figure 1: Derivative of the total Hamiltonian of the system with respect to time for several discrete λ values for the LJ decoupling of benzene in *n*-hexane.

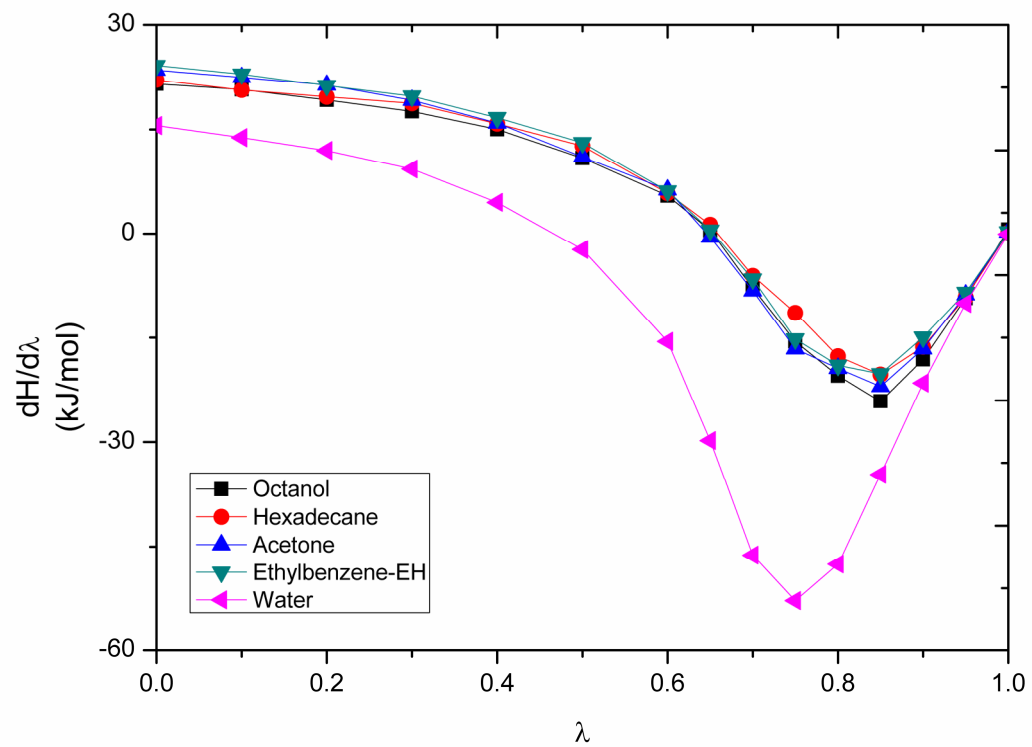


Figure 2: Derivative of the total Hamiltonian of the system with respect to the coupling parameter λ for the LJ decoupling of propane in different solvents.

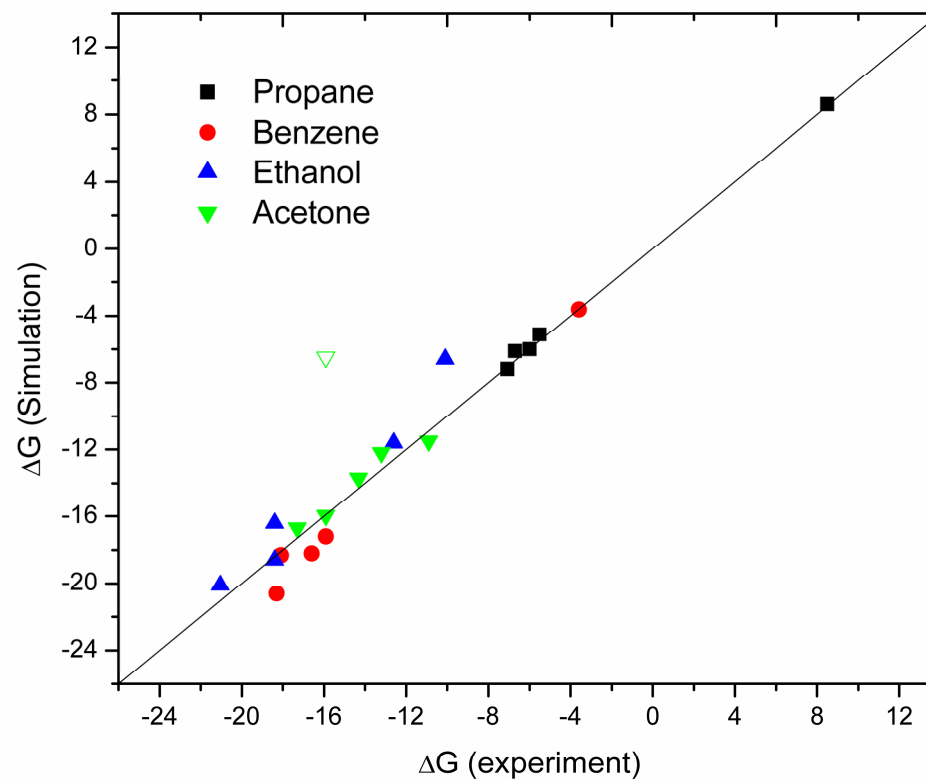


Figure 3: Overall comparison of experimental data and MD predictions for the Gibbs energy of solvation (in kJ/mol) of different solutes (propane, benzene, ethanol and acetone) in solvents of different polarity. For acetone, the open symbol refers to the original TraPPE point charges in the water phase (see Table 5).

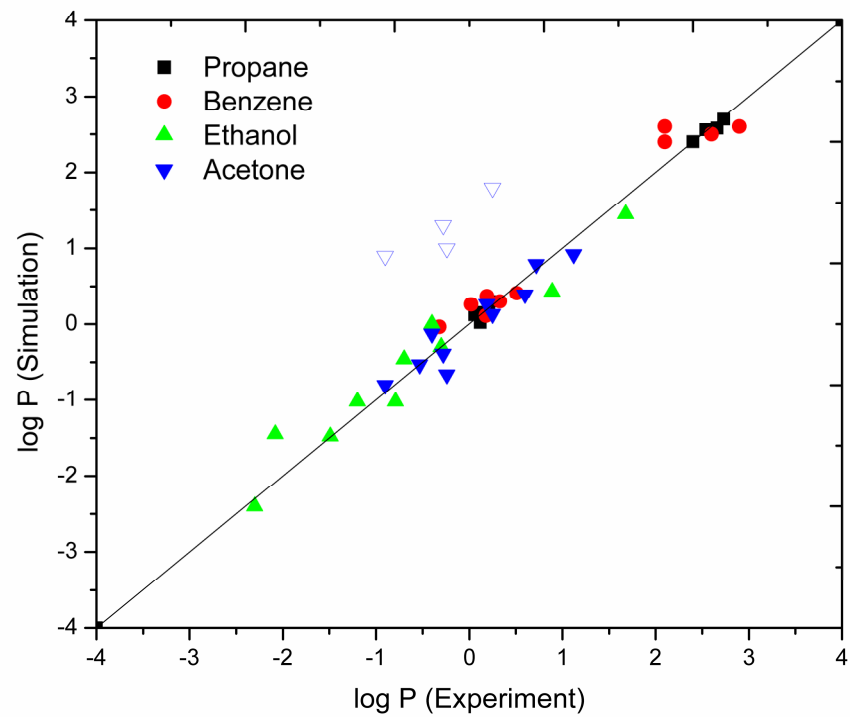


Figure 4: Overall comparison of experimental data or “estimated” experimental data against MD predictions for the log P of different solutes in various solvent pairs. Open symbols refer to predicted log P values of acetone in systems containing water using the original TraPPE point charges (See Table 6 for details).