J Comput Aided Mol Des (2010) 24:307–316
DOI 10.1007/s10822-010-9343-7

Predicting hydration free energies using all-atom molecular dynamics simulations and multiple starting conformations

Pavel V. Klimovich · David L. Mobley

Received: 16 November 2009 / Accepted: 16 March 2010 / Published online: 6 April 2010

Abstract Molecular dynamics simulations in explicit solvent were applied to predict the hydration free energies for 23 small organic molecules in blind SAMPL2 test. We found good agreement with experimental results, with an RMS error of 2.82 kcal/mol over the whole set and 1.86 kcal/mol over all the molecules except several hydroxyl-rich compounds where we find evidence for a systematic error in the force field. We tested two different solvent models, TIP3P and TIP4P-Ew, and obtained very similar hydration free energies for these two models; the RMS difference was 0.64 kcal/mol. We found that preferred conformation of the carboxylic acids in water differs from that in vacuum. Surprisingly, this conformational change is not adequately sampled on simulation timescales, so we apply an umbrella sampling technique to include free energies associated with the conformational change. Overall, the results of this test reveal that the force field parameters for some groups of molecules (such as hydroxyl-rich compounds) still need to be improved, but for most compounds, accuracy was consistent with that seen in our previous tests.

Keywords Hydration · Alchemical · Free energy · Molecular dynamics

Introduction

Hydration free energies, describing the free energy of transfer of a small molecule from gas to water, have become a hot topic in recent literature. Here, our focus is on computing these free energies from molecular simulations. Some early simulation studies computed hydration free energies of amino acid sidechain analogs [7, 15, 29, 38], but more recent work has moved to computing hydration free energies of diverse sets of small molecules, both in explicit solvent [19, 24, 25] and in implicit solvent [22, 36]. Two key facts drive this interest in hydration free energies. First, hydration free energies can now be computed from molecular simulations quite precisely. This provides a fundamental point of contact with experiment and a way of assessing the quality of force fields and potentially identifying systematic errors [24]. Second, hydration free energies are seen as an important component of a binding free energy calculation, since binding involves desolvation of a ligand and a receptor binding site as the two form a complex. Thus, they provide some idea of the level of accuracy that might be expected when simulations are applied in drug discovery applications. Hydration free energies are also related to physical properties of interest in drug discovery (such as water-octanol parititioning coefficients [37] and solubilities [14]) and can provide fundamental insight into how water organizes around solutes [21].

Blind challenge tests are important in assessing the accuracy of methods for at least two reasons. First, if we want methods that can be applied reliably in a predictive context, we should test them in a related context. Second, when results are known in advance, it is easy for bias to creep in, even unintentionally. One example of this would be if a group performs an initial set of calculations and obtains results that have some distribution of errors relative to experiment, and then picks those calculations which performed worst relative to experiment and repeats *only* those calculations, perhaps using different starting velocities. It is likely that, due to statistical fluctuations, some of

P. V. Klimovich · D. L. Mobley (✉)
Department of Chemistry, University of New Orleans,
2000 Lakeshore Drive, 70148 New Orleans, LA, USA
e-mail: dmobley@gmail.com

these results will be better than the original set, so the average or RMS error will decrease simply due to statistical fluctuations (though, with sufficiently precise results, this effect will be minimal). Unfortunately, this also means that the overall accuracy of the approach will, in such a case, appear better in this test than it would in a real-world application or in a blind test.

This manuscript reports the results of a blind test of hydration free energy calculations using alchemical free energy calculations based on molecular dynamics simulations. Alchemical methods provide a rigorous way of computing true free energies, which include entropic contributions, as well as effects of any conformational changes. The present study marks our third such blind test of alchemical free energy calculations on hydration free energies with these methods [23, 25]. Most recently, we entered the blind challenge immediately preceding the one reported in this special issue, and other participants also applied a variety of other methods to predicting hydration free energies [6, 12, 16, 23, 26, 35]. Previously, we also applied similar alchemical methods to computing binding free energies in a predictive context [3, 20]

One message from our hydration free energy studies so far in explicit solvent (with fixed-charge force fields) has been that we can expect an RMS error relative to experiment in the 1–2 kcal/mol range in many applications. Over all 599 (net neutral) molecules we studied to date [19, 23–25] the RMS error was 1.7 kcal/mol, with an average absolute error of 1.2 kcal/mol. And these errors are not distributed randomly; some functional groups tend to be particularly challenging [23, 24]. RMS errors were also substantially higher for our 2008 blind test, though this was mostly caused by the many hypervalent sulfur-containing compounds in the set that were particularly poorly described. Ideally, here, we would see similar RMS errors, in the neighborhood of 1.3–1.8 kcal/mol. However, RMS errors for polyfunctional or highly polar molecules tend to be higher, which could push this up [23, 25].

In previous studies, we tested several different charge models [19, 23, 25]. We found that generally, charge models based on higher levels of quantum mechanics or other enhancements did not drastically improve results [19, 25], though the possibility remains open that performance benefits might be seen for some classes of molecules or on large enough sets. In view of our past work on charge models, then, here we focus on conformational sampling and convergence issues, seeking to ensure that we obtain correct hydration free energies that are unbiased by starting conformation. Previous work in implicit solvent suggested that some conformational transitions for small molecules can be relatively slow [22], so adequately sampling conformational changes is one focus of this study.

Simulation methods

Overview

In this work, we computed hydration free energies in explicit solvent for 31 small organic molecules, using an all-atom force field in explicit water. The approach was similar to that used in previous explicit solvent studies [19, 23, 25]. We used explicit solvent molecular dynamics simulations with two types of water models: TIP3P [11] and TIP4P-Ew [8]. Free energies were computed using alchemical free energy techniques [30, 32].

These techniques allow computation of free energy difference between different thermodynamic states (here, the solute in gas versus the solute in water) by introducing a series of intermediate "alchemical" states spanning between these. Each state is simulated separately. Here, these intermediate states involve first turning off the solute partial charges in water, then turning off Lennard-Jones interactions between the solute and water, then restoring the solute partial charges in gas. From these simulations, free energy differences can be computed as discussed below.

System preparation

Starting mol2 structure files were provided by the organizers of OpenEye's SAMPL2 event.

Since one of our purposes here was to test the conformation dependence of results as a check for convergence, we needed multiple distinct starting conformations. For the molecules treated with the TIP3P water model, several molecular conformations, from 1 to 4 per molecule, were generated from mol2 files using OMEGA. In addition, we retained the starting conformation provided by the organizers, for a total of 2–5 conformations per molecule. Simulations were conducted beginning from each of these (potentially diverse) starting conformations.

Simulations were performed in GROMACS 3.3.3 [33] using Generalized Amber Force Field (GAFF) [39, 40] small molecule parameters. Partial charges were computed using AM1-BCC [9, 10] as implemented in the AmberTools v. 1.1 version of Antechamber [40]. Charges were computed for just one conformer of each molecule (for the single top-ranked conformation produced by OMEGA) and used for all conformers of that molecule. The molecules were solvated using GROMACS utilities in a dodecahedral simulation box with at least 1.2 nm from the solute to the nearest simulation box edge. As noted previously, we made minor modifications to GAFF such as adding some missing improper torsions (with the default values) [23, 24], and a change to a Lennard-Jones

well depth for triple bonded carbons (which did not affect any compounds in this set).

General simulation parameters

Simulations were conducted at many intermediate alchemical states (with corresponding λ values) where each λ value is simulated separately. First, solute electrostatics are gradually turned off with the variable λ (where $\lambda = 0$, 0.25, 0.5, 0.75, 1.0 in turn). Second, solute-water Lennard-Jones interactions are turned off in water using soft core potentials [2] with the parameters suggested by Shirts et al. [29] and Mobley et al. [17] ($\alpha = 0.5$, with a soft core exponent of 1), as previously [24]. For this step, λ values were 0.0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1.0. Finally, solute electrostatics were turned back on in vacuum (with $\lambda = 0$, 0.25, 0.5, 0.75, 1.0). The free energy of each of these component steps was computed using the Bennett acceptance ratio (BAR) [1] and then the total hydration free energy was computed as $\Delta G_{\mathrm{hyd}} = \Delta G_{\mathrm{chg,vac}} - \Delta G_{\mathrm{chg}} - \Delta G_{\mathrm{LJ}}$, where ΔG_{chg} denotes the free energy of turning off the electrostatics in water, $\Delta G_{\mathrm{chg,vac}}$ denotes the same quantity for vacuum, and ΔG_{LJ} denotes the free energy of turning off the solute-water Lennard-Jones interactions in water.

Simulation protocols were virtually identical to those described in reference [23]. Langevin dynamics was used for temperature control. The reference temperature was 300 K and the frictional constant was 10 ps^{-1}. At each λ value, the minimized (steepest descents) starting structure was run through an equilibration procedure consisting of 10 ps of constant volume equilibration, followed by 100 ps of constant pressure equilibration with the Berendsen weak-coupling scheme with a time constant of 0.5 ps, a reference pressure of 1.0 atm, and an isothermal compressibility of 4.5×10^{-5} bar. The production part of the simulation was 5 ns at each λ, at constant volume.

Since the production simulations were at constant volume, it was important to ensure that the density and pressure were correct during these simulations. The box size was fixed at the end of constant pressure equilibration, and, to ensure the correct density, the box volume was set to the average volume from this equilibration period prior to production. This was done using an affine transformation to rescale the coordinates and distances a small amount to obtain the correct volume. Then an additional 100 ps of data was discarded to equilibration before collecting data for analysis. In principle, we are interested in free energies at constant pressure, but we are unsatisfied with the barostat options in GROMACS, and hence chose to perform the simulations at constant volume (ensuring correct densities) and correct for the $p\Delta V$ work.

The confine-and-release cycle was applied to handle some slow conformational changes

As mentioned above, in the TIP3P calculations, we conducted simulations from different starting configurations. Then, as a test for convergence, we compared the hydration free energies and uncertainties for each configuration of a molecule. A few molecules were found to be problematic, in that the hydration free energies calculated from these different conformations differed by more than the computed uncertainties, and hence were not converged. Therefore, we applied the confine-and-release framework [18] to include free energies associated with the conformational changes, which we separately calculated using umbrella sampling, as described below.

There were relatively few cases where hydration free energies depended on starting conformation. Many of these were carboxylic acids, where calculated hydration free energies appeared to vary with the dihedral angle H–O–C–O. Thus, the potential of mean force (PMF) for rotating that dihedral angle was calculated by means of the umbrella sampling, using the dihedral restraints option in GROMACS. Umbrella sampling was done using 24 equally spaced umbrella potentials with centers separated by 15 degrees. Spring constants were 300 kJ mol^{-1} nm^{-2}. Minimization and equilibration protocols were the same as for the hydration calculations, and production simulations for carboxylic acids were 500 ps rather than 5 ns, but otherwise the same.

Analysis of the umbrella sampling was done using the MBAR algorithm [28] to reweight the potentials and compute the free energy differences between neighboring umbrella windows and produce a PMF. This procedure was applied to all of the carboxylic acids examined: acetylsalicylic acid, diflunisal, flurbiprofen, ibuprofen, ketoprofen, and naproxen. We also found that d-xylose has two substantially different (and slow to interconvert) ring conformations that differed in their pucker, resulting in different hydration free energies for simulations beginning from the different conformations. We were able to compute the free energy landscape for altering the pucker by umbrella sampling one of the torsions within the ring, which altered the ring conformation in sufficiently long simulations. To achieve this, we had to increase the length of the production simulations in umbrella sampling to 7.5 ns, and we discarded the first 2 ns to equilibration. These are very long simulations, but this was still faster than sampling the ring pucker in unbiased simulations, where we were not able to achieve adequate sampling even on these timescales. We also repeated these umbrella sampling calculations beginning from an alternate conformation of d-xylose in order to ensure we obtained the same PMF regardless of starting conformation.

For both sets of umbrella sampling simulations, we applied these techniques both in water (TIP3P and TIP4P-Ew) and in vacuum, in order to apply the confine-and-release cycle to compute correct free energies [18]. The results reported below include contributions from these conformational changes unless otherwise noted, as these contributions were included in our predictions.

Results and discussion

Overall agreement with experiment was relatively good for most compounds

As discussed in the Introduction, we hoped that our RMS error for our predictions might be in the 1.3–1.7 kcal/mol range, based on previous data. However, challenging chemical groups could make this RMS error higher. Here, we found that our overall RMS error for the 23 predictions was 2.82 ± 0.08 kcal/mol (TIP3P) and 3.07 ± 0.09 kcal/mol (TIP4P-Ew). Correlation coefficients (R^2) were 0.66 ± 0.03 and 0.63 ± 0.03, respectively.

Including eight additional "explanatory" molecules which were part of the SAMPL2 test set resulted in overall RMS errors of 2.72 ± 0.09 kcal/mol (TIP3P) and 2.99 ± 0.09 kcal/mol (TIP4P-Ew), and R^2 values of 0.81 ± 0.03 and 0.79 ± 0.03, respectively (Figs. 1, 2).

These numbers should be compared with our overall error on the 599 neutral molecules we had studied prior to this SAMPL (2), 1.7 kcal/mol. Hence, performance is substantially worse overall. Note, however, from the figure,

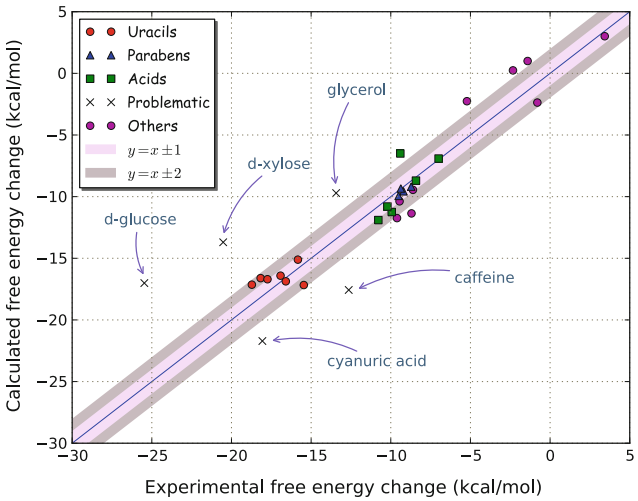

Fig. 1 Computed versus experimental hydration free energies for the TIP3P water model. The *black* $x = y$ *line* denotes perfect agreement with experiment, and *diagonal gray bars* indicate 1 and 2 kcal/mol errors relative to experiment. Several compounds discussed in the text are highlighted, and the remainder are colored by common chemical functionality, if any

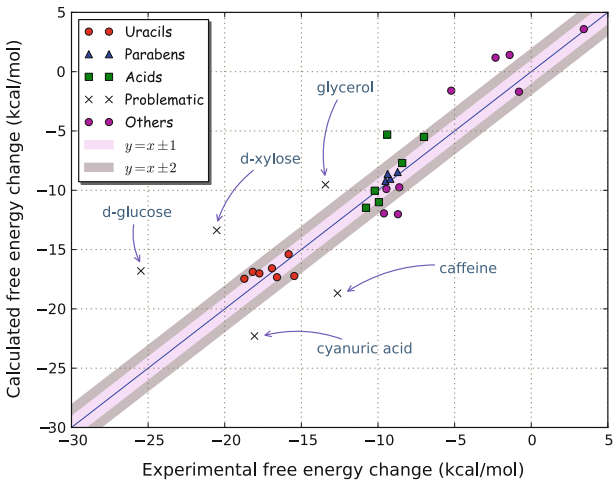

Fig. 2 Computed versus experimental hydration free energies for the TIP4P-Ew water model. The *black* $x = y$ *line* denotes perfect agreement with experiment, and *diagonal gray bars* indicate 1 and 2 kcal/mol errors relative to experiment. Several compounds discussed in the text are highlighted, and the remainder are colored by common chemical functionality, if any

that most of the error is concentrated in just five molecules: three hydroxyl-rich compounds, and cyanuric acid plus caffeine. Figure 6 shows results from SAMPL2 together with all of our previous results. A table of computed hydration free energies is available in the Supporting Information.

It is also interesting to note that the two solvent models perform very similarly here, both in terms of overall performance, but also in terms of agreement with one another: The RMS difference between the TIP3P and TIP4P-Ew results is only 0.64 ± 0.03 kcal/mol.

This is in contrast to our previous work, where we had found that performance of TIP4P-Ew was substantially worse [19]. However, apparently those earlier calculations were affected by a bug in GROMACS that caused errors for free energies computed with four-point water models, so they are being repeated and an erratum is forthcoming. Given the similarity in the results from the two water models applied here, the analysis below will focus on results with just TIP3P.

Overall among participants of SAMPL2, our computed hydration free energies were reasonably good, though there were many participants with overall RMS errors near or even somewhat better than ours. RMS errors ranged from 1.69 up to 12.03 kcal/mol, with most in the range 2–4 kcal/mol. Thus, our approach did not particularly stand out from the pack in terms of RMS error. It was notable, however, that the majority of our error came from just a few molecules (discussed below) and seems to be systematic, suggesting that with rather modest further force field development efforts it may be possible to do much better. A table showing RMS difference between

our method and those of others, as well as comparing RMS errors of different methods, is available in the Supporting Information.

Analysis of problem compounds suggests possible force field limitations

As noted, most of the overall error is caused by just two groups of molecules which were poorly predicted. One is a group of hydroxyl-rich compounds which are all off by a similar amount in the same direction (Fig. 1). One possible source of error is poor convergence. However, we began our hydration free energy calculations for each molecule from several different conformations, and our results from these different conformations generally agree with one another within calculated uncertainty (with the exception of d-xylose, discussed below in section on slow conformational changes).

If the error for these compounds is not caused by a convergence error, then it suggests a possible problem with the force field. This could be any number of things. However, our previous extensive (504 molecule) study [24] analyzed functional groups for potential systematic errors, and found that alcohols were overrepresented in compounds with high errors, suggesting that there might be a systematic error in the description of alcohols. Furthermore, all of the alcohols in that study except one (triacetyl glycerol, which was likely wrong for several other reasons) were off by similar amounts in the same direction (Fig. 3). On average across all alcohols, the mean error per hydroxyl was 1.29 kcal/mol with a standard deviation of 1.00 kcal/mol; without triacetyl glycerol, the mean and standard deviation were 1.44 and 0.34 kcal/mol, respectively.

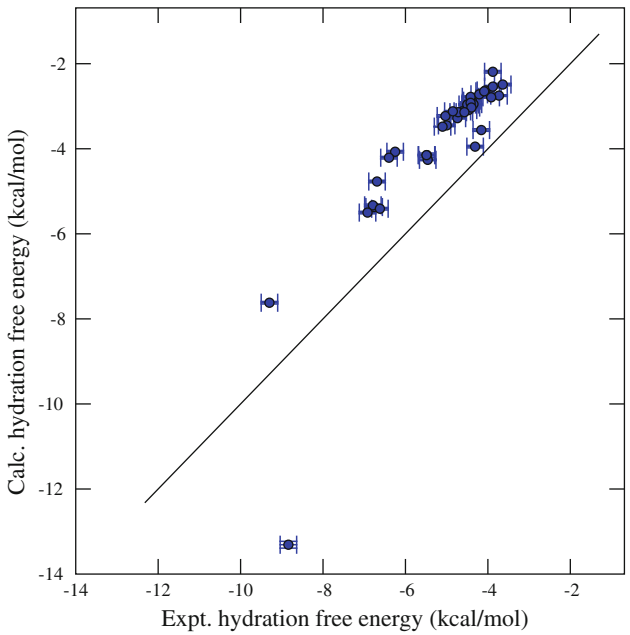

Fig. 3 Computed versus experimental hydration free energies for a set of alcohols from a previous study [24]. Over the whole set, the mean error per hydroxyl is 1.29 kcal/mol with a standard deviation of 1.00 kcal/mol; without triacetyl glycerol, the one compound which is off in the negative direction, the mean and standard deviation were 1.44 and 0.34 kcal/mol, respectively

The hydroxyl-rich compounds here are d-xylose, d-glucose, and glycerol. Might the majority of the error be explained by a systematic error of the kind observed previously for alcohols? Here, the error per hydroxyl for d-glucose, d-xylose, and glycerol, was 1.69, 1.36, and 1.24 kcal/mol, respectively. This is certainly consistent with the 1.44 kcal/mol error per hydroxyl observed previously for alcohols and suggests there may be a systematic error in the force field description of hydroxyls. Without these compounds, the RMS error would be 1.86 ± 0.10 kcal/mol.

The remaining two outliers are cyanuric acid and caffeine, which are both nitrogen-rich heterocycles. On the first of these, the initial result for cyanuric acid provided by the organizers of SAMPL2 disagreed wildly with most of the predictions. However, this was apparently due to a glitch in the analysis of experimental data (J. Peter Guthrie, personal communication) and the correct experimental value for cyanuric acid is −18.06 ± 0.27 kcal/mol, much closer to what we (and many of the other participants) had obtained initially. Cyanuric acid has two major tautomers, however, and the organizers provided only one. Calculations were done on that tautomer, as we currently have no framework for handling changes of tautomeric state. So, calculations were done for the keto form, and our value was −21.71 ± 0.07 kcal/mol, an error of 3.65 ± 0.28 kcal/mol. After the SAMPL2 challenge, we repeated the calculation using the enol form, and obtained −16.50 ± 0.09 kcal/mol, reducing the error to 1.56 ± 0.28 kcal/mol. Part of the error here may be computing the hydration free energies with the wrong tautomeric state. (This, however, may not be the final verdict—the keto form may predominate in solution, though the protonation state is also quite sensitive to the pH of the solution and at basic pH multiple tautomeric states may be relevant [4, 13, 27].) It is also worth pointing out that our approach assumes the molecule stays in one fixed tautomeric state in both gas and water. This may not be the case, in reality—the molecule may undergo a change of tautomeric state on transfer from gas to water, or it may exist as a mix of tautomers in one or both environments, and these would be missed in our approach.

Caffeine, also an outlier, is structurally similar to uracils, which were well predicted. In spite of the structural similarity, the computed free energy is about 5 kcal/mol more negative than the experimental value and we have not yet been able to identify the origin of this error (Fig. 1).

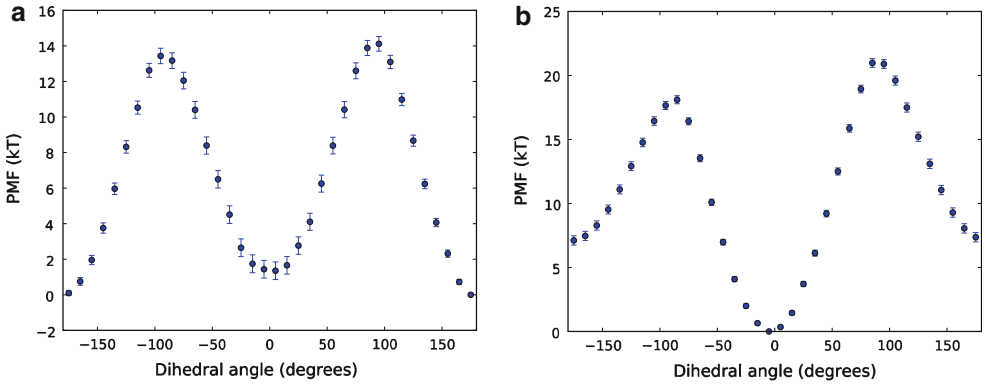

Fig. 4 PMFs for rotating the orientation of the hydroxyl hydrogen around the H–O–C–O torsion in ibuprofen. (**a**) Ibuprofen in water; (**b**) ibuprofen in vacuum. The free energy landscape/PMF is shown (in units of k_BT) for reorienting the hydroxyl in ibuprofen, both in water and in vacuum. The well at 0° is the preferred conformation in vacuum, where the hydrogen points towards the other oxygen, while the alternate conformation is preferred in water

Some conformational changes are quite slow and present sampling challenges

For the carboxylic acids, the experimental hydration free energies are adjusted to measure transfer of the neutral form of the molecule to water (J. Peter Guthrie, personal communication). Hence, we retain the hydroxyl protons on transfer of the molecule from gas to water in our simulations and report transfer free energies for the neutral form from gas to water.

As mentioned above, the calculated hydration free energies for carboxylic acids were found to vary depending on what position the hydroxyl was in during the calculations.

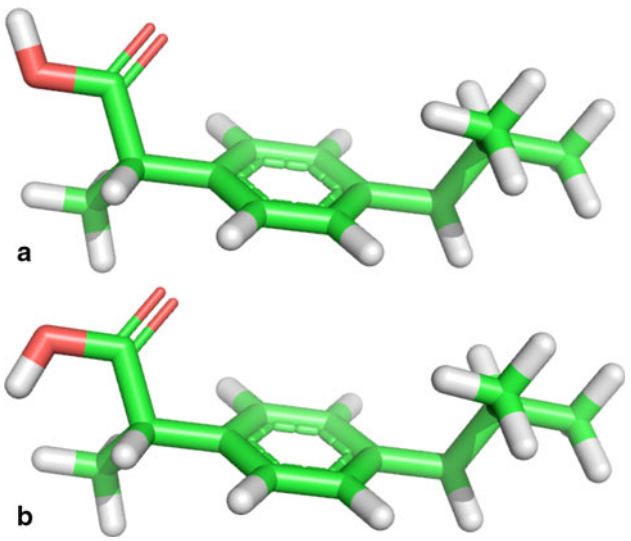

Fig. 5 Illustration of the preferred conformation of the carboxylic acid in vacuum versus water. (**a**) Ibuprofen in vacuum/gas; (**b**) ibuprofen in water. Shown is a *cartoon* illustrating the preferred conformations of the carboxylic acid portion of ibuprofen in vacuum versus in water. This shows only the rotation of the hydroxyl; the rest of the molecule is left in the same conformation

The torsion governing the orientation of the hydroxyl hydrogen (H–O–C–O) has a very large barrier, meaning that over the course of our simulations, the hydrogen either tends to remain pointing towards the other oxygen, or away from it (in what we call the "opposed" conformation) (Fig. 5). Depending on where this hydrogen points, computed free energies change by more than 2 kcal/mol from different starting conformations.

This is due to kinetic trapping in the simulations, and as discussed in the Methods section, we correct for this trapping by computing the free energy landscape (PMF) for reorienting the hydroxyl in gas and in water, and completing a confine-and-release thermodynamic cycle to get correct hydration free energies that include the effects of any conformational changes. When we include these contributions, results are consistent regardless of starting conformation. Thus, the confine-and-release approach was used when making predictions for SAMPL2.

An additional benefit of computing these free energy landscapes is that we gain some insight into the importance of conformational change. Here, we find that the preferred conformation of H–O–C–O group varies depending on environment: In vacuum, carboxylic acids strongly prefer the typical conformation where the hydroxyl hydrogen pointed towards the other oxygen (Fig. 5). But, at least according to the force field, in water there is a slight preference for the opposed conformation, where the hydroxyl hydrogen points away from the oxygen. Whether this slight preference is correct or not, it seems clear that the first conformation should be substantially less favorable in water than in vacuum, simply due to shielding of the strong electrostatic interactions that dominate in water.

Potentials of mean force for ibuprofen in vacuum and water are shown in Fig. 4. Without including these contributions for carboxylic acids, RMS errors would have been up to 2.45 kcal/mol depending on starting conformation for

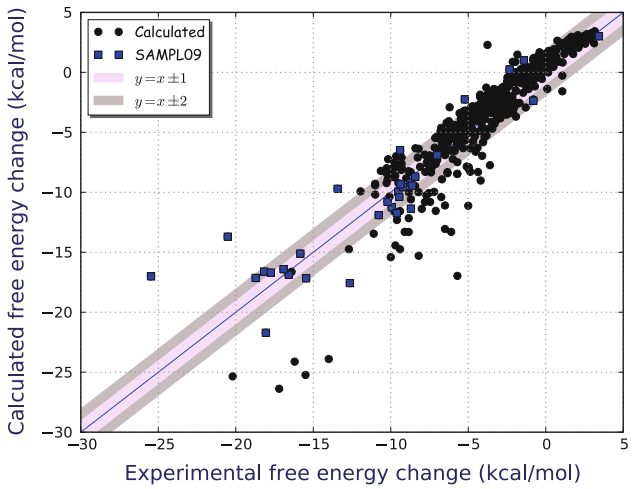

Fig. 6 Computed versus experimental hydration free energies for the TIP3P water model, across all compounds studied to date. The *black* $x = y$ *line* denotes perfect agreement with experiment, and *diagonal gray bars* indicate 1 and 2 kcal/mol errors relative to experiment. *Blue symbols* denote compounds from this SAMPL; *black symbols* are those studied previously [19, 23–25].

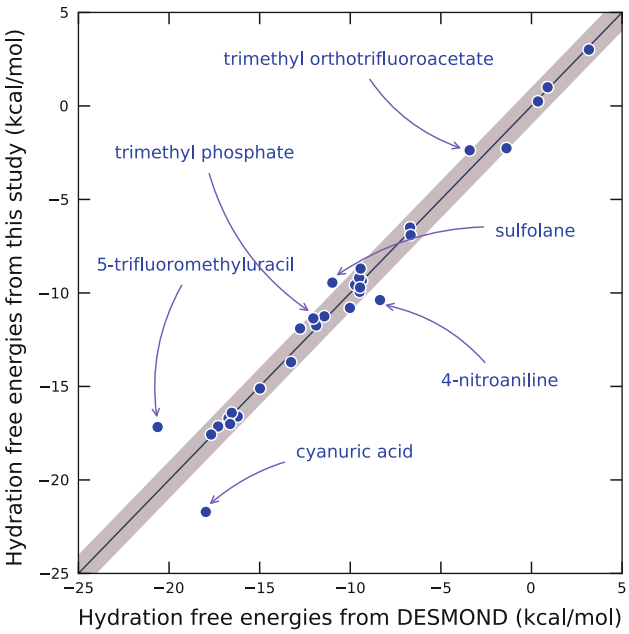

Fig. 7 Hydration free energies computed with DESMOND versus those from this study (GROMACS). Shown are hydration free energies computed using the same force field parameters and TIP3P water model (but slightly different AM1-BCC partial charges) in DESMOND plotted versus those computed here. The diagonal gray line denotes differences of ± 1 kcal/mol

these molecules, while including these contributions, the RMS error across the carboxylic acids was only 1.40 kcal/mol.

Ab initio calculations (on acetic acid) also indicate the conformation of 5(a) is preferred in vacuum, by about 4.3 kcal/mol (7.2 kT), not too far off from the 7 kT observed here for ibuprofen. The ab initio calculations also found the energetic barrier for reorienting the hydroxyl is substantial, in the neighborhood of 13.3 kcal/mol [5].

We also found some slow conformational changes in the ring pucker in d-xylose, and applied the confine-and-release technique to handle these conformational changes as well. Here, there is no change of conformation on transfer between environments, only a change in how strongly the preferred conformation is preferred over the alternate conformation, so this is a relatively minor issue. Still, simulations begun from the wrong ring conformation would have led to biased hydration free energies.

Overall, one take-home message is that even for small molecules, conformational changes can be quite slow. For ibuprofen's carboxylic acid, barriers to the important conformational change are 12–14 k_BT, meaning that enhanced or biased sampling techniques (like those employed here) or *extremely* long simulations are necessary to obtain correct hydration free energies (Fig. 6).

Comparison with another molecular dynamics package and the same force field gives fairly good agreement

We have compared our data with the results from another group (Fig. 7, Huafeng Xu, personal communication) which were using the same force field and charge model (though computed using a different implementation – the Amber 8 version of Antechamber) but performing calculations in different simulation package, DESMOND, from D. E. Shaw Research and Development. Ideally, two sets of calculations using the same force field but different programs would agree within statistical uncertainty. Here, we generally agree quite well, with an RMS difference of only 1.10 ± 0.05 kcal/mol. The fact that the two data sets are consistent with each other indicates that we have reached the point where different groups can generally reproduce one another's results in different simulation packages.

Despite the overall good agreement between the two sets of results, there are some discrepancies for several molecules. AM1-BCC charges were used in both sets of calculations, but these were computed using different implementations and conformations. We repeated our hydration free energy calculations for five molecules for which there were significant differences between the two sets of results: 4-nitroaniline, 5-trifluoromethyluracil, sulfolane, trimethyl orthotrifluoroacetate, and trimethyl phosphate. The differences for these were originally −2.02, 3.48, 1.57, 1.01, and 0.78 kcal/mol. Redoing our GROMACS calculations using the partial charges used in the DESMOND calculations, the differences were −1.12, 0.78, 0.05, 0.96, and 0.70 kcal/mol. So, for the first three of the molecules, the difference in partial charges leads to substantial differences in the results. (Unlike in our work, the

partial charges used for the DESMOND work were computed using the conformations provided by the organizers, probably explaining the slight differences in partial charges and hence the variation in computed free energies). The remaining differences remain to be tracked down.

Differences can, in some cases, be due to statistical fluctuations. But the differences reported here for some compounds are far outside computed uncertainties from both groups. Our calculated uncertainties are designed to represent the standard error in the mean, which measures how precisely (to our best estimate) we know the true hydration free energy for our parameters. The fact that differences are larger than this error suggests that either (a) uncertainties are vastly underestimated, or (b) force field or methodological differences may underly the discrepancies. For most of the molecules here, we favor explanation (b). Uncertainties may be slightly underestimated (as it is always difficult to estimate whether future measurements one is unaware of may affect free energy estimates) but in our experience this underestimation is not vast, except in cases where results are far from converged due to unsampled conformational changes. Here, we used multiple starting conformations for our calculations to help assess convergence. The fact that our results are, by and large, similar beginning from these substantially different conformations suggests that our error estimates may be reasonable and the parameters or methodology may be the source of any discrepancies. This is especially true for relatively rigid molecules like 4-nitroaniline; trimethyl phosphate is relatively flexible and a difference of 0.70 kcal/mol could be due to unsampled conformational changes, though we have no evidence for this.

Methodologies or details of simulation parameters can make a difference, however, as (in principle) could choice of simulation package depending on how algorithms are implemented. It is difficult to assess the potential for variation due to such issues, as it is relatively rare that results from different simulation packages using the same parameters are compared. We are aware of only one previous such comparison [34], where the value of −2.15 kcal/mol was obtained for the hydration free energy of toluene in AMBER, as in previous work with the same parameters in GROMACS [19]. This is consistent with our comparison here with DESMOND—modern simulation packages are getting to the point where many results are reproducible (which makes discrepancies even more interesting). Other simulation parameters, such as Lennard-Jones cutoff [31], can make significant differences in some cases, but are not likely to be the origin of discrepancies here. Different densities can lead to different results as well [31], but care is taken here to ensure correct densities. Other choices, such as the choice of thermodynamic integration (TI) versus BAR, make less difference. For example, we typically analyze all our results with both TI and BAR and check for consistency. Both analysis methods generally agree within computed uncertainties, and when they do not, we follow up to track down any discrepancies. On the other hand, small differences in partial charges, in our experience, can change computed hydration free energies substantially. Anecdotal reports from some others in the field suggest this as well. This further supports our focus on differences in partial charges as a source of discrepancies here.

Conclusions

We predicted the hydration free energies for 31 diverse organic molecules using two solvent models, TIP3P and TIP4P-Ew. These two water models gave similar performance overall, and trends were very similar, suggesting the water model is not the leading source of error in these calculations. Overall RMS errors were somewhat higher than our previous average (2.82 kcal/mol overall here, vs. 1.7 kcal/mol there). Without three hydroxyl-rich compounds, which seem to have a clear systematic error which our data suggests is caused by an error on hydroxyls, our RMS error is only 1.86 kcal/mol.

We also found, while making our predictions, that slow sampling of molecular conformational changes could lead to biased free energies, and introduce substantial errors when conformational changes are important on transfer between environments. It is important to note that conformational sampling can be a challenge even for small molecules.

Somewhat surprisingly, we found that the force field predicts carboxylic acids, in their neutral form, would undergo a conformational change when transferred from vacuum to water. We found that the hydroxyl alters its preferred orientation, pointing towards the other oxygen in vacuum, and away from it in water.

One of our other problem molecules, cyanuric acid, has two different potential tautomers. Initially, the molecule was treated as the keto form (as provided by the organizers), resulting in a hydration free energy very far off from the experimental value. If we instead compute the hydration free energy for the enol tautomer, it is substantially more accurate (an improvement of 5.21 ± 0.11 kcal/mol in the hydration free energy). However, it is not clear if this is the final verdict, as the molecule might well exist in solution as a mix of tautomers.

Overall, these results further support the notion that physical force fields can make accurate predictions of fundamental thermodynamic properties in favorable cases, but that fundamental tests like this on predicting hydration free energies reveal deficiencies that need to be resolved in order for these calculations to become much more accurate.

Acknowledgments We are grateful to Huafeng Xu (D. E. Shaw Research and Development) for providing his DESMOND data, and to Christopher J. Fenell (University of California, San Francisco) for bringing the tautomerism of cyanuric acid to our attention. We are grateful to the Louisiana Optical Network Initiative Institute, supported by the Louisiana Board of Regents Post-Katrina Support Fund Initiative grant LEQSF(2007-12)-ENH-PKSFI-PRS-01 for an allocation of computer time (to DLM) and a graduate fellowship (to PVK). We thank the University of New Orleans for startup funding and for a SCoRE grant.

References

1. Bennett CH (1976) Efficient estimation of free energy differences from Monte Carlo data. J Comp Phys 22:245–268
2. Beutler TC, Mark AE, van Schaik RC, Gerber PR, van Gunsteren WF (1994) Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. Chem Phys Lett 222:529–539
3. Boyce SE, Mobley DL, Rocklin GJ, Graves AP, Dill KA, Shoichet BK (2009) Predicting ligand binding affinity with alchemical free energy methods in a polar model binding site. J Mol Biol 394:747–763
4. Budavari S (eds) (1996) The Merck index: an encyclopedia of chemicals, drugs and biologicals, 12th edn. Merck Research Laboratories
5. Buemi G (2009) DFT study of the hydrogen bond strength and ir spectra of formic, oxalic, glyoxylic and pyruvic acids in vacuum, acetone and water solution. J Phys Org Chem 22:933–947
6. Guthrie JP (2009) A blind challenge for computational solvation free energies: introduction and overview. J Phys Chem B 113(14):4501–4507
7. Hess B, van der Vegt NFA (2006) Hydration thermodynamic properties of amino acid analogues: a systematic comparison of biomolecular force fields and water models. J Phys Chem B 110:17616–17626
8. Horn HW, Swope WC, Pitera JW, Madura JD, Dick TJ, Hura GL, Head-Gordon T (2004) Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. J Chem Phys 120(20):9665
9. Jakalian A, Bush BL, Jack DB, Bayly CI (2000) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. method. J Comput Chem 21:132–146
10. Jakalian A, Jack DB, Bayly CI (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. parameterization and validation. J Comput Chem 23:1623–1641
11. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935
12. Klamt A, Eckert F, Diedenhofen M (2009) Prediction of the free energy of hydration of a challenging set of pesticide-like compounds. J Phys Chem B 113:4508–4510
13. Klotz IM, Askounis T (1947) Absorption spectra and tautomerism of cyanuric acid, melamine and some related compounds. J Am Chem Soc 69:801–803
14. Llinàs A, Glen RC, Goodman JM (2008) Solubility challenge: Can you predict solubilities of 32 molecules using a database of 100 reliable measurements?. J Chem Inf Model 48:1289–1303
15. Maccallum JL, Tieleman DP (2003) Calculation of the water-cyclohexane transfer free energies of neutral amino acid side-chain analogs using the OPLS all-atom force field. J Comp Chem 24(15):1930–1935
16. Marenich AV, Cramer CJ, Truhlar DG (2009) Performance of SM6, SM8, and SMD on the SAMPL-1 test set for the prediction of small-molecule solvation free energies. J Phys Chem B 113:4438–4543
17. Mobley DL, Chodera JD, Dill KA (2006) On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. J Chem Phys 125:084902
18. Mobley DL, Chodera JD, Dill KA (2007a) Confine-and-release method: Obtaining correct binding free energies in the presence of protein conformational change. J Chem Theory Comput 3:1231–1235
19. Mobley DL, Dumont È, Chodera JD, Dill KA (2007b) Comparison of charge models for fixed-charge force fields: small-molecule hydration free energies in explicit solvent. J Phys Chem B 111:2242–2254
20. Mobley DL, Graves AP, Chodera JD, McReynolds A, Shoichet BK, Dill KA (2007c) Predicting absolute ligand binding free energies to a simple model site. J Mol Biol 371:1118–1134
21. Mobley DL, Barber II AE, Fennell CJ, Dill KA (2008a) Charge asymmetries in hydration of polar solutes. J Phys Chem B 112:2405–2414
22. Mobley DL, Chodera JD, Dill KA (2008b) Entropy and conformational change in implicit solvent simulations of small molecules. J Phys Chem B 112:938–946
23. Mobley DL, Bayly CI, Cooper MD, Dill KA (2009a) Predictions of hydration free energies from all-atom molecular dynamics simulations. J Phys Chem B 113:4533–4537
24. Mobley DL, Bayly CI, Cooper MD, Shirts MR, Dill KA (2009b) Small molecule hydration free energies in explicit solvent: An extensive test of fixed-charge atomistic simulations. J Chem Theory Comput 5(2):350–358
25. Nicholls A, Mobley DL, Guthrie JP, Chodera JD, Pande VS (2008) Predicting small-molecule solvation free energies: a blind challenge test for computational chemistry. J Med Chem 51:769–778
26. Nicholls A, Wlodek S, Grant JA (2009) The sample solvation challenge: Further lessons regarding the pitfalls of parametrization. J Phys Chem B 113:4521–4532
27. Sancier KM, Brady AP, Lee WW (1964) Absorption spectra of solutions of cyanuric acid and its chlorinated derivatives. Spectrochimica Acta 20:397–403
28. Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. J Chem Phys 129:124105
29. Shirts MR, Pande VS (2005) Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. J Chem Phys 122:134508
30. Shirts MR, Mobley DL, Chodera JD (2007a) Alchemical free energy calculations: ready for prime time? Ann Rep Comp Chem 3:41–69
31. Shirts MR, Mobley DL, Chodera JD, Pande VS (2007b) Accurate and efficient corrections for missing dispersion interactions in molecular simulations. J Phys Chem B 111:13052–13063
32. Shirts MR, Mobley DL, Brown SP (2010) Free energy calculations in structure-based drug design. In: Kenneth M, Merz DR, Reynolds CH (eds) Structure based drug design, Cambridge University Press, New York (forthcoming)
33. van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) Gromacs: fast, flexible, and free. J Comp Chem 26:1701–1718
34. Steinbrecher T, Mobley DL, Case DA (2008) Nonlinear scaling schemes for Lennard-Jones interactions in free energy calculations. J Chem Phys 127:214108
35. Sulea T, Wanapun D, Dennis S, Purisima EO (2009) Prediction of SAMPL-1 hydration free energies using a continuum electrostatics-dispersion model. J Phys Chem B 113:4511–4520
36. Tan C, Tan YH, Luo R (2007) Implicit nonpolar solvent models. J Phys Chem B 111:12263–12274

37. Tetko IV, Tanchuk VY, Villa AE (2001) Prediction of n-octanol water partition coefficients from PHYSPROP database using artificial neural networks and e-state indices. J Chem Inform Comput Sci 41(5):1407–1421
38. Villa A, Mark AE (2002) Calculation of the free energy of solvation for neutral analogs of amino acid side chains. J Comp Chem 23(5):548–553
39. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general AMBER force field. J Comput Chem 25:1157–1174
40. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. J of Mol Graphics Modell 26:247–260