

Molecular Simulation of Absolute Hydration Gibbs Energies of Polar Compounds

N. M. Garrido^{1,2}, A. J. Queimada¹, M. Jorge¹, I.G. Economou^{2,3}, E.A. Macedo^{1*}

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

² Molecular Thermodynamics and Modeling of Materials Laboratory, Institute of Physical Chemistry, National Center for Scientific Research “Demokritos”, GR-153 10, Aghia Paraskevi Attikis, Greece

³ The Petroleum Institute, Department of Chemical Engineering, PO Box 2533, Abu Dhabi, United Arab Emirates

* Author to whom all correspondence should be addressed at: eamacedo@fe.up.pt.

Revised version for publication in *Fluid Phase Equilibria*

Special issue for

VIII IBEROAMERICAN CONFERENCE ON PHASE EQUILIBRIA AND FLUID PROPERTIES
FOR PROCESS DESIGN – EQUIFASE 2009

Algarve – Praia de Rocha, Portugal 17 – 21 October 2009

Abstract

In this work, we present simulation-based predictions of the absolute hydration energy for several simple polar molecules with different functional groups, as well as for more complex multifunctional molecules. Our calculations were performed using the Thermodynamic Integration methodology where electrostatic and non-polar interactions were treated separately, allowing for a stable transition path between the end-points of the integration. An appropriate methodology for the analytical integration of the simulation data was applied. We compare the performance of three popular molecular mechanics force fields: TraPPE, Gromos and OPLS-AA for the description of solute atoms in MSPC/E water. It is observed that these force fields generally perform well for the simpler molecules, but are less accurate when multifunctional molecules are considered.

Keywords: Gibbs Hydration Energy, Molecular Modeling, Thermodynamic Integration

1. Introduction

Solvation Gibbs energy is an important thermodynamic function that provides a quantitative understanding of solute-solvent interactions. Solvation energy is fundamentally different from the dissolution Gibbs energy – while in the former only solute-solvent interactions are taken into account, in the second both solute-solute and solute-solvent interactions are considered. Although the solubility can only be directly obtained from the dissolution Gibbs energy, the solvation energy provides the energy requirements to solvate/dessolvate a particular solute in a given solvent and may be used to indirectly predict solubility [1, 2]. Thus, this information is of extreme importance in different research areas such as pharmaceutics, biochemistry and biomedicine. Solvation effects, as well as phase partitioning properties, play a vital role in many biomolecular processes, such as protein folding or active transport across cell membranes [3, 4]. Additionally, free energy differences are a vital quantity in computational drug discovery using current docking and scoring algorithms.

The continuous increase of computational power has led to faster calculations as well as to more efficient calculation algorithms, making routine free energy calculations possible with sufficient accuracy in a reasonable amount of time. In recent years, numerous predictions of the absolute hydration (water solvation) Gibbs energy have been reported in the literature, using several biomolecular force fields, with special emphasis on amino-acid side chain analogs [5-13]. Gibbs energy predictions from molecular simulation are highly sensitive to the details of the force field (FF) parameters used to describe interactions between solute and solvent molecules, such as atomic partial charges, Lennard-Jones (LJ) radii and well depth, as well as to the technical details of the molecular simulations [14]. In this context, prediction of free energy differences constitutes a critical test to the accuracy of a given FF. Specifically, the Gibbs energy is primarily dependent on the solute charges and relatively insensitive to the van der Waals parameters [9]. Indeed, electrostatic interactions play a major role in bio-molecular systems (highly polar solutes) when compared to van der Waals interactions [15]. Thus, it is of great importance to test different solute force field parameters, evaluating their predictive capability.

With this motivation, we have recently reported calculations of the solvation energy of different alkanes both in water and in 1-octanol using various FF, and went on to predict liquid phase partitioning in 1-octanol-water systems [16]. In the current work, we present an extension of these calculations to several molecules involving different functional groups (hydroxyl, amine, carbonyl and carboxyl). Moreover, we also include more complex polyfunctional molecules of particular importance for the pharmaceutical industry, such as benzoic acid (BA), acetylsalicylic acid (ASA), and ibuprofen (IBP) (Figure 1). The first two molecules are usually considered as models for general drug substances, since substituted aromatic rings and carboxyl groups are frequently found in many active pharmaceutical ingredients. ASA can be seen as an extension of the BA molecule obtained by adding a substituent in the *ortho* position, which may sterically hinder access of solvent molecules to

the carboxyl group, while IBP is a widely used pharmaceutical molecule. Experimental solvation characteristics of these molecules are well documented in the literature [17, 18]. In the present work, we compare the performance of three popular FFs to represent such molecules: the united atom (UA) TraPPE [19-25] and Gromos [9] and the all atom (AA) OPLS-AA [26, 27]. The remainder of this paper is organized as follows: in Section 2 we present the methodology adopted for the estimation of the Gibbs energy as well as the main simulation details. Results and discussion about the performance of each FF are detailed in Section 3 while the main conclusions are summarized in Section 4.

2. Theoretical and Simulation details

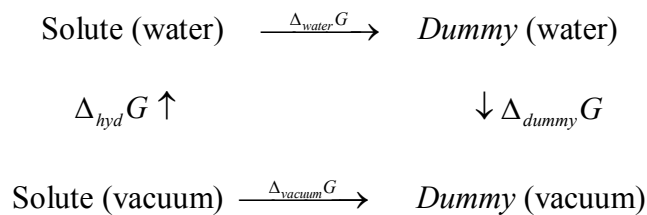
Solvation Gibbs energy represents the work required to transfer a solute molecule from the ideal gas phase to solution, at constant pressure and temperature:

$$\Delta_{hyd}G^{G \rightarrow L}(T, p) = \mu^L - \mu^G = \Delta G_{Elec}^{G \rightarrow L} + \Delta G_{LJ}^{G \rightarrow L} \quad (1)$$

This Gibbs energy of transfer from the gas to the liquid phase, $\Delta G^{G \rightarrow L}$, expressed in terms of the solute chemical potentials, can be estimated as the sum of electrostatic and non-polar (Lennard-Jones) contributions, where the latter includes the work for cavity formation (insertion of an uncharged van der Waals molecule into the solvent).

The solvation free energy can be experimentally determined from concentration measurements in different phases [28, 29]. Several experimental data have been reported in the literature for aqueous systems [30-33] providing a direct comparison between experiments and estimated values. For many complex molecules, however, experimental measurement of the solvation energy may not be possible as some solutes may not be chemically stable in one or more of the phases being evaluated. Additionally, whenever we wish to screen a large set of solutes, or when these are not available, as is the case for many drug candidates in early discovery stages, the availability of adequate computational tools to provide accurate estimates is strongly desirable.

In fact, the solvation Gibbs energy can also be calculated from an appropriate model using molecular dynamics (MD) simulations. **Within this approach, the hydration Gibbs energy, $\Delta_{hyd}G$, can be estimated based on the following thermodynamic cycle:**



In short, the thermodynamic cycle consists of three stages: i) turning the solute into a *dummy* molecule (non-bonded interactions are switched off) in vacuum ($\Delta_{water}G$); ii) solvating the *dummy* molecule ($\Delta_{dummy}G$); iii) turning the dummy molecule into the solute in water ($\Delta_{vacuum}G$). *Dummy* molecules are fictitious molecules that do not interact with their environment, so the free energy of the second stage is zero by definition. In practice, these molecules have no electrostatic or van der Waals interactions, but their intramolecular bonded interactions are the same as in the solute molecules. From this cycle, the hydration Gibbs energy can be simply estimated taking the difference between the vacuum and water stages [16, 34]. The separate calculation in vacuum is necessary to compensate for changes in solute-solute intramolecular non-bonded interactions that take place when the intermolecular interactions are switched off. For each case (solvent and vacuum), the associated free energy is estimated here using the thermodynamic integration method whose algorithm, free of hysteresis, has been described elsewhere [3]. In short, if we consider two generic well-defined states, a coupling parameter, λ , can be added to the total Hamiltonian of the system and used to describe the transition between the two states. Considering several discrete and independent λ values, equilibrium averages can be used to evaluate derivatives of the free energy with respect to λ . One then integrates the derivatives of the free energy along a continuous path connecting the initial and final states in order to obtain the free energy between them. In practice, the solvation free energy can be estimated as follows: i) simulate the system in water at different λ values; ii) simulate the system in vacuum at different λ values; iii) compute the solvation free energy from equation (2):

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

During the decoupling process, the electrostatic interactions were linearly interpolated between neighboring states while the LJ interactions were interpolated via soft-core interactions [35]. This soft-core dependence eliminates singularities in the calculation as the LJ interactions are turned-off [36]. We have used a power of λ , $p = 1$ for the λ dependence, and a soft-core parameter of 0.5.

MD simulations were performed using the GROMACS simulation suite [37]. As already found in other studies, the selection of an appropriate force field is the critical issue to obtain the best estimates of solvation energies from molecular simulations. In this study we have tested both united-atom and all-atom force fields. For solute description we have tested the TraPPE UA FF [20-24] together with the improvements proposed by Kamath et al. for carboxylic acids [25] and for esters [19]. Contrary to the original version of the TraPPE FF where all bond lengths are fixed, bond stretching was modeled here by a harmonic potential with force constants taken from the CHARMM FF [38]. As a second UA FF we have tested the biomolecular GROMOS FF as proposed by Oostenbrink et al. [9].

Finally we have tested the widely used OPLS-AA FF [26] with the optimized parameters for amines from Rizzo et al. [27] and for esters from Charifson et al. [39]. Solvated systems consisted of one solute molecule and 500 water molecules represented by the MSPC/E model [40] at 298 K and 1 bar. Covalent bonds involving hydrogen atoms were constrained using the LINCS algorithm [41] while the water geometry was fixed with the SETTLE algorithm [42]. For efficiency reasons, the reaction-field method [43], which approximates the medium beyond a cut-off distance of 1 nm by a dielectric continuum of uniform permittivity $\epsilon_{\text{rf}} = 80$, was used to handle long-range electrostatics. Simulations were performed using periodic boundary conditions in all directions, and Newton’s equations of motion for all the species were integrated using the leap-frog dynamic algorithm [44]. The remaining cut-off radii used were 1 nm for the short-range neighbor list and a 0.8-0.9 nm switched cut-off for the LJ interactions. We have applied long range corrections for energy and pressure as suggested by Shirts et al. [8]. Langevin stochastic dynamics [45] was used to control the temperature, with a frictional constant of 1 ps^{-1} while for constant pressure runs the Berendsen barostat [46], with a time constant of 0.5 ps and an isothermal compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$, was used to enforce pressure coupling. Vacuum simulations were performed without cutoffs.

For each simulation, we first run an energy minimization (the Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm [47] during 5000 steps followed by a steepest descent minimization during 500 steps) followed by a constant volume equilibration (100 ps), a constant pressure equilibration (500 ps) enough to obtain complete equilibration of the box volume, and finally a 5 ns NPT production stage, using a time step of 2 fs. This procedure was repeated for several discrete λ values, allowing for separate minimization, from $\lambda = 0$ (fully interacting solute) to $\lambda = 1$ (non-interacting solute). **It is worth noticing that 5 ns is, in principle, enough to observe several transitions between stable configurations indicating sufficient sampling of different torsional degrees of freedom, even for the more complex compounds [14].**

In the present work the reported statistical uncertainties were obtained from block averaging [48] and the integration of the Hamiltonian derivatives was carried out by fitting the data to an appropriate fitting function and integrating the curve analytically. For the electrostatic component, a fourth degree polynomial was used:

$$\left(\frac{\partial H}{\partial \lambda} \right)_{Ch} = \sum_{i=0}^4 A_i \lambda^i \quad (3)$$

For the LJ component, we have used a previously developed function based on a physically-based approximation to the cavity formation and dispersion interaction terms [49]:

$$\left(\frac{\partial \mathcal{H}}{\partial \lambda}\right)_{LJ} = B_0 \lambda^2 + B_1 \lambda - \frac{B_2}{\lambda^2 - B_3 \lambda + B_4} + \frac{B_2}{B_4} \quad (4)$$

In equations (3) and (4) A_i and B_i are fitting parameters. This integration procedure has recently been shown to increase the precision of the calculated free energies and decrease the number of necessary intermediate points, relative to commonly used numerical integration methods [49]. In this work, we have observed that all the hydration free energies calculated using this improved integration procedure were closer to the experimental values than estimates based on the widely employed trapezoidal rule.

3. Results and Discussion

Results for the LJ and electrostatic contributions to the total hydration free energy are presented in Figures 2 and 3, respectively, for the OPLS-AA FF. Similar plots were obtained for the other FFs. In general, smooth paths were obtained between the initial and end-state points. The behavior of the LJ component is non-monotonic, increasing up to about $\lambda = 0.8$ and decreasing for higher λ values. This behavior reflects two counterbalancing factors: unfavorable excluded volume effects, due to cavity creation in the solvent, and favorable solute-solvent dispersion interactions. In contrast, the behavior of the electrostatic component is monotonic for all solutes in the entire interval of integration. Results of the integration of the $\partial \mathcal{H} / \partial \lambda$ plots and corresponding calculated values of $\Delta_{hyd} G$ (as well as their components $\Delta_{hyd} G^C$ and $\Delta_{hyd} G^{LJ}$) via equation (2) are shown in Table 1, together with the corresponding estimated uncertainties. A global comparison of the different FF predictions against experimental data is provided in Figure 4.

For the smaller molecules (methanol, propanol, acetone, ethylamine and acetic acid), the description of the different functional group representatives is generally quite satisfactory and the absolute average deviations (AAD) between simulations and experiments are 6.4, 3.4 and 5.5 kJ/mol for TraPPE, Gromos and OPLS-AA FFs respectively. TraPPE is a united atom force field where (CH_x) groups are treated as pseudo atoms located at the sites of the carbon atoms, whereas all other atoms (e.g. hydroxyl O and H) are modeled explicitly. Such UA force fields have the advantage, over explicit hydrogen force fields (e.g. OPLS-AA), of being more computationally efficient by decreasing the number of degrees of freedom evaluated in the simulations. During the parameterization of the TraPPE force field, the main purpose was the ability to transfer atoms or groups of atoms between different molecules, keeping a good description of the phase equilibrium properties. It is important to notice that the parameterization of each functional group was performed independently from the others, and that multifunctional molecules were not included in the parameterization set. Furthermore, no solvation effects were included in the parameterization. Nevertheless, TraPPE provides a reasonable description

of the hydration free energy of alcohols (see Figure 4) and, as shown in our previous paper, of alkanes [16]. However, it significantly underestimates the hydration energy of ethylamine, acetone and acetic acid. For the case of acetone, it seems that the partial charges presented in the TraPPE FF are not able to correctly describe the electrostatic term (the value of $\Delta_{hyd}G^C$ for acetone with this FF is significantly below the results for the other two FFs, as shown in Table 1), and consequently the total $\Delta_{hyd}G$ term is quite underestimated. An improvement of the TraPPE FF parameters for carbonyl groups, particularly the point charges, would be desirable. The same could be said of acetic acid, but intriguingly TraPPE performs quite well for benzoic acid, precisely because it predicts a relatively low electrostatic contribution. This suggests that a correct parameterization of carboxylic groups in different molecules may be more challenging than previously thought.

Not surprisingly, the best overall predictions among these smaller molecules were afforded by the Gromos force field, as this parameter set has included hydration energies, together with heats of vaporization and liquid densities, in its recent reparameterization [9]. In the case of the Gromos force field, the philosophy underlying its parameterization was a correct balance between an accurate description of the interaction energy as a function of conformation together with a relatively simple functional form. The main groups used in the test set of this last reparameterization were proteins, nucleotides, sugars and lipids. The Gromos FF shares with TraPPE the advantage of a more efficient united atom description of (CH_x) groups. Gromos yields an excellent description of alkanes [16], alcohols, ethylamine and acetone, but strongly overestimates the free energy of acetic acid and benzoic acid. The reason for this is a much more negative electrostatic contribution predicted by Gromos for these molecules, compared to the other two FFs (Table 1). Therefore, there may be much to gain by reparameterizing the Gromos partial charges for carboxylic acids.

Finally, the original OPLS-AA force field was parameterized to reproduce condensed-phase properties (densities and heats of vaporization) and gas-phase torsional profiles. However, some subsequent improvements, as in the case of amines [27], included relative free energy differences between molecules of homologous series in the parameterization. Of the three force fields studied here, OPLS-AA is the one that includes the largest parameter set and has been most widely applied to molecules with different functionalities. Our previous work showed that this FF was slightly worse than the two UA models in predicting hydration energies of alkanes, but better for describing alkane solvation in 1-octanol [16]. Here, we see that OPLS-AA yields reasonable descriptions of the hydration energy of the small solutes, with the exception of ethylamine. This is a deficiency already reported in the literature [10, 11, 50-52], where the correct description of amine (and amide) groups with the OPLS-AA force field remains problematic even after their reparameterization. This reparameterization solved the question of the relative hydration Gibbs energy between progressively more substituted amines [27], but apparently is still not able to accurately predict the absolute

hydration energy of individual amines. Interestingly, OPLS-AA provides the best description of acetic acid and benzoic acid, which suggests that the point charges for carboxylic groups in this force field are more accurate than in the two UA force fields.

Although the simulation predictions for the smaller test set using Gromos (the best performing FF) deviate on average from experiments “only” 3.4 kJ/mol, this quantity still represents more than 10% of the experimental values. However, it is worthwhile to notice that this is a purely predictive method and thus requires no additional experimental information. When compared to our previous work for non-polar molecules [16], the deviations from experiments have increased from 1.2 kJ/mol to 3.4 kJ/mol as observed here (Gromos FF). However, contrary to non-polar solutes, two components (LJ and electrostatic) need to be estimated to evaluate the total hydration energy of polar solutes. In light of this fact, and of the high sensitivity of the results to the values of the point charges of the solutes, we consider that general purpose FFs such as those studied here are able to provide a reasonable description of the hydration energies of small mono-functional molecules.

Considering the full data set, the AAD increases up to 6.4, 8.8 and 7.1 kJ/mol for TraPPE, Gromos and OPLS-AA, respectively. In this case, deviations are considerably larger indicating that for multifunctional molecules the force field parameters used cannot cope with the increased number of degrees of freedom of the molecular models. As a comparison, for practical applications in pharmaceutical research, the accuracies required for predicting protein-ligand binding affinity (much more complex systems) are 2-4 kJ/mol [53]. As we can observe, current parameter sets have not yet achieved this goal. This is most likely due to the fact that the FF parameters are obtained from calculations on simple mono-functional molecules, thus ignoring possible interactions between different functional groups that may exist on the same solute molecule. Also, the polarization of a given functional group may change according to its environment, and this effect is also hard to capture when only mono-functional molecules of the same type are considered.

We also noted that although some differences between the three force fields were found for the total hydration energy values, those differences are considerably higher when we consider each of the contributing terms separately (LJ and electrostatic). This suggests that there might be some degree of error compensation in the parameterization of the LJ parameters and partial charges for some (or all) of the force fields. However, as we have observed good predictions for non-polar species (where solely the LJ term is evaluated) we suggest that results for polar species can be improved by tuning atomic partial charges. Furthermore, the transferability of these FFs to different solvents appears to be very limited, as a given fixed-charge parameter set is parameterized for a specific solvent. In this case, polarizable FF [54-57] must be object of further study.

4. Conclusions

In this study we have computed the absolute hydration free energy for a diverse set of small and complex polar molecules using thermodynamic integration of MD simulations in explicit solvent, and compared our results against experimental data. We have evaluated the performance of three different popular force-fields: Gromos, TraPPE and OPLS-AA. The hydration free energies were predicted without any experimental information and without any fitting of parameters. A recently proposed integration method, based on the use of appropriate analytic fitting functions for the Hamiltonian derivative, was employed, leading to estimates that were closer to experimental data.

In general, all three force fields provide a reasonable description of the hydration energies of small mono-functional molecules. Gromos seems to give the best overall estimates, which is perhaps to be expected since this force field used free energy calculations in its parameterization. We also found that using the all-atom OPLS-AA FF does not necessarily improve the solvation free energy calculations, in some cases even leading to worse estimates than those provided by Gromos or TraPPE. Each FF was found to be particularly problematic for a given functional group (carbonyl for TraPPE, carboxyl for Gromos and amine for OPLS-AA), particularly in the description of the electrostatic contribution to the free energy. Significant improvements are therefore to be expected if the point charges of these groups are reparameterized in the respective force fields. With respect to larger multifunctional solutes, more similar to typical drug molecules, we have observed that the current force fields are somewhat limited, since the errors increase substantially compared to the mono-functional molecules. This suggests that more care should be taken when transferring parameters from simple to more complex molecules, at least in what free energy calculations are concerned. A systematic study of the effect of the interaction between different functional groups present in the same solute molecule on the hydration free energy results would be highly desirable and is suggested as future work.

5. Acknowledgments

This work was in part supported by LSRE financing by FEDER/POCI/2010, for which the authors are thankful. NMG acknowledges his *Fundação para a Ciência e a Tecnologia* (Lisboa, Portugal) Ph.D. scholarship SFRH/BD/47822/2007. AJQ acknowledges POCI.

6. References

- [1] D. S. Palmer, A. Llinas, I. Morao, G. M. Day, J. M. Goodman, R. C. Glen, J. B. O. Mitchell, *Molecular Pharmaceutics* 5 (2008) 266-279.
- [2] J. Westergren, L. Lindfors, T. Hoglund, K. Luder, S. Nordholm, R. Kjellander, *J. Phys. Chem. B* 111 (2007) 1872-1882.
- [3] W. L. Jorgensen, *Science* 303 (2004) 1813-1818.
- [4] P. A. Kollman, *Acc. Chem. Res.* 29 (1996) 461-469.
- [5] J. M. Wang, P. Cieplak, P. A. Kollman, *J. Comp. Chem.* 21 (2000) 1049-1074.
- [6] B. Hess, N. F. A. van der Vegt, *J. Phys. Chem. B* 110 (2006) 17616-17626.
- [7] M. R. Shirts, V. S. Pande, *J. Chem. Phys.* 122 (2005) 134508.
- [8] M. R. Shirts, J. W. Pitner, W. C. Swope, V. S. Pande, *J. Chem. Phys.* 119 (2003) 5740-5761.
- [9] C. Oostenbrink, A. Villa, A. E. Mark, W. F. Van Gunsteren, *J. Comp. Chem.* 25 (2004) 1656-1676.
- [10] M. Udier-Blagovic, P. M. De Tirado, S. A. Pearlman, W. L. Jorgensen, *J. Comp. Chem.* 25 (2004) 1322-1332.
- [11] D. L. Mobley, E. Dumont, J. D. Chodera, K. A. Dill, *J. Phys. Chem. B* 111 (2007) 2242-2254.
- [12] D. Shivakumar, Y. Deng, B. Roux, *J. Chem. Theory Comput.* 5 (2009) 919-930.
- [13] J. Tomasi, M. Persico, *Chem. Rev.* 94 (1994) 2027-2094.
- [14] N. M. Garrido, M. Jorge, A. J. Queimada, I. G. Economou, E. A. Macedo, *Fluid Phase Equilib* 289 (2010) 148-155.
- [15] W. F. van Gunsteren, D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D. P. Geerke, A. Glättli, P. H. Hünenberger, M. A. Kastenholz, C. Oostenbrink, M. Schenk, D. Trzesniak, N. F. A. van der Vegt, H. B. Yu, *Angew. Chem. Int. Ed.* 45 (2006) 4064-4092.
- [16] N. M. Garrido, A. J. Queimada, M. Jorge, E. A. Macedo, I. G. Economou, *J. Chem. Theory Comput.* 5 (2009) 2436-2446.
- [17] G. L. Perlovich, A. Bauer-Brandl, *Pharm. Res.* 20 (2003) 471-478.
- [18] G. L. Perlovich, S. V. Kurkov, A. N. Kinchin, A. Bauer-Brandl, *AAPS PharmSci* 6 (2004) 1-9.
- [19] G. Kamath, J. Robinson, J. J. Potoff, *Fluid Phase Equilibria* 240 (2006) 46-55.
- [20] B. Chen, J. J. Potoff, J. I. Siepmann, *J. Phys. Chem. B* 105 (2001) 3093-3104.
- [21] M. G. Martin, J. I. Siepmann, *J. Phys. Chem. B* 102 (1998) 2569-2577.
- [22] M. G. Martin, J. I. Siepmann, *J. Phys. Chem. B* 103 (1999) 4508-4517.
- [23] J. M. Stubbs, J. J. Potoff, J. I. Siepmann, *J. Phys. Chem. B* 108 (2004) 17596-17605.
- [24] C. D. Wick, J. M. Stubbs, N. Rai, J. I. Siepmann, *J. Phys. Chem. B* 109 (2005) 18974-18982.
- [25] G. Kamath, F. Cao, J. J. Potoff, *Journal of Physical Chemistry B* 108 (2004) 14130-14136.
- [26] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* 118 (1996) 11225-

11236.

- [27] R. C. Rizzo, W. L. Jorgensen, *J. Am. Chem. Soc.* 121 (1999) 4827-4836.
- [28] R. Wolfenden, L. Andersson, P. M. Cullis, C. C. B. Southgate, *Biochem.* 20 (1981) 849-855.
- [29] A. Ben-Naim, Y. Marcus, *J. Chem. Phys.* 81 (1984) 2016-2027.
- [30] L. Michielan, M. Bacilieri, C. Kaseda, S. Moro, *Bioorg. Med. Chem.* 16 (2008) 5733-5742.
- [31] K. N. Rankin, T. Sulea, E. O. Purisima, *J. Comp. Chem.* 24 (2003) 954-962.
- [32] R. C. Rizzo, T. Aynechi, D. A. Case, I. D. Kuntz, *J. Chem. Theory Comput.* 2 (2006) 128-139.
- [33] S. Cabani, P. Gianni, V. Mollica, L. Lepori, *J. Sol. Chem.* 10 (1981) 563-595.
- [34] A. Leach, *Molecular Modeling: principles and applications*, Prentice-Hall, 2001.
- [35] T. M. Beuler, R; van Schaik, RC; Gerber, PR; van Gunsteren, WF, *Chem. Phys. Lett.* 222 (1994) 529-539.
- [36] J. W. Pitera, W. F. Van Gunsteren, *Molec. Sim.* 28 (2002) 45-65.
- [37] D. Van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, *J. Comput. Chem.* 26 (2005) 1701-1718.
- [38] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* 102 (1998) 3586-3616.
- [39] P. S. Charifson, R. G. Hiskey, L. G. Pedersen, *J. Comput. Chem.* 11 (1990) 1181-1186.
- [40] G. C. Boulougouris, I. G. Economou, D. N. Theodorou, *J. Phys. Chem. B* 102 (1998) 1029-1035.
- [41] B. Hess, H. Bekker, H. J. C. Berendsen, J. Fraaije, *J. Comp. Chem.* 18 (1997) 1463-1472.
- [42] S. Miyamoto, P. A. Kollman, *J. Comp. Chem.* 13 (1992) 952-962.
- [43] F. S. Lee, A. Warshel, *J. Chem. Phys.* 97 (1992) 3100-3107.
- [44] W. van Gunsteren, H. Berendsen, *Mol. Sim.* 1 (1988) 173-185.
- [45] W. F. Van Gunsteren, H. J. C. Berendsen, *Mol. Phys.* 45 (1982) 637-647.
- [46] H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola, J. R. Haak, *J. Chem. Phys.* 81 (1984) 3684-3690.
- [47] D. C. Liu, J. Nosedal, *Math. Programm.* 45 (1989) 503-528.
- [48] B. Hess, *J. Chem. Phys.* 116 (2002) 209-217.
- [49] M. Jorge, N. M. Garrido, A. J. Queimada, I. G. Economou, E. A. Macedo, *J. Chem. Theory Comput.* (2010) in Press.
- [50] D. J. Tannor, B. Marten, R. Murphy, R. A. Friesner, D. Sitkoff, A. Nicholls, M. Ringnalda, W. A. Goddard, B. Honig, *J. Am. Chem. Soc.* 116 (1994) 11875-11882.
- [51] P. Y. Morgantini, P. A. Kollman, *J. Am. Chem. Soc.* 117 (1995) 6057-6063.

- [52] Y. B. Ding, D. N. Bernardo, K. Kroghjerspersen, R. M. Levy, *J. Phys. Chem.* 99 (1995) 11575-11583.
- [53] G. Kaminski, E. M. Duffy, T. Matsui, W. L. Jorgensen, *J. Phys. Chem.* 98 (1994) 13077-13082.
- [54] P. Y. Ren, J. W. Ponder, *J. Comp. Chem.* 23 (2002) 1497-1506.
- [55] G. A. Kaminski, H. A. Stern, B. J. Berne, R. A. Friesner, Y. X. X. Cao, R. B. Murphy, R. H. Zhou, T. A. Halgren, *J. Comp. Chem.* 23 (2002) 1515-1531.
- [56] G. A. Kaminski, H. A. Stern, B. J. Berne, R. A. Friesner, *J. Phys. Chem. A* 108 (2004) 621-627.
- [57] S. Patel, C. L. Brooks, *J. Comp. Chem.* 25 (2004) 1-15.

Table 1: Electrostatic and Lennard-Jones contributions to the total hydration free energy (kJ/mol). Simulations and experimental data[18, 30].

Solute	Force Field	$\Delta_{Hyd}G^C$	$\Delta_{Hyd}G^{LJ}$	$\Delta_{Hyd}G^{Calc}$	$\Delta_{Hyd}G^{Exp}$
Methanol	TraPPE	-30.43 ± 0.05	11.56 ± 0.10	-18.98 ± 0.11	-21.5
	Gromos	-29.20 ± 0.07	6.47 ± 0.10	-22.73 ± 0.12	
	OPLS-AA	-26.08 ± 0.04	7.02 ± 0.09	-19.06 ± 0.10	
1-Propanol	TraPPE	-30.27 ± 0.06	5.86 ± 0.07	-24.41 ± 0.09	-20.6
	Gromos	-29.20 ± 0.07	7.17 ± 0.15	-22.03 ± 0.17	
	OPLS-AA	-27.19 ± 0.05	9.14 ± 0.14	-18.05 ± 0.15	
Ethylamine	TraPPE	-17.61 ± 0.08	4.30 ± 0.28	-13.31 ± 0.29	-19.5
	Gromos	-26.31 ± 0.05	11.32 ± 0.14	-14.99 ± 0.15	
	OPLS-AA	-13.81 ± 0.04	8.63 ± 0.12	-5.19 ± 0.13	
Acetone	TraPPE	-12.17 ± 0.05	5.04 ± 0.30	-7.13 ± 0.30	-16.1
	Gromos	-20.43 ± 0.03	4.09 ± 0.10	-16.34 ± 0.10	
	OPLS-AA	-18.73 ± 0.10	5.39 ± 0.29	-13.34 ± 0.31	
Acetic Acid	TraPPE	-21.31 ± 0.03	3.54 ± 0.11	-17.77 ± 0.11	-28.1
	Gromos	-38.50 ± 0.18	0.72 ± 0.07	-37.78 ± 0.19	
	OPLS-AA	-26.16 ± 0.04	3.62 ± 0.09	-22.54 ± 0.09	
IBP	TraPPE	-19.13 ± 0.28	-0.60 ± 0.43	-19.73 ± 0.51	-12.8
	Gromos	-39.13 ± 0.14	2.90 ± 1.17	-36.23 ± 1.18	
	OPLS-AA	-34.90 ± 0.07	11.75 ± 0.33	-23.15 ± 0.34	
ASA	TraPPE	-38.92 ± 0.10	3.72 ± 0.33	-35.20 ± 0.34	-23.9
	Gromos	-38.80 ± 0.03	-2.60 ± 0.61	-41.40 ± 0.61	
	OPLS-AA	-44.63 ± 0.07	2.15 ± 0.33	-42.48 ± 0.34	
Benzoic Acid	TraPPE	-18.91 ± 0.05	4.26 ± 0.04	-14.65 ± 0.06	-15.6
	Gromos	-27.05 ± 0.14	-1.75 ± 0.43	-28.10 ± 0.45	
	OPLS-AA	-17.75 ± 0.29	2.47 ± 0.25	-15.28 ± 0.38	

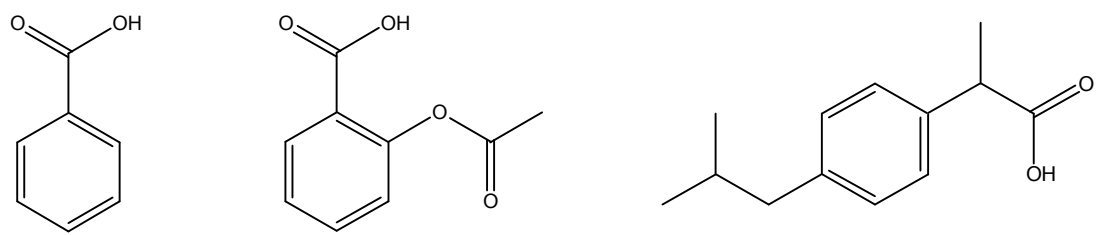


Figure 1: Benzoic acid ($C_7H_6O_2$), Acetylsalicylic acid ($C_9H_8O_4$) and Ibuprofen ($C_{13}H_{18}O_2$) molecules

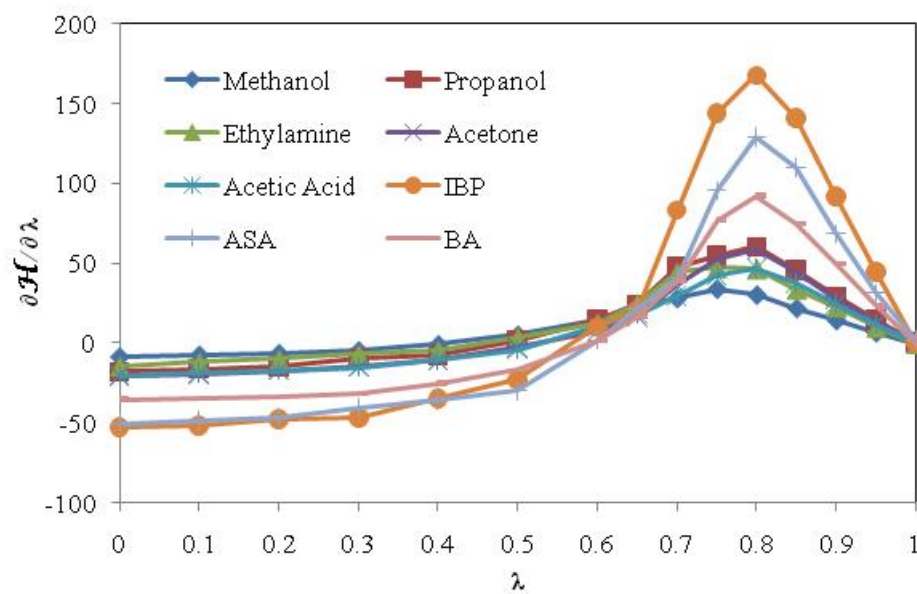


Figure 2: Lennard-Jones contributions (vacuum – water) to the derivative of the Hamiltonian (kJ/mol) with respect to λ for the OPLS-AA force field.

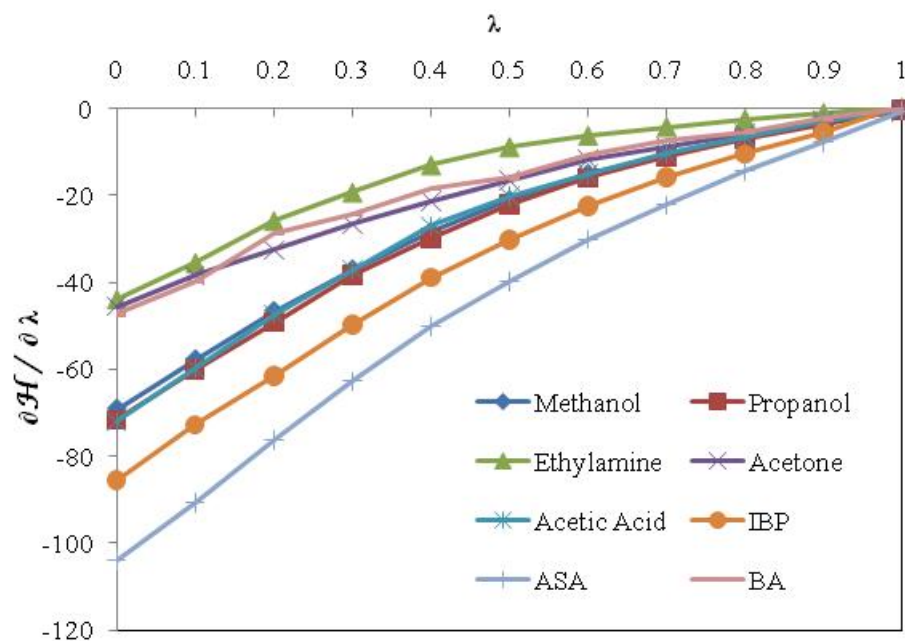


Figure 3: Electrostatic contributions (vacuum – water) to the derivative of the Hamiltonian (kJ/mol) with respect to λ for the OPLS-AA force field.

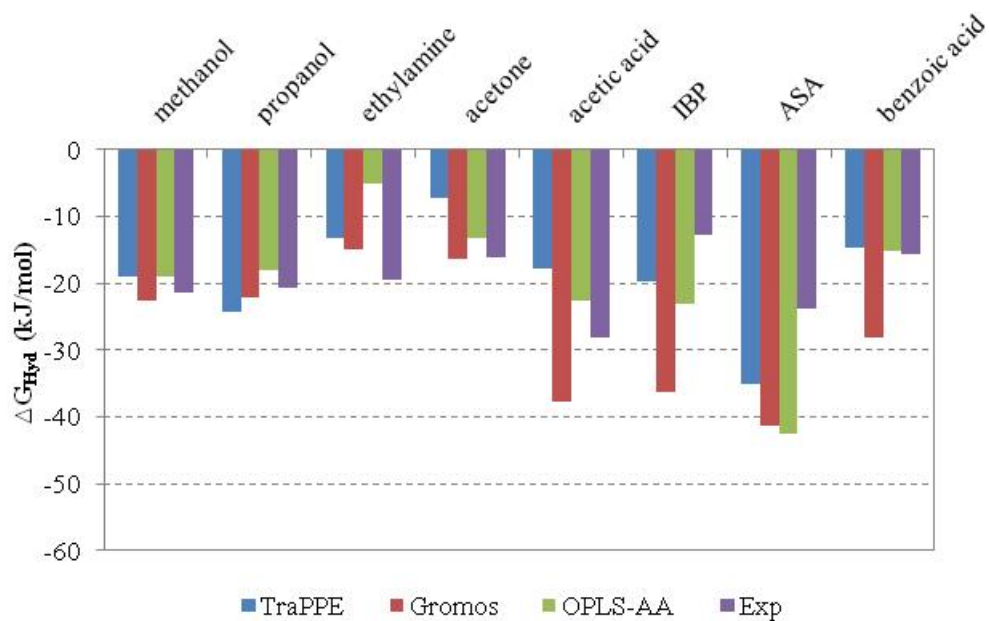


Figure 4: Overview of the performance of different force fields in predicting hydration free energies.

List of symbols

G	Gibbs Energy
T	Temperature
p	Pressure
\mathcal{H}	Total Hamiltonian of the system
μ	Chemical potential
λ	Coupling parameter
ε	dielectric permittivity