

Approaches for calculating solvation free energies and enthalpies demonstrated with an update of the FreeSolv database

Guilherme Duarte Ramos Matos,[†] Daisy Y. Kyu,[‡] Hannes H. Loeffler,[¶] John D. Chodera,[§] Michael R. Shirts,^{||} and David L. Mobley^{*,⊥}

Department of Chemistry, University of California, Irvine, Department of Pharmaceutical Sciences, University of California, Irvine, Scientific Computing Department, STFC, UK, Computational and Systems Biology Program, Sloan Kettering Institute, Department of Chemical and Biological Engineering, University of Colorado Boulder, Boulder, and Departments of Pharmaceutical Sciences and Chemistry, University of California, Irvine

E-mail: dmobley@uci.edu

Abstract

Solvation free energies can now be calculated precisely from molecular simulations, providing a valuable test of the energy functions underlying these simulations. Here, we briefly review “alchemical” approaches for calculating the solvation free energies of small, neutral organic molecules from molecular simulations, and illustrate by applying them to calculate aqueous solvation free energies (hydration free energies). These

*To whom correspondence should be addressed

[†]University of California, Irvine

[‡]University of California, Irvine

[¶]STFC Daresbury, Warrington, WA4 4AD, United Kingdom

[§]Memorial Sloan Kettering Cancer Center, New York

^{||}University of Colorado, Boulder

[⊥]University of California, Irvine

approaches use a non-physical pathway to compute free energy differences from a simulation or set of simulations and appear to be a particularly robust and general-purpose approach for this task. We also present an update (version 0.5) to our FreeSolv database of experimental and calculated hydration free energies of neutral compounds and provide input files in formats for several simulation packages. This revision to FreeSolv provides calculated values generated with a single protocol and software version, rather than the heterogeneous protocols used in the prior version of the database. We also further update the database to provide calculated enthalpies and entropies of hydration and some experimental enthalpies and entropies, as well as electrostatic and nonpolar components of solvation free energies.

Introduction

Solvation free energies give the free energy change associated with the transfer of a molecule between ideal gas and solvent at a certain temperature and pressure. While solvation free energies (ΔG^{solv}) in general, and hydration free energies (ΔG^{hyd} , solvation in water) in particular might not seem to have far reaching implications, in fact, researchers in diverse areas can benefit from their prediction, because such solvation free energies are related to a broad range of physical properties such as infinite dilution activity coefficients, Henry's law constants, solubilities, and distribution of chemical species between immiscible solvents or different phases. Solvation free energies (ΔG^{solv})—especially hydration free energies (ΔG^{hyd} , aqueous solvation free energies)—are related to a broad range of physical properties such as infinite dilution activity coefficients, Henry's law constants, solubilities, and distribution of chemical species between immiscible solvents or different phases.

Solvation free energies are differences in thermodynamic potentials which describe the relative populations of a chemical species in solution and gas phase at equilibrium.^{1,2} In the thermodynamic limit in the solvated phase and the ideal gas limit in the gas phase, ΔG^{solv} of component i is equal to $\mu_{i,solv} - \mu_{i,gas}$, the difference in chemical potentials in the two

phases. In the additional limit of one molecule of component i at infinite dilution, these become the infinite dilution excess chemical potentials in the respective solvents.

Solvation free energies not only tell us how much a molecule prefers one phase over another, but they also can provide insight into how solvent behaves in different environments. For example, water solvates molecules of opposite polarity differently, due to its inherent asymmetry,³ surfaces also have asymmetric effects on ion pairing which depend on the curvature of the surface,⁴ and molecular geometry and chemical environment affects hydrophobic solvation.⁵ Although they can be difficult to measure experimentally, ΔG^{sol} and ΔG^{hyd} can be calculated to a precision better than $0.4 \text{ kJ}\cdot\text{mol}^{-1}$, even with a relatively modest investment of simulation time, for relatively diverse small neutral molecules⁶ such as those seen in the FreeSolv database of hydration free energies⁷ and in recent blind challenges such as the Statistical Assessment of the Modeling of Protein and Ligands (SAMPL) challenges. These challenges aim to improve the quality of predictive computational tools in drug design,^{1,6,8-21} and have leveraged solvation free energies to help drive improvements in modeling.

Since the solvation free energy of neutral compounds is an aggregate measure of many competing interactions and entropic effects that can span many kJ/mol, comparison of computed solvation free energies to experiment has proven to be an exacting test of force field quality that has been useful in revealing deficiencies in small molecule force fields.^{3,22,23} The relative ease by which solvation free energies can be calculated – as opposed to protein-ligand binding free energies, which are fraught with a variety of sampling issues – also makes them attractive for this purpose^a. For instance, SAMPL has frequently (in SAMPL1 through SAMPL4) included blind predictions of hydration free energies in particular.^{1,8-14} However, to our knowledge, no laboratories are currently measuring hydration free energies, leading the field to search for other simple physical properties that can be rapidly computed – such as relative solubilities,²⁴ distribution coefficients,²⁵ and solvation free energies in organic

^aBut see the Supporting Information for how protonation state/tautomer challenges may apply here, as in protein-ligand binding.

solvents²⁶ – as a tool to assess and improve small molecule force fields. In computational chemistry, hydration free energies are of particular importance because they are frequently used in force field parameterization^{26–29} and in the testing of free energy methods and force fields.^{1,8–14,30–37}

Solvation free energies are often calculated by alchemical free energy methods,³⁸ which simulate a series of alchemical intermediates to compute the free energy of transferring a solute from solution to gas phase (as here) or vice versa. This alchemical path provides an efficient way to move the solute from solution to the gas phase by perturbing its interactions in a non-physical way. Since free energy is path-independent, this non-physical process still yields the free energy change for transfer of the solute from solvent to gas.^{38,39} The path is formed by constructing intermediate states with interactions that modulate between the end states of interest, with the variable λ parameterizing progress along the path. A particularly efficient set of intermediate states uses a two step process, first turning off the van der Waals interactions using one parameter λ_v , and another turning off the electrostatic interactions using a second λ_e . Here, we compute the free energy change to transition between each pair of λ values, and the overall free energy change is the sum of these pairwise differences.

While other approaches have been used to calculate solvation free energies,⁴⁰ alchemical free energy calculations using explicit solvent have become a mainstream approach,^{41,42} in part because of their formal rigor. Alternative approaches include implicit solvent models,^{34–37,43} which yield ΔG^{hyd} but do not take into consideration solvent configuration around the solute, and Monte Carlo based approaches using the Gibbs ensemble^{44–50} and expanded ensemble,⁵¹ though these are most commonly used for molecules that are particularly small and/or rigid.

Hydration and solvation free energies have a range of applications

The activity coefficient γ_i of a solute species i can be calculated from ΔG^{solv} :

$$\gamma_i = \exp\left(\frac{\mu_i^{excess}}{RT}\right) = \exp\left(\frac{\Delta G_i^{solv}}{RT}\right) \quad (1)$$

where μ_i^{excess} is the excess chemical potential of i and is equal to ΔG_i^{solv} in the ideal gas limit of the vapor phase, R is the universal gas constant and T , the absolute temperature. For instance, solvation free energies are used to estimate infinite dilution activity coefficients (γ_i^∞) in many solvents by using a single molecule of solute i .⁵²⁻⁵⁸ Experimental results obtained from gas chromatography^{59,60} can be compared to γ_i^∞ obtained from ΔG^{solv} to further test models and methodologies that use these free energy calculations.

Solubility prediction is another field where $\Delta G^{solv/hyd}$ prediction can have great value. One methodology computes the free energy of solubility by computing both the sublimation free energy (from solid to gas) and hydration free energy (from gas to water).⁶¹ Another way to predict molecular solid solubilities depends on excess chemical potential calculations. The chemical potential, μ , of a species is calculated at different concentrations to build the concentration-dependent chemical potential curve of solutions⁶²⁻⁶⁵ in order to discover phase equilibrium conditions. Free energies of solvation in pure melts and pure amorphous matter have been used to find upper bounds for solubilities given that most drug-like compounds have crystal polymorphs.⁶⁶⁻⁶⁹ Relative solubilities of a given chemical species between different solvents can also be assessed with these calculations.^{24,70} Henry's law solubility constants⁷¹ and solubilities in supercritical fluids⁷² can also be predicted using solvation free energies.

The latest SAMPL challenge, SAMPL5, included blind prediction of distribution coefficients between cyclohexane and water for 53 solutes.^{32,33,73,74} Distribution and partition

coefficients are important properties for toxicology and pharmacology because they play a major part in predicting absorption and distribution of a substance in different tissues.⁷⁵ Partition coefficients – which are the distribution coefficients of the neutral form of a compound – can be estimated from the difference between solvation free energies of the chemical species in two different solvents,²¹ as shown in equation 2:

$$\log_{10} P_{A \rightarrow B} = \frac{\Delta G^{solv,A} - \Delta G^{solv,B}}{RT \ln(10)} \quad (2)$$

where $\Delta G^{solv,A}$ and $\Delta G^{solv,B}$ are the solvation free energies of a molecule in solvents A and B , respectively. While in principle, the calculation could be done by transferring the solute between phases, in many software implementations it is more straightforward to simply compute the solvation free energy in each phase separately, or the free energy of removing the solute from each phase. Thus, solvation free energy calculations have found relatively widespread application in calculating partition coefficients, including in SAMPL5.¹⁵⁻²¹ Hydration free energies themselves are valuable quantities in drug design^{42,76} and can be used to understand the impact desolvation of the ligand has in the binding process^{77,78} or can be utilized as QSAR descriptors.⁷⁹

Theory and practical aspects of alchemical calculations

Solvation free energies can be calculated in various ways. In this paper we focus in alchemical free energy calculations, which have been one of the most consistently reliable methods in recent applications such as the SAMPL series of challenges.^{1,8-14,25} Consider a pair of end states A and B , and their respective Hamiltonians $H_A(\mathbf{q}, \mathbf{p}; \lambda)$ and $H_B(\mathbf{q}, \mathbf{p}; \lambda)$.

$$H(\mathbf{q}, \mathbf{p}; \lambda) = f(\lambda)H_A(\mathbf{q}, \mathbf{p}; \lambda) + g(\lambda)H_B(\mathbf{q}, \mathbf{p}; \lambda) \quad (3)$$

where $f(\lambda)$ and $g(\lambda)$ are functions of λ used to mix the Hamiltonians, typically set such that $H = H_A$ at $\lambda = 0$. With $H(\mathbf{q}, \mathbf{p}; \lambda)$ we can calculate the free energy difference between A and B :

$$\Delta G = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (4)$$

This method, called thermodynamic integration (TI),⁸⁰ performs similarly to more robust methods when the integrand is smooth.^{81–83} However, it can break down when the integrand is not smooth, and it can be difficult to capture numerical integration errors in resulting uncertainty estimates.

Exponential averaging (EXP), also known as Free Energy Perturbation (FEP), was introduced by Zwanzig.⁸⁴ In this method, the free energy difference between two states A and B is given by:

$$\Delta G = \frac{1}{\beta} \ln \langle e^{-\beta[H_B(\mathbf{q}, \mathbf{p}; \lambda) - H_A(\mathbf{q}, \mathbf{p}; \lambda)]} \rangle_A \quad (5)$$

where $\beta = (k_B T)^{-1}$. Although equation 5 is exact in the limit of large numbers of samples, EXP is inefficient and prone to shot noise and other large biases when configurations sampled in one state are very unlikely to be found in the other state, and vice-versa. The probability that describes this likelihood is called the phase-space overlap between the two states. EXP convergence is far from ideal, requiring states to have sizable phase-space overlap with one another.^{38,82,85} Thus, addition of intermediate states (with values of λ between 0 and 1) can improve overlap dramatically and thus the quality of the final result.⁸⁶ Another issue is an asymmetric bias depending on which direction the free energy difference extrapolation is performed,^{87,88} so other analysis methods are now preferred.³⁸ In the limit of adequate sampling, EXP converges to the same free energy value in both directions, but there are other ways to calculate free energies more efficiently.

An alternate method, Bennett's acceptance ratio (BAR), uses the information from both

directions to derive the following relationship:

$$\sum_{i=1}^{n_i} \frac{1}{1 + \exp(\ln(n_i/n_j) + \beta\Delta H_{ij} - \beta\Delta G)} - \sum_{j=1}^{n_j} \frac{1}{1 + \exp(\ln(n_j/n_i) + \beta\Delta H_{ji} + \beta\Delta G)} = 0 \quad (6)$$

where n_i and n_j are the number of samples between each state and ΔH_{ij} is the enthalpy difference between these states.

This expression minimizes the free energy variance⁸⁹ and makes BAR much more efficient than EXP.^{87,88} The Multistate Bennett acceptance ratio (MBAR) is an extension of BAR that considers the overlap between a given state and all the others in the path between the end states.⁹⁰ BAR and MBAR perform similarly when the spacing between intermediate states is moderate⁸³ and therefore only neighboring states have phase-space overlap. WHAM^{91,92} is essentially an approximation to MBAR, and thus also gives very similar results when carefully done with appropriately small bins. MBAR performs consistently well, and indeed is perhaps the most consistently well-performing free energy estimator,⁸³ thus we recommend it as the analysis method of choice whenever possible. In many cases, TI can perform as well as BAR or MBAR if the integrand is smooth.^{81,82} EXP should generally be avoided due to its asymmetric bias and sensitivity to the tails of the distribution.⁸⁷

Choice of alchemical pathway

Alchemical free energy calculations were given this name because the pathway involves unphysical changes to the atomic identities, such as to the interactions between components.^{42,93} Solvation free energy calculations can use several different approaches to modulating interactions. One approach, called *decoupling*, modulates only the interactions between the solute and its surroundings, retaining internal interactions (which is the approach we use here). An alternative approach, called *annihilation*, removes internal non-bonded interactions within the solute as well as those with the surroundings. Mixtures of the two approaches are also possible, such as annihilating internal electrostatic interactions while decoupling non-polar

interactions. Here, three main states are considered: a single, non-interacting molecule of the solute in a box of solvent; the solute molecule that interacts with its surroundings through nonpolar (dispersion and exclusion) forces; and a fully interacting system, in which solvent molecules interact with the solute molecule through both electrostatic and nonpolar (dispersion and exclusion) forces. Simulations are then conducted over a series of intermediates connecting these states: going through a phase which changes electrostatic interactions only, and another phase which modifies van der Waals interactions only (figure 1). Each of these intermediates has high configuration space overlap with at least neighboring states, allowing precise calculation of free energy differences.⁹⁴⁻⁹⁷

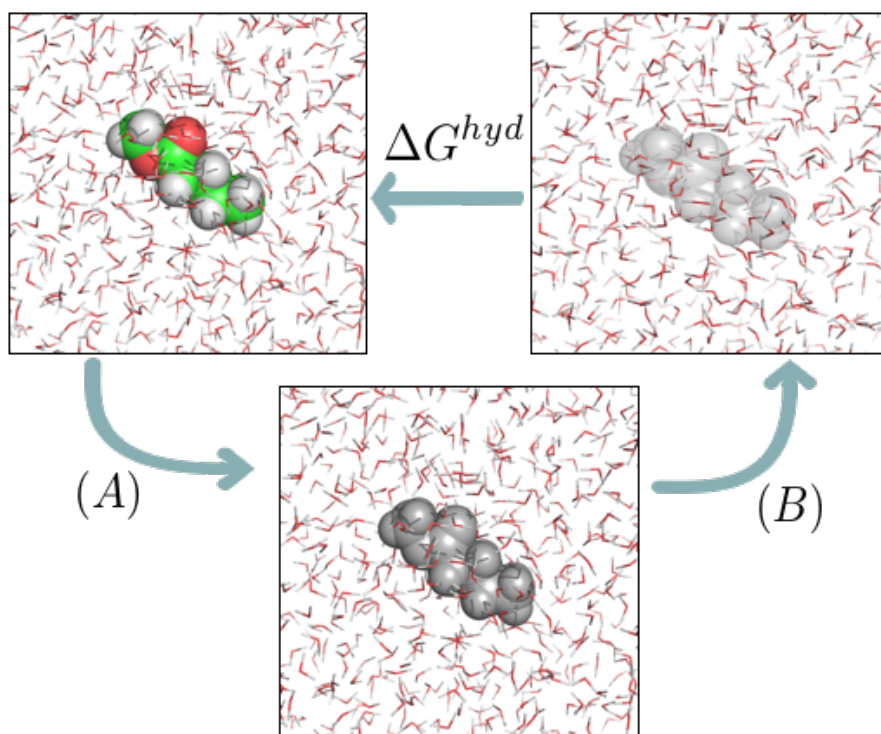


Figure 1: Thermodynamic cycle used to calculate hydration free energies (or, more generally, solvation free energies). In (A), we have states in which charge-charge interactions between the solute and its environment are progressively turned off. In (B) dispersion interactions between solute and water are progressively turned off. Colored atoms (green for carbon, red for oxygen, white for hydrogen) have electrostatic and nonpolar interactions with the environment; gray atoms retain only nonpolar interactions; and transparent atoms have no interactions with their environment (and thus represent the solute in vacuum).

The most straightforward way to switch between states is the linear pathway

$$H(\mathbf{q}, \mathbf{p}; \lambda) = (1 - \lambda)H_A(\mathbf{q}, \mathbf{p}) + \lambda H_B(\mathbf{q}, \mathbf{p}) \quad (7)$$

but this pathway is in general problematic for solvation of all but the smallest molecules. This is because repulsive forces are often handled by a $1/r^{12}$ term (such as in the Lennard-Jones functional form) which leads to non-integrable singularities in $\langle \partial H / \partial \lambda \rangle$ at terminal λ values due to sudden changes in the potential at small r . This is not a problem which is specific to TI; rather, this issue can still result in numerical instabilities or large errors in calculated free energies even with other analysis approaches.^{42,98,99} Thus, more complicated λ pathways are required, such as soft-core potentials, which should in general be used to avoid such numerical problems.^{94,98,99} A common soft-core form for Lennard-Jones potential between two particles i and j is:

$$U_{ij}^{LJ}(r_{ij}, \lambda) = \lambda^n 4\epsilon_{ij} \left(\frac{1}{[\alpha(1 - \lambda)^m + (r_{ij}/\sigma_{ij})^6]^2} - \frac{1}{\alpha_{LJ}(1 - \lambda)^m + (r_{ij}/\sigma_{ij})^6} \right) \quad (8)$$

where ϵ_{ij} and σ_{ij} are the Lennard-Jones parameters and α is a positive constant which should typically be set to 0.5.^{99,100} The exponents m and n are most efficient at $n = 1$ and $m = 1$ but other values have been used too.^{96,99-101} Improvements have been achieved by new soft-core functions that ease the problem with additional minima within the formulation of the original soft-core potential,¹⁰² and alternate potentials that construct near optimal paths for alchemical simulations.¹⁰³ Linear basis functions can be used as an alternative to soft-core potentials that approaches the minimum variance possible over all pair potentials;^{97,104} these can also enhance the efficiency of alchemical calculations."

The use of soft-core potentials promotes better convergence in many circumstances, and provides much lower variance free energy estimates given a fixed amount of simulation time,^{94,96,98,100,103} thus their use is highly recommended for successful free energy calculations. Without soft-core potentials, convergence is much more difficult or nearly impossible

to achieve in many types of solvation free energy calculations.

Considerations for successful alchemical calculations

The accuracy of these calculations is affected by at least three factors:^{105,106} Is our sampling representative and adequate? Is the free energy estimator good enough? Is the force field adequate for the system? For solvation free energies of small molecules in solvents with relatively fast dynamics, such as water, sampling is typically adequate with a few nanoseconds of dynamics per lambda window (at least for relatively rigid solutes), and the free energy estimators above are robust when applied carefully.

However, when designing new studies, it is still important to choose robustly performing estimators and ensure adequate sampling. As discussed above, we recommend MBAR as the best and most reliable general-purpose estimator.⁸³ Sampling remains a critical issue,^{105,107} both as the solute size and flexibility grows and as solvent dynamics or environment become heterogeneous, for example, for solvation free energies in octanol which can form local clusters of hydrophilic and hydrophobic sites,²¹ or in mixed solvents.²⁵

We updated FreeSolv, the free community solvation free energy database

FreeSolv⁷ is a hydration free energy database for neutral^b compounds that contains experimental and calculated hydration free energy values, SMILES strings, PubChem compound IDs, IUPAC names, and now (as of version 0.5, presented in this work) calculated enthalpies and entropies of hydration of 643 small organic molecules. Since experimental and calculated hydration free energies, ΔG^{hyd} , can be computed quite precisely for quantitative comparison, FreeSolv can provide information for force field development,²⁶⁻²⁹ and can assist the testing of new solvation free energy methods.^{108,109} While calculated hydration free energies for

^bFor additional discussion of why the focus is on neutral compounds, see the Supporting Information.

all compounds have been in FreeSolv since the database was constructed,⁷ previous values had been calculated with somewhat heterogeneous protocols in a variety of different studies spread over roughly 10 years.^{2,6,11,13,23,41,110,111} In this work, we have updated FreeSolv by repeating all of the calculations under a single unified protocol, and additionally obtained enthalpies and entropies of hydration.

We obtained FreeSolv's calculated hydration free energies using alchemical free energy calculations, connecting the end states (corresponding to the solute in vacuum and in solution) via a λ path with 20 intermediate states (full details in SI). The first five states corresponded to changes in electrostatic interactions, while the last 15 modified the Lennard-Jones terms in the potential. This separation allows electrostatic interactions to be changed linearly, and soft-core potentials to be used only when changing non-polar interactions.⁹⁷ We ran 5 nanoseconds of Langevin dynamics per state with 2 femtosecond time steps in GROMACS 4.6.7¹¹²⁻¹¹⁷ at 298.15K. Pressure was maintained at 1 atm by the Parrinello-Rahman barostat.¹¹⁸ Full details can be found in the supporting materials.

Input files for version 0.5 of FreeSolv were constructed from scratch from the isomeric SMILES strings for the compounds which are deposited in the database. From these SMILES strings, we used the OpenEye Python toolkits¹¹⁹⁻¹²¹ to generate molecular structures and assign AM1-BCC partial charges,^{122,123} then charged mol2 files were written out. The AMBER Antechamber package was then used to assign parameters from the GAFF²⁰ small molecule force field (version 1.7), and these were then converted to GROMACS format and solvated with the TIP3P water model.¹²⁴ The script which performs the setup and regenerates all input and molecular structure files in the database is available in the scripts directory of FreeSolv and provides full details. Following the calculations, MBAR hydration free energies were obtained using `alchemical-analysis.py`.⁹³ Additional details can be found in the supporting material.

Computed hydration free energies are compared with experiment in figure 2.

In the calculations described in this study, we found an average error of 1.3 ± 0.3 kJ·mol⁻¹,

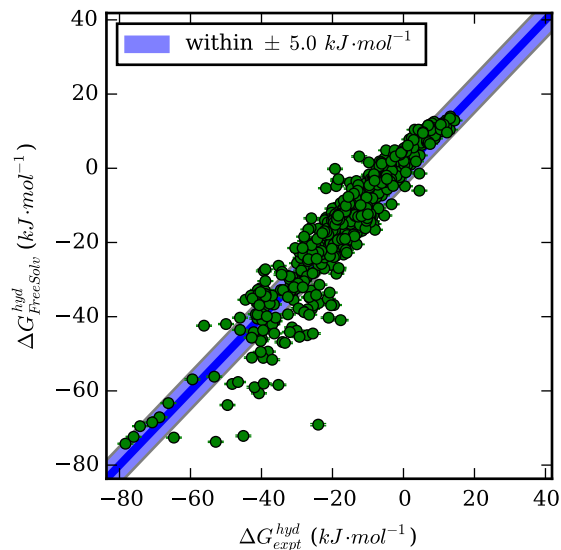


Figure 2: Calculated versus experimental hydration free energies for the compounds in FreeSolv version 0.5.

RMS error of $6.4 \pm 0.3 \text{ kJ}\cdot\text{mol}^{-1}$, average absolute error of $4.7 \pm 0.2 \text{ kJ}\cdot\text{mol}^{-1}$, Kendall τ of 0.80 ± 0.01 , and Pearson R of 0.933 ± 0.008 , comparable to those in the original FreeSolv set,⁷ though some individual compounds have reasonably significant discrepancies (see SI). This level of accuracy is consistent with what is often seen from classical fixed-charge force fields, which typically yield RMS errors around 4-8 kJ/mol in computed hydration free energies.⁴²

In addition to experimental and calculated values, FreeSolv now includes the free energy of decoupling the solute-solvent electrostatic interactions (ΔG^q) and the free energy of decoupling the nonpolar interactions in water (ΔG^{vdW}) (available at github.com/mobleyleab/FreeSolv). These quantities have been used for various purposes, including to assist in the study, development, and test of implicit solvent models.^{125,126} However, it is important to remember that these components come from our particular decomposition of the free energy,¹²⁷⁻¹³⁰ and are not state functions; other decompositions are possible, so considerable care needs to be taken in interpreting these components. For example, annihilation rather than decoupling of Coulomb interactions would result in somewhat different decompositions due to electrostatics-induced conformational differences while van der Waals interactions are

being decoupled.

In addition to hydration free energies, we have also computed enthalpies (ΔH^{hyd}) and entropies of hydration (ΔS^{hyd}), and have added these to the database. Enthalpies of transfer, due to their larger dynamic range and lack of compensating entropic effects, are generally more sensitive to force field parameters than free energies,^{131–133} and thus can be sensitive probes of force field accuracy, providing an additional point of comparison to experiments. While only a few hydration enthalpies are available experimentally, there are a sufficient number to note that significant discrepancies between experiment and computed values exist for some compounds (SI Figure 2 and SI Table 1). There seems to be little correlation between compounds which have accurate hydration free energies versus which have accurate hydration enthalpies; for example, the calculated hydration free energy of benzene is within error of the experimental value, but the enthalpy is off by approximately 12 kJ/mol. In contrast, the hydration free energy of cyclohexanol is off by more than 5 kJ/mol but the enthalpy is within error of the experimental value. Thus, clearly these quantities yield different information. We postulate that, should more experimental enthalpies of hydration become available, this information may provide additional insight into force field deficiencies.

To compute hydration free energies via simulation, we used a difference in potential energies between a water box solvating the compound and a neat water box with the compound removed to vacuum:

$$\Delta H^{hyd} = \langle U_{solution} \rangle - (\langle U_{water} \rangle + \langle U_{vacuo} \rangle) \quad (9)$$

Here, $\langle U_{solution} \rangle$ is the internal energy of the solution (containing the solute); $\langle U_{water} \rangle$ is the internal energy of a box of the same number of water molecules (under the same conditions) without the solute; and $\langle U_{vacuo} \rangle$ is the internal energy of the solute molecule alone in vacuum. We have neglected the pressure-volume contribution to the enthalpies, $P\Delta V$, since for solutes of this size, the contribution is much smaller than our typical uncertainties of $\approx 2.9 \text{ kJ} \cdot \text{mol}^{-1}$

;¹³¹ at larger pressures or for larger solutes than in this set, this term could become significant. Notably, this scheme also omits other contributions that may be relevant in making direct comparison with experimental enthalpies of hydration, including contributions from the cost of polarizing the molecule from vacuum to solvated phase charges (relevant to fixed-charge force fields), corrections to the vibrational modes due to the quantum chemical nature of real solutes, nonideality of the gas phase, and the fact that the simulation of the liquid is carried out at atmospheric pressure rather than at the vapor pressure of the gas phase; for a review of these contributions, see.¹³⁴ We note that other groups have also omitted these contributions, which still await a thorough assessment of relative magnitude for small molecule hydration enthalpies.¹³¹

To calculate the average potential energies, we ran 60 nanosecond Langevin dynamics simulations, with two femtosecond timesteps at 298.15 K and 1 atm in water and in vacuo for each molecule in the database. These long simulations were necessary to reduce error bars on the computed enthalpies to levels around $2.9 \text{ kJ} \cdot \text{mol}^{-1}$, roughly the level of typical thermal energy ($1 k_B T$) as done in, for example, host-guest binding calculations.¹³² In order to obtain consistent results, we used simulation boxes with 1,309 water molecules and one solute molecule. The same system parameters and water model were used as in the free energy calculations. Hydration entropies are calculated via the equation:

$$\Delta S = \frac{\Delta H - \Delta G}{T} \quad (10)$$

with ΔG and ΔH calculated as described previously.

Calculated hydration enthalpies exhibit some correlation with calculated hydration free energies, but the correlation is not perfect, indicating that enthalpies can indeed provide additional constraints on the force field.¹³² The Kendall τ and the Pearson R between the calculated ΔH^{hyd} and ΔG^{hyd} respectively were 0.76 ± 0.02 and 0.943 ± 0.005 (see supplementary information).

Our conclusion that enthalpies can provide an additional constraint on the force field is further supported by comparison to experimental data. Specifically, 11 experimental hydration enthalpies and entropies from ORCHYD, a database of experimental hydration properties,¹³⁵ were added to FreeSolv. Calculated and experimental enthalpies have a Kendall τ of 0.77 ± 0.05 , and a Pearson R of 0.87 ± 0.03 (see SI). These values indicate that the computed hydration free energies are relatively predictive of experimental values, though there is also clear room for improvement. Calculated hydration enthalpies and their experimental counterparts show significant differences that are not observed in the plot of experimental versus calculated free energies of the same 11 compounds, suggesting (as in previous studies¹³¹) that enthalpies provide additional information on the thermodynamics and constraints on the force field (though as noted above, additional enthalpy corrections may be needed¹³⁴). More details can be found in the supporting information.

We also partitioned the hydration enthalpy, ΔH , into two components: a solvent interaction term and a conformational change term, ΔH_{int}^{hyd} and ΔH_{conf}^{hyd} , respectively, in order to understand how much the solvation enthalpy is influenced by the solute conformation, and how much solute conformation is modulated by solvation. We obtained the solvent interaction component by taking the average energy of the solute in water and subtracting off the solute internal energy and the energy of a corresponding box of pure water, leaving only the enthalpy change due to changing solute-solvent interactions and solvent reorganization:

$$\Delta H_{int}^{hyd} = \langle U_{solv} \rangle_s - \langle U_{vac} \rangle_s - \langle U_{water} \rangle_w \quad (11)$$

where $\langle U_{solv} \rangle_s$ is the average potential energy over the original solvated trajectory, $\langle U_{vac} \rangle_s$ is the average potential energy of the solute molecule in the solvated trajectory after removing its water molecules, and $\langle U_{water} \rangle_w$ is the average potential of a simulation with an equivalent box with pure water. ΔH_{int}^{hyd} thus corresponds to the change in solvation enthalpy due to transferring a solute molecule from vacuum to water with a *fixed set of configurations* (as

given by the solvated trajectory) – i.e., it treats the solute as if there is no conformational change going from gas to water, so it includes only changes in solvent structure and solute-solvent interactions.

The conformational change component of the enthalpy is obtained by taking the change in solute internal energy on going from gas to water, which we can evaluate as follows:

$$\Delta H_{conf}^{hyd} = \langle U_{vac} \rangle_s - \langle U_{vac} \rangle_v \quad (12)$$

where $\langle U_{vac} \rangle_v$ is the potential energy of the solute molecule in vacuum evaluated from the trajectory run in vacuum, and $\langle U_{vac} \rangle_s$ is the potential energy of the solute molecule in vacuum evaluated from the trajectory run in solvent (after stripping the solvent molecules). ΔH_{conf}^{hyd} thus gives to the enthalpy change due to solute conformational changes on solvation; these occur because interactions with water can stabilize configurations that are not common in vacuum. If a compound's distribution of configurations is unchanged on transfer to solvent, ΔH_{conf}^{hyd} will be zero. It can trivially be verified that these components still sum to the total enthalpy change:

$$\Delta H^{hyd} = \Delta H_{solv}^{hyd} + \Delta H_{conf}^{hyd} \quad (13)$$

These components, while certainly not a unique decomposition of the total enthalpy, do provide a way to intuitively understand one important set of contributions to the enthalpy of hydration in a way which provides some insight into changes undergone by the solute and environment. For example, solutes which undergo significant conformational changes on solvation may tend to have a large change in the conformational component of the hydration enthalpy (fig. 3). This happens because solutes that make hydrogen bonds with water or have strong internal electrostatic interactions in the gas phase can assume conformations that were energetically unfavorable *in vacuo* when solvated.

Our intention is that FreeSolv serve as an updateable, extensible community resource. While it already covers a large number of molecules, we would be delighted to include input

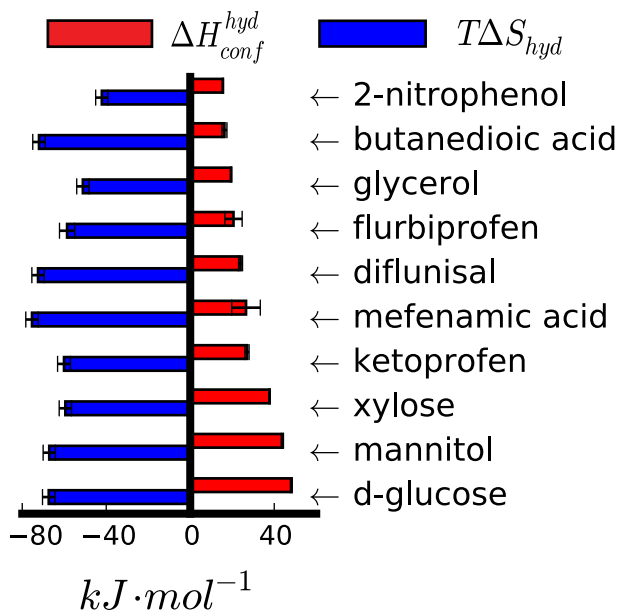


Figure 3: Conformational enthalpies and associated entropies of compounds with highest ΔH_{conf}^{hyd} .

files and calculated values from other force fields and/or methods so it can further serve as a benchmark of methods, simulation packages, and so on. Additionally, while hydration free energy data is not abundant, certainly at least some data is available that is not presently included in FreeSolv, so community contributions of experimental data with references will be appreciated. Additional curation of the experimental data already present is likely needed – for example, much of the experimental data still needs to be tracked back to its original source material rather than literature compilations of data which are currently cited. FreeSolv is available on GitHub at <http://github.com/mobleyleab/FreeSolv> and contributions are welcomed there.

Conclusions

Solvation free energies have been the subject of considerable scientific interest for many years because they are related to a large number of physical properties. Here, we have provided a short review of alchemical methods for computing solvation free energies of small organic

molecules, and discussed their application to hydration free energies. Solvation free energies for such molecules can be calculated precisely and effectively using alchemical free energy calculations, as described here, preferably using MBAR as the free energy estimator.

We also introduced an update to FreeSolv⁷ (v0.5), a database of calculated and experimental hydration free energies, enthalpies and entropies. The database was designed to be easily incorporated into automated workflows: we provide IUPAC names, PubChem compound IDs and SMILES strings, as well as topology and coordinate files, but additional data is welcome. Additionally, we provide calculated and experimental free energy values that can be used to assist method and force field development. Unfortunately, experimental hydration enthalpies and entropies are not available for every compound.

Calculated free energies show reasonable agreement with experimental values (fig. 2) with an RMS error around $6 \text{ kJ} \cdot \text{mol}^{-1}$ and an average error close to $1 \text{ kJ} \cdot \text{mol}^{-1}$. With the aid of ORCHYD,¹³⁵ we were able to extend FreeSolv to contain experimental hydration enthalpies for a few (11) compounds for the first time. We observe significant errors for hydration enthalpies that are much larger than those for hydration free energies, so further investigation will be needed. This result also suggests that enthalpies can be used as additional constraints in force field development.

We believe that this update of FreeSolv will assist future efforts in force field development and development and testing of new methods. We also hope that FreeSolv's new features help serve the scientific community, and provide a valuable resource the community will help extend.

Acknowledgement

DLM and GDRM appreciate the financial support from the National Science Foundation (CHE 1352608), and computing support from the UCI GreenPlanet cluster, supported in part by NSF Grant CHE-0840513. GDRM appreciates support from the Brazilian agency CAPES - Science without Borders program (BEX 3932-13-3). JDC acknowledges partial

support from NIH grant P30 CA008748. HHL is supported through an EPSRC provided SLA, funding the core support of CCPBioSim. CCPBioSim is the Collaborative Computational Project for Biomolecular Simulation funded by EPSRC grants EP/J010588/1 and EP/M022609/1. We particularly appreciate Kyle Beauchamp (Counsyl, South San Francisco, CA) and Lee-Ping Wang (UC Davis) for input on curation of the FreeSolv database. We also thank Gaetano Calabrò and Caitlin Bannan for their support and assistance, and sharing of knowledge.

DLM is a member of the Scientific Advisory Board for OpenEye Scientific Software. JDC is a member of the Scientific Advisory Board for Schrödinger, LLC.

Supporting Information Available

The following files are available free of charge.

- Supplement.pdf: document containing correlation plots between calculated free energies, enthalpies and entropies.
- GROMACS 4.6.7 .mdp files: GROMACS input files containing all the details of the simulations.
- FreeSolv can be obtained free of charge at <http://github.com/mobleylab/FreeSolv>.

This material is available free of charge via the Internet at <http://pubs.acs.org/>.

References

- (1) Guthrie, J. P. *The Journal of Physical Chemistry B* **2009**, *113*, 4501–4507.
- (2) Klimovich, P. V.; Mobley, D. L. *Journal of Computer-Aided Molecular Design* **2010**, *24*, 307–316.
- (3) Mobley, D. L.; Baker, J. R.; Barber, A. E.; Fennell, C. J.; Dill, K. A. *The Journal of Physical Chemistry. B* **2008**, *112*, 2405–2414.

- (4) Chorny, I.; Dill, K. A.; Jacobson, M. P. *The Journal of Physical Chemistry. B* **2005**, *109*, 24056–24060.
- (5) Harris, R. C.; Pettitt, B. M. *Proceedings of the National Academy of Sciences of the United States of America* **2014**, *111*, 14681–14686.
- (6) Nicholls, A.; Mobley, D. L.; Guthrie, J. P.; Chodera, J. D.; Bayly, C. I.; Cooper, M. D.; Pande, V. S. *Journal of Medicinal Chemistry* **2008**, *51*, 769–779.
- (7) Mobley, D. L.; Guthrie, J. P. *Journal of Computer-Aided Molecular Design* **2014**, *28*, 711–720.
- (8) Geballe, M. T.; Skillman, A. G.; Nicholls, A.; Guthrie, J. P.; Taylor, P. J. *Journal of Computer-Aided Molecular Design* **2010**, *24*, 259–279.
- (9) Geballe, M. T.; Guthrie, J. P. *Journal of Computer-Aided Molecular Design* **2012**, *26*, 489–496.
- (10) Gallicchio, E.; Levy, R. M. *Journal of Computer-Aided Molecular Design* **2012**, *26*, 505–516.
- (11) Mobley, D. L.; Liu, S.; Cerutti, D. S.; Swope, W. C.; Rice, J. E. *Journal of Computer-Aided Molecular Design* **2012**, *26*, 551–562.
- (12) Gallicchio, E.; Deng, N.; He, P.; Wickstrom, L.; Perryman, A. L.; Santiago, D. N.; Forli, S.; Olson, A. J.; Levy, R. M. *Journal of Computer-Aided Molecular Design* **2014**, *28*, 475–490.
- (13) Mobley, D. L.; Wymer, K. L.; Lim, N. M.; Guthrie, J. P. *Journal of Computer-Aided Molecular Design* **2014**,
- (14) Muddana, H. S.; Sapra, N. V.; Fenley, A. T.; Gilson, M. K. *Journal of Computer-Aided Molecular Design* **2014**, *28*, 277–287.

- (15) Michel, J.; Orsi, M.; Essex, J. W. *The Journal of Physical Chemistry B* **2008**, *112*, 657–660.
- (16) Genheden, S. *Journal of Chemical Theory and Computation* **2016**, *12*, 297–304.
- (17) Garrido, N. M.; Queimada, A. J.; Jorge, M.; Macedo, E. A.; Economou, I. G. *Journal of Chemical Theory and Computation* **2009**, *5*, 2436–2446.
- (18) Garrido, N. M.; Economou, I. G.; Queimada, A. J.; Jorge, M.; Macedo, E. A. *AIChE Journal* **2012**, *58*, 1929–1938.
- (19) Yang, L.; Ahmed, A.; Sandler, S. I. *Journal of Computational Chemistry* **2013**, *34*, 284–293.
- (20) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *Journal of Computational Chemistry* **2004**, *25*, 1157–1174.
- (21) Bannan, C. C.; Calabrò, G.; Kyu, D. Y.; Mobley, D. L. *Journal of Chemical Theory and Computation* **2016**, *12*, 4015–4024.
- (22) Shirts, M. R.; Pitner, J. W.; Swope, W. C.; Pande, V. S. *The Journal of Chemical Physics* **2003**, *119*, 5740–5761.
- (23) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. *The Journal of Physical Chemistry. B* **2007**, *111*, 2242–2254.
- (24) Liu, S.; Cao, S.; Hoang, K.; Young, K. L.; Paluch, A. S.; Mobley, D. L. *Journal of Chemical Theory and Computation* **2016**, *12*, 1930–1941.
- (25) Bannan, C. C.; Burley, K. H.; Chiu, M.; Shirts, M. R.; Gilson, M. K.; Mobley, D. L. *Journal of Computer-Aided Molecular Design* **2016**,
- (26) Zhang, J.; Tuguldur, B.; van der Spoel, D. *Journal of Chemical Information and Modeling* **2015**, *55*, 1192–1201.

- (27) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. *Journal of Computational Chemistry* **2004**, *25*, 1656–1676.
- (28) Fennell, C. J.; Wymer, K. L.; Mobley, D. L. *Journal of Physical Chemistry B* **2014**, *118*, 6438–6446.
- (29) Jämbeck, J. P. M.; Lyubartsev, A. P. *The Journal of Physical Chemistry B* **2014**, *118*, 3793–3804.
- (30) Bosisio, S.; Mey, A. S. J. S.; Michel, J. *Journal of Computer-Aided Molecular Design* **2016**, 1–10.
- (31) Harwood, D. B.; Peters, C. J.; Siepmann, J. I. *Fluid Phase Equilibria* **2016**, *407*, 269–279.
- (32) Klamt, A.; Eckert, F.; Reinisch, J.; Wichmann, K. *Journal of Computer-Aided Molecular Design* **2016**, 1–9.
- (33) Rustenburg, A. S.; Dancer, J.; Lin, B.; Feng, J. A.; Ortwine, D. F.; Mobley, D. L.; Chodera, J. D. *bioRxiv* **2016**, 063081.
- (34) Nicholls, A.; Honig, B. *Journal of Computational Chemistry* **1991**, *12*, 435–445.
- (35) Grant, J. A.; Pickup, B. T.; Nicholls, A. *Journal of Computational Chemistry* **2001**, *22*, 608–640.
- (36) Knight, J. L.; Brooks, C. L. *Journal of Computational Chemistry* **2011**, *32*, 2909–2923.
- (37) Yang, P.-K. *Chemical Physics* **2016**, *472*, 229–240.
- (38) Shirts, M. R.; Mobley, D. L.; Chodera, J. D. *Annual Reports in Computational Chemistry* **2007**, *3*, 41–59.
- (39) Shirts, M. R. *Methods in Molecular Biology*; 2012; Vol. 819; pp 425–467.

- (40) Skyner, R. E.; McDonagh, J. L.; Groom, C. R.; van Mourik, T.; Mitchell, J. B. O. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6174–6191.
- (41) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Dill, K. A. *The Journal of Physical Chemistry B* **2009**, *113*, 4533–4537.
- (42) Shirts, M. R.; Mobley, D. L.; Brown, S. P. *Drug Design* **2010**, 61–86.
- (43) Cramer, C. J.; Truhlar, D. G. *Accounts of Chemical Research* **2008**, *41*, 760–768.
- (44) Panagiotopoulos, A.; Quirke, N.; Stapleton, M.; Tildesley, D. *Molecular Physics* **1988**, *63*, 527–545.
- (45) Chen, B.; Siepmann, J. I. *Journal of the American Chemical Society* **2000**, *122*, 6464–6467.
- (46) Anderson, K. E.; Siepmann, J. I. *Developments and Applications in Solubility*; IUPAC, Oxford, 2007; pp 171–187.
- (47) Martin, M. G.; Siepmann, J. I. *Theoretical Chemistry Accounts* **1998**, *99*, 347–350.
- (48) McGrath, M. J.; Kuo, I.-F. W.; Ngouana W, B. F.; Ghogomu, J. N.; Mundy, C. J.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G.; Siepmann, J. I. *Physical Chemistry Chemical Physics* **2013**, *15*, 13578–13585.
- (49) Huang, Y.-L.; Miroshnichenko, S.; Hasse, H.; Vrabec, J. *International Journal of Thermophysics* **2009**, *30*, 1791.
- (50) Zhang, L.; Siepmann, J. I. *Theoretical Chemistry Accounts* **2006**, *115*, 391–397.
- (51) Lyubartsev, A. P.; Jacobsson, S. P.; Sundholm, G.; Laaksonen, A. *The Journal of Physical Chemistry B* **2001**, *105*, 7775–7782.
- (52) Slusher, J. T. *Fluid Phase Equilibria* **1998**, *153*, 45–61.

- (53) da Silva, E. F. *Fluid Phase Equilibria* **2004**, *221*, 15–24.
- (54) Tong, C.; Clegg, S. L.; Seinfeld, J. H. *Atmospheric Environment* **2008**, *42*, 5459–5482.
- (55) Gerber, R. P.; Soares, R. d. P. *Industrial & Engineering Chemistry Research* **2010**, *49*, 7488–7496.
- (56) Ge, M.-L.; Deng, X.-M.; Zhang, L.-H.; Chen, J.-Y.; Xiong, J.-M.; Li, W.-H. *The Journal of Chemical Thermodynamics* **2014**, *77*, 7–13.
- (57) Ge, M.-L.; Lu, C.-Y.; Liu, X.-Y.; Li, X.-B.; Chen, J.-Y.; Xiong, J.-M. *The Journal of Chemical Thermodynamics* **2015**, *91*, 279–285.
- (58) Ge, M.-L.; Zhang, Q.; Li, S.-N.; Li, Y.-J.; Zhang, X.-Z.; Mu, Z. *The Journal of Chemical Thermodynamics* **2016**, *102*, 303–309.
- (59) Krummen, M.; Gruber, D.; Gmehling, J. *Industrial & Engineering Chemistry Research* **2000**, *39*, 2114–2123.
- (60) Martínez, R.; Sanz, M. T.; Beltrán, S.; Corcuera, E. *Journal of Chemical & Engineering Data* **2012**, *57*, 1480–1485.
- (61) Schnieders, M. J.; Baltrusaitis, J.; Shi, Y.; Chattree, G.; Zheng, L.; Yang, W.; Ren, P. *Journal of Chemical Theory and Computation* **2012**, *8*, 1721–1736.
- (62) Ferrario, M.; Ciccotti, G.; Spohr, E.; Cartailier, T.; Turq, P. *The Journal of Chemical Physics* **2002**, *117*, 4947–4953.
- (63) Sanz, E.; Vega, C. *The Journal of Chemical Physics* **2007**, *126*, 014507.
- (64) Aragoes, J. L.; Sanz, E.; Vega, C. *The Journal of Chemical Physics* **2012**, *136*, 244508.
- (65) Benavides, A. L.; Aragoes, J. L.; Vega, C. *The Journal of Chemical Physics* **2016**, *144*, 124504.

- (66) Westergren, J.; Lindfors, L.; Höglund, T.; Lüder, K.; Nordholm, S.; Kjellander, R. *The Journal of Physical Chemistry. B* **2007**, *111*, 1872–1882.
- (67) Lüder, K.; Lindfors, L.; Westergren, J.; Nordholm, S.; Kjellander, R. *The Journal of Physical Chemistry. B* **2007**, *111*, 1883–1892.
- (68) Lüder, K.; Lindfors, L.; Westergren, J.; Nordholm, S.; Kjellander, R. *The Journal of Physical Chemistry. B* **2007**, *111*, 7303–7311.
- (69) Lüder, K.; Lindfors, L.; Westergren, J.; Nordholm, S.; Persson, R.; Pedersen, M. *Journal of Computational Chemistry* **2009**, *30*, 1859–1871.
- (70) Paluch, A. S.; Parameswaran, S.; Liu, S.; Kolavennu, A.; Mobley, D. L. *The Journal of Chemical Physics* **2015**, *142*, 044508.
- (71) Olsen, R.; Kvamme, B.; Kuznetsova, T. *Fluid Phase Equilibria* **2016**, *418*, 152–159.
- (72) Noroozi, J.; Ghotbi, C.; Sardroodi, J. J.; Karimi-Sabet, J.; Robert, M. A. *The Journal of Supercritical Fluids* **2016**, *109*, 166–176.
- (73) Genheden, S.; Essex, J. W. *Journal of Computer-Aided Molecular Design* **2016**,
- (74) Chung, K.-C.; Park, H. *Journal of Computer-Aided Molecular Design* **2016**,
- (75) Bodor, N.; Buchwald, P. *Retrometabolic Drug Design and Targeting*; John Wiley & Sons, Inc., 2012; pp 9–38.
- (76) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. *Current Opinion in Structural Biology* **2011**, *21*, 150–160.
- (77) Deng, Y.; Roux, B. *The Journal of Physical Chemistry B* **2009**, *113*.
- (78) Michel, J.; Essex, J. W. *Journal of Computer-Aided Molecular Design* **2010**, *24*, 639–658.

- (79) Zafar, A.; Reynisson, J. *Molecular Informatics* **2016**, *35*, 207–214.
- (80) Kirkwood, J. G. *Journal Of Chemical Physics* **1935**, *3*, 300–313.
- (81) Ytreberg, F. M.; Swendsen, R. H.; Zuckerman, D. M. *Journal of Chemical Physics* **2006**, *125*, 184114.
- (82) Shirts, M. R.; Pande, V. S. *The Journal of Chemical Physics* **2005**, *122*, 144107.
- (83) Paliwal, H.; Shirts, M. R. *Journal of Chemical Theory and Computation* **2011**, *7*, 4115–4134.
- (84) Zwanzig, R. *The Journal of Chemical Physics* **1954**, *22*, 1420–1426.
- (85) Lu, N.; Singh, J. K.; Kofke, D. A. *The Journal of Chemical Physics* **2003**, *118*, 2977–2984.
- (86) Wu, D.; Kofke, D. A. *Journal of Chemical Physics* **2005**, *123*.
- (87) Wu, D.; Kofke, D. A. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics* **2004**, *70*, 066702.
- (88) Wu, D.; Kofke, D. A. *Journal of Chemical Physics* **2005**, *123*.
- (89) Bennett, C. H. *Journal of Computational Physics* **1976**, *22*, 245–268.
- (90) Shirts, M. R.; Chodera, J. D. *The Journal of Chemical Physics* **2008**, *129*, 124105.
- (91) Ferrenberg, A. M.; Swendsen, R. H. *Physical Review Letters* **1989**, *63*, 1195–1198.
- (92) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *Journal of Computational Chemistry* **1992**, *13*, 1011–1021.
- (93) Klimovich, P. V.; Shirts, M. R.; Mobley, D. L. *Journal of Computer-Aided Molecular Design* **2015**, *29*, 397–411.

- (94) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *The Journal of Chemical Physics* **1994**, *100*, 9025–9031.
- (95) Wang, J.; Hou, T.; Xu, X. *Current Computer - Aided Drug Design* **2006**, *2*, 287–306.
- (96) Steinbrecher, T.; Mobley, D. L.; Case, D. A. *The Journal of Chemical Physics* **2007**, *127*, 214108.
- (97) Naden, L. N.; Shirts, M. R. *Journal of Chemical Theory and Computation* **2015**, *11*, 2536–2549.
- (98) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chemical Physics Letters* **1994**, *222*, 529–539.
- (99) Steinbrecher, T.; Joung, I.; Case, D. A. *Journal of Computational Chemistry* **2011**, *32*, 3253–3263.
- (100) Pham, T. T.; Shirts, M. R. *The Journal of Chemical Physics* **2011**, *135*, 034114.
- (101) Shirts, M. R.; Pande, V. S. *The Journal of Chemical Physics* **2005**, *122*, 134508.
- (102) Gapsys, V.; Seeliger, D.; de Groot, B. L. *Journal of Chemical Theory and Computation* **2012**, *8*, 2373–2382.
- (103) Pham, T. T.; Shirts, M. R. *The Journal of Chemical Physics* **2012**, *136*, 124120.
- (104) Naden, L. N.; Pham, T. T.; Shirts, M. R. *Journal of Chemical Theory and Computation* **2014**, *10*, 1128–1149.
- (105) Christ, C. D.; Fox, T. *Journal of Chemical Information and Modeling* **2014**, *54*, 108–120.
- (106) Hansen, N.; Van Gunsteren, W. F. *Journal of Chemical Theory and Computation* **2014**, *10*, 2632–2647.

- (107) Mobley, D. L. *Journal of Computer-Aided Molecular Design* **2012**, *26*, 93–5.
- (108) Wang, B.; Wei, G. W. *The Journal of Chemical Physics* **2015**, *143*, 134119.
- (109) Nessler, I. J.; Litman, J. M.; Schnieders, M. J. *Physical Chemistry Chemical Physics* **2016**, *18*, 30313–30322.
- (110) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. *Journal of Chemical Theory and Computation* **2009**, *5*, 350–358.
- (111) Mobley, D. L.; Dill, K. A.; Chodera, J. D. *The Journal of Physical Chemistry B* **2008**, *112*, 938–946.
- (112) Berendsen, H. J. C.; Van Der Spoel, D.; van Drunen, R. *Computational Physics Communications* **1995**, *91*, 43–56.
- (113) Lindahl, E.; Hess, B.; van der Spoel, D. *Journal of Molecular Modeling* **2001**, *7*, 306–317.
- (114) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *Journal of Computational Chemistry* **2005**, *26*, 1701–1718.
- (115) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *Journal of Chemical Theory and Computation* **2008**, *4*, 435–447.
- (116) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. *Bioinformatics (Oxford, England)* **2013**, *29*, 845–854.
- (117) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. *SoftwareX* **2015**, *1-2*, 19–25.
- (118) Parrinello, M.; Rahman, A. *Physical Review Letters* **1980**, *45*, 1196–1199.
- (119) OpenEye Scientific Software, I. *OEChem*; 2010.

- (120) OpenEye Scientific Software, I. *QUACPAC 1.7.0.2*.
- (121) Hawkins, P. C. D.; Nicholls, A. *Journal of Chemical Information and Modeling* **2012**, *52*, 2919–2936.
- (122) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *Journal of Computational Chemistry* **2000**, *21*, 132–146.
- (123) Jakalian, A.; Jack, D. B.; Bayly, C. I. *Journal of Computational Chemistry* **2002**, *23*, 1623–1641.
- (124) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *The Journal of Chemical Physics* **1983**, *79*, 926–935.
- (125) Gallicchio, E.; Levy, R. M. *Journal of Computational Chemistry* **2004**, *25*, 479–499.
- (126) Gallicchio, E.; Paris, K.; Levy, R. M. *Journal of Chemical Theory and Computation* **2009**, *5*, 2544–2564.
- (127) Boresch, S.; Archontis, G.; Karplus, M. *Proteins: Structure, Function, and Bioinformatics* **1994**, *20*, 25–33.
- (128) Zacharias, M.; Straatsma, T. P. *Molecular Simulation* **1995**, *14*, 417–423.
- (129) Dill, K. A. *Journal of Biological Chemistry* **1997**, *272*, 701–704.
- (130) Levy, R. M.; Gallicchio, E. *Annual Reviews Physical Chemistry* **1997**, *49*, 701–704.
- (131) Hess, B.; van der Vegt, N. F. A. *The Journal of Physical Chemistry B* **2006**, *110*, 17616–17626.
- (132) Henriksen, N. M.; Fenley, A. T.; Gilson, M. K. *Journal of Chemical Theory and Computation* **2015**, *11*, 4377–4394.

- (133) Fenley, A. T.; Muddana, H. S.; Gilson, M. K. *Proceedings of the National Academy of Sciences* **2012**, *109*, 20006–20011.
- (134) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *The Journal of Chemical Physics* **2004**, *120*, 9665–9678.
- (135) Plyasunova, N. V.; Plyasunov, A. V.; Shock, E. L. *International Journal of Thermodynamics* **2004**, *25*, 351–360.

Graphical TOC Entry

