UNIVERSITY OF CALIFORNIA,
IRVINE

Free Energy Calculations in Action: Theory, Applications and Challenges of Solvation Free Energies

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Chemistry

by

Guilherme Duarte Ramos Matos

Dissertation Committee:
Professor David L. Mobley, Chair
Professor Ioan Andricioaei
Professor Thomas Poulos

2018

DEDICATION

To my grandparents, Ana and José, who passed away while I was living in a foreign land.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to acknowledge assistance and financial support for the work presented here on a by-chapter basis:

- Chapter 2: I appreciate the support from the National Science Foundation (CHE 1352608) and the Brazilian agency CAPES - Science without Borders program (BEX 3932-13-3), and computing support from the UCI GreenPlanet cluster, partly supported by NSF Grant CHE-0840513. I thank Hannes Loeffler, Michael Shirts, and John Chodera for the intense scientific debate, and Daisy Kyu for being a good mentee.

- Chapter 3: I appreciate the support from the National Science Foundation (CHE 1352608) and the Brazilian agency CAPES - Science without Borders program (BEX 3932-13-3), and computing support from the UCI GreenPlanet cluster, partly supported by NSF Grant CHE-0840513. I would like to thank OpenEye Software, Gaetano Calabrò, Christopher Bayly, Anthony Nicholls and the Orion team, especially Quinn Bailey and Craig Bruce.

- Chapter 4: I appreciate the support from the National Science Foundation (CHE 1352608) and the Brazilian agency CAPES - Science without Borders program (BEX 3932-13-3), and computing support from the UCI GreenPlanet cluster, partly supported by NSF Grant CHE-0840513. I thank Stefano Bosisio, Dongyukh Suh, Hannes Loeffler, Julien Michel and Benoit Roux for the debates and for the collaborative effort.

- Chapter 5: I appreciate the support from the National Science Foundation (CHE 1352608) and the Brazilian agency CAPES - Science without Borders program (BEX 3932-13-3), and computing support from the UCI GreenPlanet cluster, partly supported by NSF Grant CHE-0840513.

I would like to thank Prof. Mobley for his support, mentorship, and patience as well as his wife, Maura, for her hospitality on Thanksgiving. Prof. Mobley taught me the importance of being organized, and I will try to do my best to follow his example. I thank each one of my fellow lab mates for their friendship and fruitful scientific debates, especially Caitlin Bannan, Gaetano Calabrò, Léa el Khoury and Sukanya Sasmal for the adventures around the Orange County. Many thanks to my great friends Ilya V. Vinogradov, Mayukh Banik, Kevin Rothi, Geoffrey Richards for our adventures in the wilderness of California. A special thanks to Ilya for five years of friendship, support, patience and a lot of scientific debate. My eternal gratitude to my great new friend Arshia Montasser-Kouhssari for the companionship, philosophical discussions, and shared interest in classical art, culture and the little pleasures of life.

The text of Chapter 2 is a minimally modified reprint of the material as it appears in Guilherme D. R. Matos, Daisy Y. Kyu, Hannes H. Loeffler, John D. Chodera, Michael R. Shirts, David L. Mobley. Approaches for Calculating Solvation Free Energies and Enthalpies

Demonstrated with an Update of the FreeSolv Database. *Journal of Chemical & Engineering Data* **62**(5), 1559-1569.

The text of Chapter 3 is a minimally modified reprint of the material as it appears in Guilherme D. R. Matos, Gaetano Calabro, David L. Mobley. Infinite Dilution Activity Coefficients as Constraints for Force Field Parameterization and Method Development. *ChemRxiv*, preprint published under CC-BY that might later be published by ACS.

The text of Chapter 4 is a minimally modified reprint of the material as it appears in Hannes H. Loeffler, Stefano Bosisio, Guilherme D. R. Matos, Donghyuk Suh, Benoit Roux, David L. Mobley, Julien Michel. Reproducibility of Free Energy Calculations Across Different Molecular Simulation Software. *Journal of Chemical Theory and Computation* **14**(11), 5567-5582.

The text of Chapter 5 is a minimally modified reprint of the material as it appears in Guilherme D. R. Matos, David L. Mobley. Challenges in the use of atomistic simulations to predict solubilities of drug-like molecules. *F1000Research* 2018, 7:686.

CURRICULUM VITAE

Guilherme Duarte Ramos Matos

EDUCATION

Doctor of Philosophy in Chemistry — **2013-2018**
University of California, Irivne — *Irvine, CA*

Master of Science in Chemistry — **2011-2013**
Universidade de Brasilia — *Brasilia, DF, Brazil*

Bachelor of Science in Chemistry — **2007-2011**
Universidade de Brasilia — *Brasilia, DF, Brazil*

RESEARCH EXPERIENCE

Graduate Student Researcher — **2013–2018**
Mobley Lab, University of California, Irvine — *Irvine, California*

M.Sc. Student CNPq Fellow — **2011–2013**
Laboratorio de Modelagem de Sistemas Complexos, UnB — *Brasilia, DF, Brazil*

TEACHING EXPERIENCE

Chem 1LD General Chemistry Lab — **Fall 2018**
University of California, Irvine — *Irvine, CA*

Chem 138 Computational Organic Chemistry — **Spring 2016**
University of California, Irvine — *Irvine, CA*

PharmSci 170B Molecular Pharmacology II — **Spring 2015**
University of California, Irvine — *Irvine, CA*

Chem H2LB Honors & Majors General Chemistry Lab — **Winter 2014**
University of California, Irvine — *Irvine, CA*

Chem H2LA Honors & Majors General Chemistry Lab — **Fall 2014**
University of California, Irvine — *Irvine, CA*

IQD 117269 Elementos de Espectroscopia Molecular — **2012**
Universidade de Brasilia — *Brasilia, DF, Brazil*

PUBLICATIONS

1. **Approaches for Calculating Solvaton 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 *Journal of Chemical & Engineering Data* **2017**, *62*, 1559-1569

2. **Challenges in the Use of Atomistic Simulations to Predict Solubilities of Drug-Like Molecules [version 1; referees: 2 approved]**, Guilherme Duarte Ramos Matos, David L. Mobley, *F1000Research* **2018**, *7*, doi:10.12688/f1000research.14960.1

3. **Reproducibility of Free Energy Calculations Across Different Molecular Simulation Software**, Hannes H. Loeffler, Stefano Bosisio Guilherme Duarte Ramos Matos, Dongyukh Suh, Benoit Roux, David L. Mobley, and Julien Michel *Journal of Chemical Theory and Computation* **2018**, **article ASAP**, doi: 10.1021/acs.jctc.8b00544

4. **Infinite Dilution Activity Coefficients as Constraints for Force Field Parameterization and Method Development** Guilherme Duarte Ramos Matos, Gaetano Calabro, David L. Mobley, **submitted**

PRESENTATIONS

Poster Presentations:

1. Solvation free energies via alchemical free energy calculations: applications and challenges. San Francisco, CA / 253^{rd} ACS Meeting, 2017.

ABSTRACT OF THE DISSERTATION

Free Energy Calculations in Action: Theory, Applications and Challenges of Solvation Free Energies

By

Guilherme Duarte Ramos Matos

Doctor of Philosophy in Chemistry

University of California, Irvine, 2018

Professor David L. Mobley, Chair

Free energy calculations play an essential role in the study of physical transformations and chemical reactions. Within the universe of free energy calculation applications, solvation free energies (ΔG^{solv}) have a leading role: they are easy to calculate and have a broad scope of applications, from the study distribution coefficients between two phases to solubility prediction. This dissertation discusses some of the uses of ΔG^{solv} as a tool for method development, as a way to calculate infinite dilution activity coefficients, and as a part of an effort to predict solubilities of molecular solids.

In chapters 2 and 3, I discuss the general features of solvation free energy calculations and their applications. We also introduced an update of FreeSolv, a hydration free energy database, and the use of infinite activity coefficient calculations (IDACs) as assisting tools (or potentially substitutes) of hydration free energies in force field parameterization and method development. Chapter 4 discusses the reproducibility of relative alchemical free energies (RAFE) across different software (AMBER, GROMACS, SOMD, CHARMM). We demonstrated that $\Delta\Delta G_{hyd}$ can be reproduced to within about $+0.2\ kcal \cdot mol^{-1}$ and we highlighted the differences and particularities of each code. I was responsible for running the simulations and analyzing the results of the GROMACS code. Chapter 5 discusses an

application of free energy calculations to the prediction of solubilities. I attempted to predict the solubility of acetylsalicylic acid (ASA) by calculating the absolute chemical potentials of the most stable polymorph of ASA and of solutions of different concentrations. Despite not finding a result that agreed with experimental error, I outlined the strengths and weaknesses of the method used and suggested improvements for future attempts.

Chapter 1

Introduction

1.1 Free energies have a central role in understanding much of Chemistry, Biochemistry and Pharmaceutical Sciences

Various chemical processes of biological and pharmacological importance cannot be predicted without understanding their associated free energy changes. We call free energies the thermodynamic potentials that can be used to generate non-expansive work in a system. The minima in the free energy hyper-surface define the equilibrium states under certain conditions. The Helmholtz free energy (A) is the thermodynamic potential for representing systems with constant number of particles (N), volume (V) and temperature (T). The Gibbs free energy (G) is the corresponding potential for systems with constant number of particles (N), pressure (P) and temperature (T). Statistical mechanics is the field of Physics that studies the connections between macroscopic quantities such as A and G to microscopic features of a system. The Helmholtz and the Gibbs free energies are defined as functions of

the system's partition function (Q):

$$\begin{aligned} A(N,V,T) &= -k_B T \ln Q(N,V,T) \\ G(N,P,T) &= -k_B T \ln Q(N,P,T) \end{aligned} \tag{1.1}$$

where k_B is the Boltzmann constant. $Q(N,V,T)$ and $Q(N,P,T)$ are respectively called the canonical and the isothermal–isobaric partition functions. They represent the sum of all available microstates for given constrained properties of their ensembles.

Obtaining accurate absolute free energies values is not an easy task because it is impossible to sample the full configuration space and accurately determine the partition function. It is often possible, however, to determine free energy differences between two end states – a reference and the system of interest. The free energy difference between two states is related to the ratio of the partition functions (Eq. 1.2) which can be numerically estimated [1, 2]. The free energy difference (ΔG) is defined by:

$$\Delta G = -k_B T \ln \frac{Q_{NPT;final}}{Q_{NPT;initial}} \tag{1.2}$$

where $Q_{NPT;final}$ and $Q_{NPT;initial}$ represent the end states of a transformation. Since free energy methods are good for sampling $Q_{final}/Q_{initial}$, [3] properties such as protein–ligand affinities, partition coefficients, activity coefficients, relative solubilities, and phase diagrams can be predicted or understood given the definition of an adequate reference state.

1.2 The theory behind free energy calculations

The theory behind free energy calculations is known since the first half of the twentieth century. John G. Kirkwood set the basis for free energy calculations in 1935 when he developed an expression for the change in free energy produced by the addition or removal

of a single molecule from a fluid system [4]. The method created by Kirkwood, currently known as Thermodynamic Integration (TI), is based on a fictitious Hamiltonian (Eq. 1.3) in which two end states of a transformation $\mathcal{H}_A$ and $\mathcal{H}_B$ are coupled by a parameter λ:

$$\mathcal{H}(\mathbf{q},\mathbf{p};\lambda) = f(\lambda)\mathcal{H}_A(\mathbf{q},\mathbf{p}) + g(\lambda)\mathcal{H}_B(\mathbf{q},\mathbf{p}) \tag{1.3}$$

where $\mathbf{q}$ and $\mathbf{p}$ represent all the position and momentum coordinates that define the system. The coupling functions $f(\lambda)$ and $g(\lambda)$ are usually set in such a way that $\mathcal{H} = \mathcal{H}_A$ at $\lambda = 0$ and $\mathcal{H} = \mathcal{H}_B$ at $\lambda = 1$. Eq. 1.4 defines the free energy difference between states A and B in TI. In practice, Molecular Dynamics or Monte Carlo simulations generate distributions of $\langle \partial\mathcal{H}/\partial\lambda \rangle_\lambda$ for discrete values of λ and the free energy difference between both states.

$$\Delta G = \int_0^1 \left\langle \frac{\partial \mathcal{H}(\mathbf{q},\mathbf{p};\lambda)}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda \tag{1.4}$$

TI performs well if the integrand is smooth [5, 6, 7], but it breaks down if this condition is not met. Numerical integration error in TI is hard to estimate because it depends on the smoothness of the integrand, which is not something usually known and complicates free energy uncertainty estimation.

A second approach to calculate the free energy difference between two states was derived by Robert W. Zwanzig in 1954 through a perturbative approach [8]:

$$\Delta G = -k_B T \ln \left\langle \mathrm{e}^{-\frac{1}{k_B T}(\mathcal{H}_B(\mathbf{q},\mathbf{p}) - \mathcal{H}_A(\mathbf{q},\mathbf{p}))} \right\rangle_A \tag{1.5}$$

where the ensemble average is calculated over the configurations of state A. Even though Eq. 1.5, the theoretical basis of Exponential averaging (EXP), is exact in the limit of large samples, it is inefficient and sensitive to the tails of the distributions sampled in the simulation. One way of improving EXP free energy estimates is through multistage sampling [9], in which a chain of bridging intermediate states with overlapping configuration distributions.

Hence:

$$\Delta G = \sum_i \Delta G_i \tag{1.6}$$

where ΔG_i the free energy difference between two states in the path.

A common problem in multistage approaches is the possible presence of nonintegrable singularities in $\langle \partial \mathcal{H}/\partial \lambda \rangle_\lambda$ at terminal λ values. These singularities can happen depending on how the stages are defined; repulsive potentials are often prone to dramatic changes when atoms are too close to each other. This issue usually results in numerical instabilities and large errors in calculated free energies [10, 11, 12, 13]. Numerical instabilities can be avoided by using soft-core potentials [10], which will be described further in Chapters 2 and 4.

Configuration sampling has always been a challenge for free energy calculations [14]. An approach for removing the systematic bias in EXP calculations combines forward and reverse transformations between the end states of a transformation [15, 16] using the Bennett Acceptance Ratio (BAR) [1]. BAR, developed by Charles Bennett in 1976, estimates the ratio between the partition function of two states [1] and yields good estimates for the free energy even if the overlap between the states is poor. The acceptance ratio estimator (Eq. 1.7) corresponds to the minimum free energy variance. The free energy difference between two states A and B is given by:

$$\left\langle \frac{1}{1 + \frac{N_A}{N_B} e^{\beta \Delta \mathcal{H}_{BA}(\mathbf{q},\mathbf{p}) - \beta \Delta G}} \right\rangle_A = \left\langle \frac{1}{1 + \frac{N_B}{N_A} e^{\beta \Delta \mathcal{H}_{AB}(\mathbf{q},\mathbf{p}) + \beta \Delta G}} \right\rangle_B \tag{1.7}$$

where N_A and N_B are the number of statistically independent samples gathered from states.

Methods such as umbrella sampling [17] and parallel tempering (replica exchange) [18, 19, 20, 21, 22, 23] were also developed to improve the efficiency of free energy calculations. In umbrella sampling, a biased weighting function forces the exploration of the configuration

space in regions that otherwise would have been insufficiently sampled. Unbiased statistics can be recovered because the bias is known. Parallel tempering/ replica exchange depends on calculating a simultaneous series of molecular dynamics trajectories, or Monte Carlo walks characterized by different values of a parameter (temperature, λ, etc). et Each simulation occasionally attempts to swap configurations between states defined by proximal parameter values. This exchange is accepted or rejected according to the Metropolis criterion [24]. The Multistate Bennett Acceptance Ratio (MBAR) [2] is an extension of BAR that optimally uses samples from multiple states to estimate the free energy difference between two states – not only the adjacent ones as BAR uses. MBAR is the most robust and well-performing free energy estimator [7]. The reliable estimation of free energy changes many times depends on the combination between two or more of the methods addressed above. Comparison between different estimators will be done more thoroughly in Chapter 2.

1.3 Free energy calculations have a wide range of applications

Physical quantities such as energy, pressure and volume can be obtained from a regular molecular dynamics or Monte Carlo simulations. As shown in Section 1.2, however, free energies cannot be found from simple ensemble averages. Because of this issue, pioneering applications of free energy calculations were restricted to simple systems. Initial successes in the study of Lennard-Jones fluids at supercritical temperatures [25], liquid argon [26], atomic clusters [27, 28], and the hydrophobic effect [29] attracted due attention and led to the development of the field. Solvation of chemical species was also an important initial target for free energy calculations [30, 31, 32, 33].

As computers became more accessible in the past few decades, free energy calculations were used in more sophisticated problems. In 1984, Tembre and McCammon outlined a perturbative (EXP-like) approach to model ligand–receptor pairs [34]. Jorgensen used a similar reasoning to estimate the relative solvation free energy difference between methanol and ethanol [15], a transformation that became a standard for relative free energy calculations (see Chapter 4). The theoretical basis for deriving protein–ligand association constant was provided by Gilson et al. [35]. Relative free energy calculations became essential tools in computational studies of mutations in proteins [36] and are frequently used in drug design. Free energy calculations also found fertile ground in non-biochemical problems. Partition and distribution coefficients can be predicted from free energy calculations [37, 38, 39, 40, 41] as well as solubilities [42, 43, 44, 45] and phase diagrams [46, 47, 48].

This work does not attempt to cover all the possible applications of solvation free energy calculations. Instead, it is focused on a few applications involving solvation and hydration free energies. In chapter 2, we briefly describe the theory of solvation free energy calculations and outline a robust protocol to calculate hydration free energies (ΔG^{hyd}). Chapter 2 also discusses FreeSolv, a hydration free energy database containing experimental and calculated free energies, hydration enthalpies and entropies. Free energies and enthalpies deserve special attention because they are routinely used to test force field parameters.

Infinite dilution activity coeffients (IDACs) are the subject of chapter 3. IDACs are promising additions to ΔG^{hyd} as force field parameterization and method development tools due to the abundance of experimental techniques to measure activity coefficients and the use of IDACs as parameters in phase equilibria studies [49, 50, 51]. We calculated 237 IDAC values and compared them to their experimental values obtained from the NIST database, observing an excellent agreement with experiment.

Chapter 4 outlines a collaborative project to determine how reproducible relative alchemical free energy calculations (RAFEs) are across different software, namely GROMACS, AMBER,

SOMD and CHARMM. We described the particularities of each code and demonstrated that relative hydration free energy calculations are reproducible to a limit of about 0.2 $kcal \cdot mol^{-1}$.

Chapter 5 discusses the problem of predicting the solubility of molecular solids with the knowledge of its crystal structure. Solvation free energies are an important component of the calculation of the chemical potential of the solution at different concentrations, but the focus of the chapter is the calculation of the absolute free energy of the solid. Using an acetylsalicylic acid crystal as a testing probe, we tested the Einstein Molecule Method and proposed modifications to improve the quality of the resulting chemical potential of the solid.

Chapter 2

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

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 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.

2.1 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 are differences in thermodynamic potentials which describe the relative populations of a chemical species in solution and gas phase at equilibrium [52, 53]. 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 [54], surfaces also have asymmetric effects on ion pairing which depend on the curvature of the surface [55], and molecular geometry and chemical environment affects hy-

drophobic solvation [55]. Although they can be difficult to measure experimentally, ΔG^{solv} and ΔG^{hyd} can be calculated to a precision better than 0.4 kJ·mol^{-1}, even with a relatively modest investment of simulation time, for relatively diverse small neutral molecules [56] such as those seen in the FreeSolv database of hydration free energies [57] 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 [56, 52, 58, 59, 60, 14, 61, 57, 62, 39, 63, 64, 65, 66, 67, 40], 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 [68, 69, 54]. 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[1]. For instance, SAMPL has frequently (in SAMPL1 through SAMPL4) included blind predictions of hydration free energies in particular [52, 58, 59, 60, 14, 61, 57, 62]. 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 [70], distribution coefficients [41], and solvation free energies in organic solvents [71] – 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 [72, 73, 74, 71] and in the testing of free energy methods and force fields [52, 58, 59, 60, 14, 61, 57, 62, 75, 76, 77, 78, 79, 80, 81, 82]. Furthermore, computed free en-

[1]But see the Supporting Information for how protonation state/tautomer challenges may apply here, as in protein-ligand binding.

ergies are in some cases found to be accurate enough to highlight problems with experiments and assist in curation of experimental data [57, 83].

Solvation free energies are often calculated by alchemical free energy methods [84], which simulate a series of non-physical intermediates to compute the free energy of transferring a solute from solution to gas phase (as here) or vise 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 [84, 85]. 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 [86], alchemical free energy calculations using explicit solvent have become a mainstream approach [87, 12], in part because of their formal rigor. Alternative approaches include implicit solvent models [79, 80, 81, 82, 88], which yield ΔG^{hyd} but do not take into consideration solvent configuration around the solute, and Monte Carlo based approaches using the Gibbs ensemble[89, 37, 43, 90, 91, 92, 93] and expanded ensemble [38], though these are most commonly used for molecules that are particularly small and/or rigid.

2.2 Hydration and solvation free energies have a range of applications

Solvation free energies are used to estimate infinite dilution activity coefficients (γ_i^∞) in many solvents by using a single molecule of solute i[94, 95, 96, 97, 98, 99, 100]. Experimental results obtained from gas chromatography [101, 102] 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 solubility free energy by computing both the sublimation free energy (from solid to gas) and hydration free energy (from gas to water). [103]. 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 solutions [42, 44, 104, 105] 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 [106, 107, 108, 109]. Relative solubilities of a given chemical species between different solvents can also be assessed with these calculations [110, 70]. Henry's law solubility constants [111, 112] and solubilities in supercritical fluids [113] 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 [77, 78, 114, 115]. 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 [116]. 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

neutral form of the chemical species in two different solvents [40], as shown in equation 2.1:

$$\log_{10} P_{A\to B} = \frac{\Delta G^{solv,A} - \Delta G^{solv,B}}{RT \ln(10)} \tag{2.1}$$

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 [39, 63, 64, 65, 66, 67, 40]. Hydration free energies themselves are valuable quantities in drug design [12, 117] and can be used to understand the impact of ligand desolvation on the binding process [118, 119] or can be utilized as QSAR descriptors [120].

2.3 Theory and practical aspects of alchemical calculations

Solvation free energies can be calculated in various ways. In this paper we focus on alchemical free energy calculations, which have been one of the most consistently reliable methods in recent applications such as the SAMPL series of challenges [52, 58, 59, 60, 14, 61, 57, 62, 41]. Consider a pair of end states A and B, and their respective Hamiltonians $\mathcal{H}_A(\mathbf{q}, \mathbf{p}; \lambda)$ and $\mathcal{H}_B(\mathbf{q}, \mathbf{p}; \lambda)$.

$$\mathcal{H}(\mathbf{q}, \mathbf{p}; \lambda) = f(\lambda)\mathcal{H}_A(\mathbf{q}, \mathbf{p}; \lambda) + g(\lambda)\mathcal{H}_B(\mathbf{q}, \mathbf{p}; \lambda) \tag{2.2}$$

where $f(\lambda)$ and $g(\lambda)$ are functions of λ used to mix the Hamiltonians, typically set such that $\mathcal{H} = \mathcal{H}_A$ at $\lambda = 0$ and $\mathcal{H} = \mathcal{H}_B$ at $\lambda = 1$; $\mathbf{q}$ and $\mathbf{p}$ represent all the positions and momenta

of the system. With $\mathcal{H}(\mathbf{q}, \mathbf{p}; \lambda)$ we can calculate the free energy difference between A and B:

$$\Delta G = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial \mathcal{H}}{\partial \lambda} \right\rangle_{\lambda} \mathrm{d}\lambda \tag{2.3}$$

This method, called thermodynamic integration (TI) [4], is implemented in practice via a numerical quadrature approach after simulations are done at a discrete set of λ values. It performs similarly to more efficient methods when the integrand is smooth [5, 6, 7]. 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 [8]. In this method, the free energy difference between two states A and B is given by:

$$\Delta G = -\frac{1}{\beta} \ln \left\langle \mathrm{e}^{-\beta[\mathcal{H}_B(\mathbf{q},\mathbf{p};\lambda) - \mathcal{H}_A(\mathbf{q},\mathbf{p};\lambda)]} \right\rangle_A \tag{2.4}$$

where $\beta = (k_B T)^{-1}$. Although equation 2.4 is exact in the limit of large numbers of samples, EXP is inefficient and particularly sensitive to the tails of the relevant distributions, leading to unstable free energy estimates 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. [84, 6, 121]. 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 [122]. Another issue is an asymmetric bias depending on which direction the free energy difference extrapolation is performed [123, 124], so other analysis methods are now preferred[84]. 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, the Bennett's acceptance ratio (BAR), uses the information from both directions to derive the following relationship (which can and has been written in numerous ways):

$$\left\langle \frac{1}{1+\frac{N_A}{N_B}e^{\beta\Delta\mathcal{H}_{BA}(\mathbf{q},\mathbf{p})-\beta\Delta G}} \right\rangle_A = \left\langle \frac{1}{1+\frac{N_B}{N_A}e^{\beta\Delta\mathcal{H}_{AB}(\mathbf{q},\mathbf{p})+\beta\Delta G}} \right\rangle_B \tag{2.5}$$

where N_A and N_B are the number of statistically independent samples gathered from states A and B, and $\Delta\mathcal{H}_{BA}(\mathbf{q},\mathbf{p}) = \mathcal{H}_B(\mathbf{q},\mathbf{p}) - \mathcal{H}_A(\mathbf{q},\mathbf{p}) = -\Delta\mathcal{H}_{AB}(\mathbf{q},\mathbf{p})$ are the Hamiltonian differences between the states at a given point in phase space.

This expression minimizes the free energy variance [1] and makes BAR much more efficient than EXP [123, 124]. 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 [2]. BAR and MBAR perform similarly when the spacing between intermediate states is moderate [7] and therefore only neighboring states have phase-space overlap. Weighted histogram analysis method (WHAM) [125, 126] 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, [7] thus we recommend it as the analysis method of choice whenever possible. TI usually is more sensitive to the choice and number of intermediate states than BAR [127], but it can perform as well as BAR and MBAR if the integrand is smooth [5, 6]. EXP should generally be avoided due to its asymmetric bias and sensitivity to the tails of the distribution [123].

2.3.1 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 [34, 12, 128]. 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 (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 thermodynamic 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 repulsion) 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 [11, 129, 130, 131].

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

$$\mathcal{H}(\mathbf{q}, \mathbf{p}; \lambda) = (1 - \lambda)\mathcal{H}_A(\mathbf{q}, \mathbf{p}) + \lambda\mathcal{H}_B(\mathbf{q}, \mathbf{p}) \tag{2.6}$$

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 a not a problem which is specific

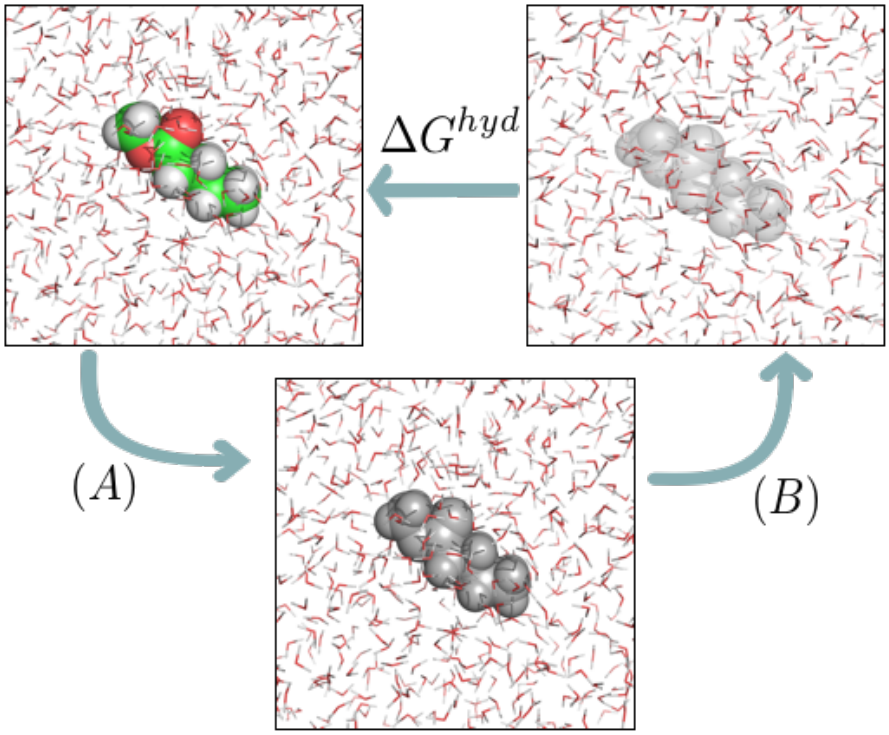

Figure 1: Thermodynamic cycle used to calculate hydration free energies (or, more generally, solvation free energies). In (A), we have states in which charge-charge interactions between the solute and its environment are progressively turned off. In (B) dispersion interactions between solute and water 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).

to TI; rather, this issue can still result in numerical instabilities or large errors in calculated free energies even with other analysis approaches [12, 10, 13]. Thus, more complicated λ pathways are required, such as soft-core potentials, which should in general be used to avoid such numerical problems [10, 11, 13]. A common soft-core form for Lennard-Jones potential between two particles i and j is:

$$U_{ij}^{LJ}(r_{ij}, \lambda) = \lambda^n 4\epsilon_{ij}\left(\frac{1}{[\alpha(1-\lambda)^m + (r_{ij}/\sigma_{ij})^6]^2} - \frac{1}{\alpha(1-\lambda)^m + (r_{ij}/\sigma_{ij})^6}\right) \tag{2.7}$$

where ϵ_{ij} and σ_{ij} are the Lennard-Jones well-depth and lengthscale parameters, respectively, and α is a positive constant which should typically be set to 0.5 [13, 132]. The exponents m and n are most efficient at $n = 1$ and $m = 1$, but other values have been used

too [6, 130, 13, 132]. 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 [133], and alternate potentials that construct near optimal paths for alchemical simulations [134]. Linear basis functions can be used as an alternative to soft-core potentials that approaches the minimum variance possible over all pair potentials [135, 131]; 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 [11, 10, 130, 132, 134], 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.

2.3.2 Considerations for successful alchemical calculations

The accuracy of these calculations is affected by at least three factors [136, 137]: Is our sampling representative and adequate? Is the free energy estimator good enough? Is the force field adequate for the system? Are there critical chemical effects omitted from the simulation, such as protonation state or tautomer effects? 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 λ 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 [7]. Sampling remains a critical issue [14, 136], 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 [40], or in mixed solvents [41].

2.4 We updated FreeSolv, the free community solvation free energy database

2.4.1 About FreeSolv

FreeSolv [57] is a hydration free energy database for neutral[2] 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. The molecular weights for compounds in FreeSolv range from 16.06 Daltons (methane) to 498.88 Daltons (decachlorobiphenyl). The number of rotatable bonds runs as high as 12, but most compounds are largely rigid. 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 [72, 73, 71, 74], and can assist the testing of new solvation free energy methods [138, 139]. One example of the use of hydration free energies as target physical properties to fit in force field development is the Automated Force Field Topology Builder and Repository (ATB) [140, 141, 142]. ATB is an online platform based in large part on FreeSolv and provides similar information, though with force field parameters of the GROMOS family. However, the database is not available in an easily downloadable public format and is only accessible via web queries. ATB partly relies on data taken from previous works from Mobley and collaborators [143], which are included in FreeSolv [57].

[2]For additional discussion of why we focus on neutral compounds, see the Supporting Information Section I.

While calculated hydration free energies for all compounds have been available in FreeSolv since the database was constructed [57], previous values had been calculated with somewhat heterogeneous protocols in a variety of different studies spread over roughly 10 years [69, 56, 143, 54, 87, 53, 14, 57]. In this work, we have updated FreeSolv by repeating all of the calculations using a single protocol, now also computing enthalpies and entropies of hydration.

2.4.2 Method details

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.[131]. Box size does not affect the result of solvation free energy calculations as long as good practices, which recommend that box edges be at least twice the Lennard-Jones cutoff distance, are followed [144]. We ran 5 nanoseconds of Langevin dynamics per state with 2 femtosecond time steps in GROMACS 4.6.7 [145, 146, 147, 148, 149, 150] at 298.15K. Van der Waals interactions were neglected beyond a smoothly switched cutoff of 1.2 nm. Different cut-off radii are commonly used, but one should be aware the choice of cut-off can affect calculated solvation free energies. However, long range dispersion corrections can be employed (as here) to remove the cutoff-dependence of calculated free energies [151]. (However, it is worth noting that in the case of heterogeneous systems, such as for binding free energy calculations, it may be necessary to use reweighting techniques instead). [151] Our choice of soft-core is the so called 1-1-6 (m and n equal to 1 in equation 2.7) which leads to statistical uncertainties approximately of the same size as uncertainties from simulations using optimized

path soft-cores [132]. Pressure was maintained at 1 atm by the Parrinello-Rahman barostat [152]. Enthalpy and entropy decomposition required 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 kJ $\cdot$ mol^{-1}, roughly the level of typical thermal energy ($1\,k_BT$) as done in, for example, host-guest binding calculations [153]. We used the default Langevin dynamics' friction coefficients implemented in GROMACS (see SI). The size of friction coefficient only affects equilibration and correlation times, but should not affect the calculated hydration free energies and enthalpies. 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. 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 toolkits [154, 155, 156] to generate molecular structures and assign AM1-BCC partial charges [157, 158], then charged mol2 files were written out. The AMBER Antechamber package (as distributed with AMBER14) was then used to to assign parameters from the GAFF[67] small molecule force field (version 1.7), and these were then converted to GROMACS format and solvated with the TIP3P water model [159]. The script which performs the setup and re-generates 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` (`github.com/mobleylab/alchemical-analysis`)[128]. Here we also introduce FreeSolv v0.51, which extends v0.5 by making the small molecule parameter sets available additionally in CHARMM, DESMOND, and LAMMPS formats. Additional details can be found in the supporting material and in the FreeSolv README files.

2.4.3 FreeSolv hydration free energy results

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

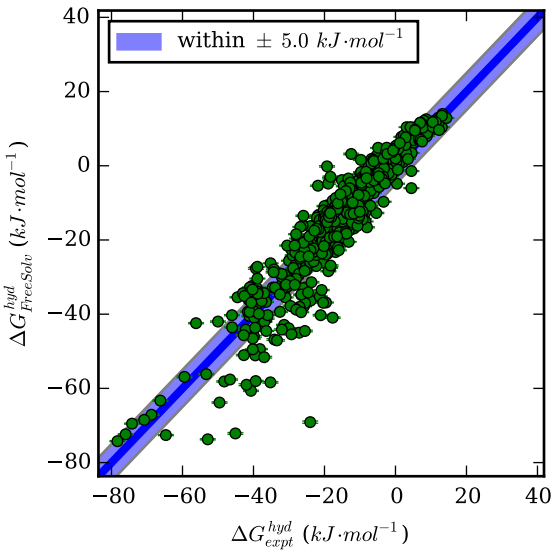

Figure 2: Calculated versus experimental hydration free energies for the compounds in FreeSolv version 0.5. Calculated values are on the vertical axis and experimental on the horizontal.

In the calculations described in this study, we found an average error of 1.3 ± 0.3 kJ·mol^{-1}, RMS error of 6.4 ± 0.3 kJ·mol^{-1}, average absolute error of 4.7 ± 0.2 kJ·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 [57], 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 [12]. We have previously used this data to address force field issues on hydroxyl groups [73], and also to highlight functional groups in the set which pose particular challenges [87]. Full details about which compounds have systematic errors, along with the functional groups represented in each compound, are present in the FreeSolv database itself.

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/mobleylab/FreeSolv`). These quantities have been used for various purposes, including to assist in the study, development, and testing of implicit solvent models [160, 161]. However, it is important to remember that these components come from our particular decomposition of the free energy [162, 163, 164, 165], 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.

2.4.4 Hydration enthalpy calculations

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 [166, 153, 167], 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 (Figure S2 and Table S1). We find that compounds which have accurate hydration free energies do not necessarily have accurate hydration enthalpies and vise versa; 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.

To compute hydration enthalpies, 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:

$$\Delta H^{hyd} = \langle U_{solution} \rangle - (\langle U_{water} \rangle + \langle U_{vacuo} \rangle) \tag{2.8}$$

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 ≈ 2.9 kJ $\cdot$ mol^{-1} [166]; 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 [168]. We note that other groups have also omitted these contributions, which still await a thorough assessment of relative magnitude for small molecule hydration enthalpies [166].

Hydration entropies are calculated via the equation:

$$\Delta S = \frac{\Delta H - \Delta G}{T} \tag{2.9}$$

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 [153].

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 [169], 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 studies [166]) that enthalpies provide additional information on the thermodynamics and constraints on the force field (though as noted above, additional enthalpy corrections may be needed [168]). While ΔH^{hyd} and ΔS^{hyd} can act as additional constraints for force field parameters, one of them can always be calculated from the other and the corresponding ΔG^{hyd}, meaning that it is not worthwhile to use all three values as constraints simultaneously. That is, ΔH^{hyd} and ΔS^{hyd} are always highly anti-correlated because of how they are calculated. More details can be found in the supporting information.

2.4.5 Components of hydration enthalpies

We also partitioned the hydration enthalpy, ΔH, into two components: a solvent interaction term and a conformational change term, ΔH^{hyd}_{int} and ΔH^{hyd}_{conf}, 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:

$$\Delta H_{int}^{hyd} = \langle U_{solv} \rangle_s - \langle U_{vac} \rangle_s - \langle U_{water} \rangle_w \tag{2.10}$$

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 box of pure water containing the same number of water molecules under the same conditions. 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:

$$\Delta H_{conf}^{hyd} = \langle U_{vac} \rangle_s - \langle U_{vac} \rangle_v \tag{2.11}$$

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 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:

$$\Delta H^{hyd} = \Delta H^{hyd}_{solv} + \Delta H^{hyd}_{conf} \tag{2.12}$$

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.

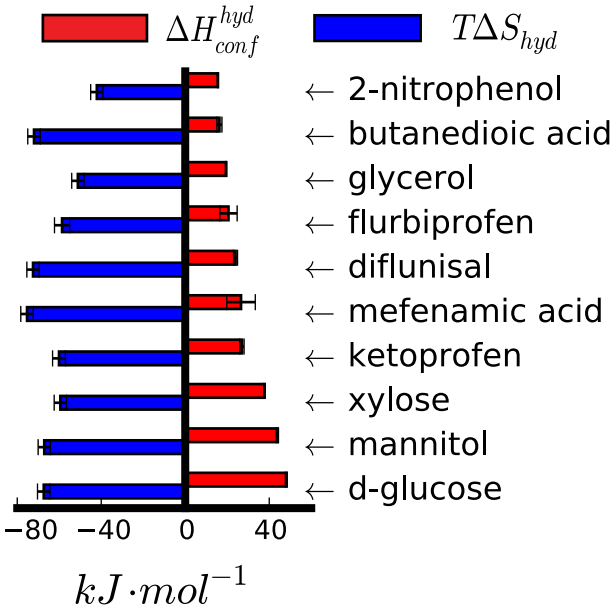

Figure 3: Conformational enthalpies and associated entropies of compounds with highest ΔH^{hyd}_{conf}. Error bars represent the standard error.

2.5 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. In our experience, BAR and MBAR require less tuning to work well, while TI requires special care to get the gradients right in rapidly varying regions and introduces unknown integration error, thus we recommend MBAR as our preferred general-purpose method, even though TI can in principle also work well. EXP should be avoided, in general, in partly because of the large bias introduced.

We also introduced an update to FreeSolv [57] (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 kJ$\cdot$ mol^{-1} and an average error close to 1 kJ$\cdot$ mol^{-1}. With the aid of ORCHYD [169], 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.

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 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/mobleylab/FreeSolv` and contributions are welcomed there.

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.

2.6 Supporting Information

2.6.1 FreeSolv has hydration free energies for neutral compounds

FreeSolv focuses on hydration free energies of neutral compounds. While many studies have computed hydration free energies for charged species, measuring hydration free energies for charged species in isolation is impossible, so extracting these can require extrathermodynamic assumptions or introduce other complexities. Thus, we agree with previous work suggesting that the main focus should be on hydration free energies of neutral compounds [54] (see particularly footnote 61), as also discussed elsewhere [57].

It is worth noting, however, that the database does contain a variety of carboxylic acids. In solution, these are typically charged at neutral pH. However, hydration free energies are typically reported for the neutral form of the molecule [57] so those are the values used here.

2.6.2 Additional practical considerations for calculation of solvation free energies

One of the appeals of hydration free energies is that they could be relatively free of the protonation state and tautomer issues which can challenge predictions of protein-ligand binding; however, this seems unlikely to be true in general (though it may be true for many of the relatively small, fragment-like compounds in FreeSolv). Particularly, small molecules can certainly have multiple relevant tautomers in solution, tautomers which change on transfer between environments (such as gas to water transfer), or tautomers which are uncertain yet important for solvation and transfer properties. While these issues may not play a major role in solvation of the present compounds, they certainly can become a factor elsewhere, as was amply illustrated in the recent SAMPL5 challenge, which focused on calculation of cyclohexane-water distribution coefficients. Many participants estimated these from solvation free energies in both solutes, and protonation and tautomer issues played an important role [41].

It is also worth noting one important issue that can affect interpretation of literature solvation free energies – these can use different standard states. Values reported in FreeSolv are for transfer free energies from gas (at a 1 M standard state) to solution (at a 1 M standard state). It is also possible to report and/or calculate values for transfer from an alternate 1 atm standard state in gas to a 1M standard state in solution [54], resulting in values which differ by an additive constant relating to the difference in gas phase standard state. Thus,

care must be taken when pulling values from the literature in order to ensure a consistent standard state is used.

2.6.3 Rebuilding the FreeSolv database

All input files deposited in FreeSolv were re-generated using the `rebuild_freesolv.py` script deposited on our GitHub repository at `github.com/mobleylab/FreeSolv`. To re-build the input files, one can simply run this script, which requires the Chodera lab's 'open-moltools' package and the Mobley Lab's 'SolvationToolkit', both of which are conda installable from the omnia channel, and are also available on GitHub at `github.com/choderalab/openmoltools` and `github.com/mobleylab/solvationtoolkit` respectively. In this particular iteration of rebuilding FreeSolv and re-running the calculations, we used openmoltools version 0.6.7.

2.6.4 Additional plots and statistics

Figure **4**(a) statistics:

- Kendall $\tau = 0.76 \pm 0.02$
- Pearson $R = 0.943 \pm 0.005$

Figure **4**(b) statistics:

- Kendall $\tau = 0.40 \pm 0.02$
- Pearson $R = 0.60 \pm 0.03$

Figure **5**(a) statistics:

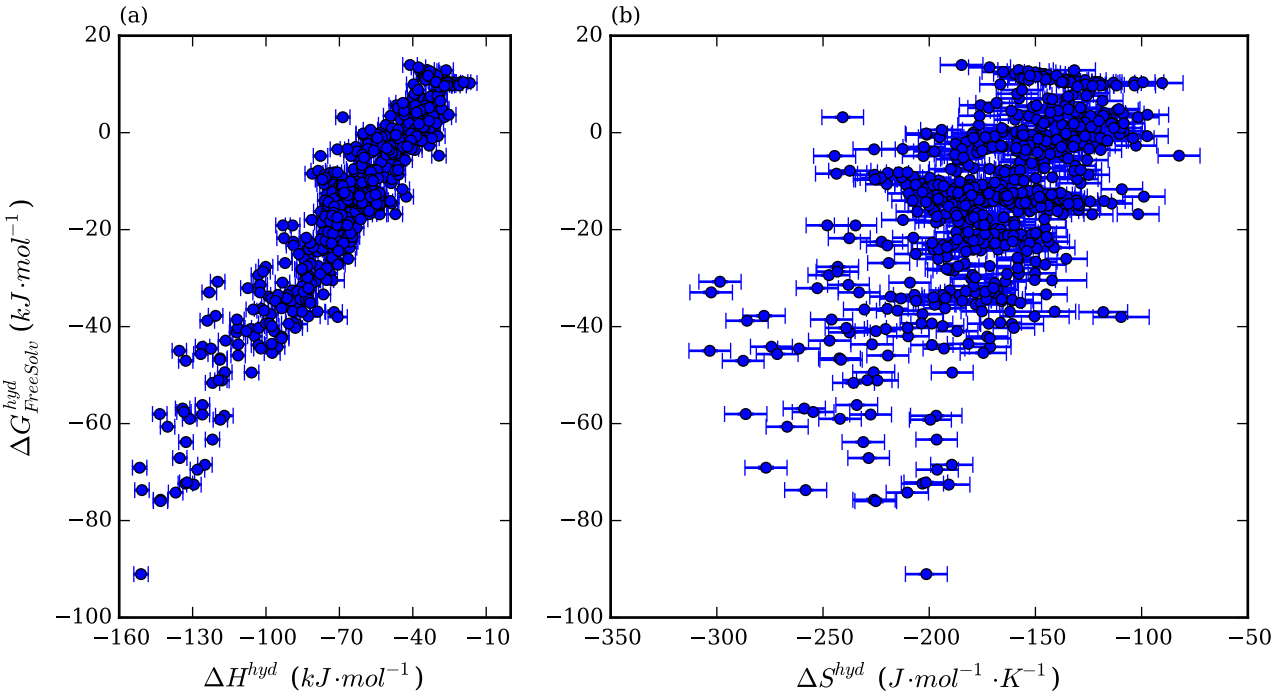

Figure 4: Correlation plots between (a) calculated enthalpies and hydration free energies in FreeSolv, and (b) calculated entropies and hydration free energies in FreeSolv. Error bars are given as standard errors in the mean.

- Average error $= -5 \pm 2\, kJ \cdot mol^{-1}$
- RMS $= 9 \pm 2\, kJ \cdot mol^{-1}$
- Average unsigned error $= 7 \pm 2\, kJ \cdot mol^{-1} \cdot K^{-1}$
- Kendall τ $= 0.7 \pm 0.2$
- Pearson R $= 0.88 \pm 0.08$

Figure **5**(b) statistics:

- Average error $= 2.3 \pm 0.8\, kJ \cdot mol^{-1}$
- RMS $= 3.2 \pm 0.9\, kJ \cdot mol^{-1}$

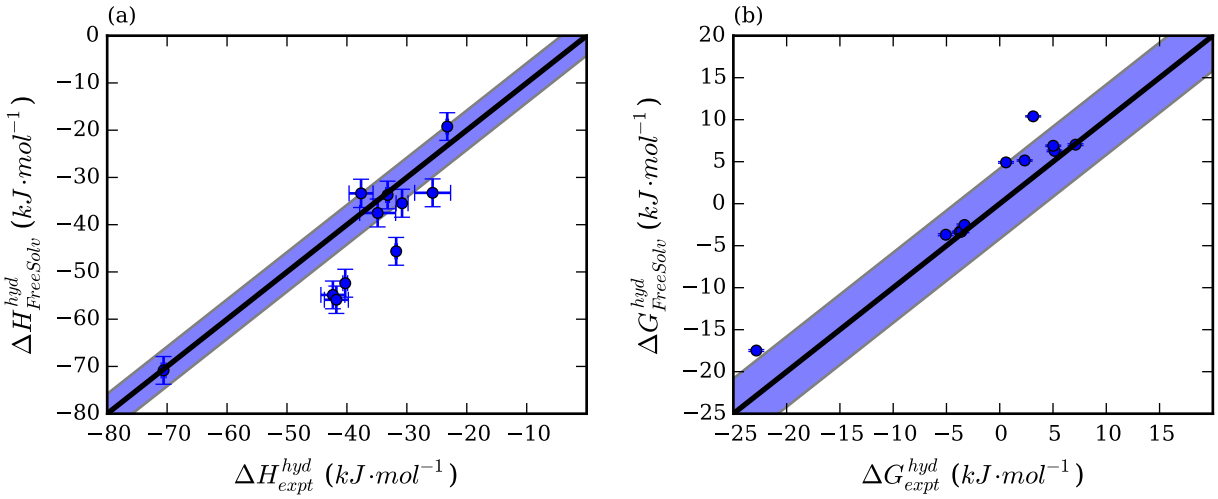

Figure 5: Correlation plots between (a) the 11 calculated enthalpies in FreeSolv, and their corresponding experimental values from ORCHYD, and (b) calculated hydration free energies for these same 11 compounds, and their corresponding experimental values. The shaded area indicates values within 4 kJ/mol of the $x = y$ line.

- Average unsigned error $= 2.3 \pm 0.8\, kJ \cdot mol^{-1} \cdot K^{-1}$
- Kendall $\tau\ = 0.85 \pm 0.2$
- Pearson $R\ = 0.96 \pm 0.05$

2.6.5 Simulation details

The following are GROMACS 4.6.7 simulation input parameters, as are the MDP files with full details which are deposited in the Supporting Information and on GitHub.

General information

- Friction coefficient $=$ mass$_{particle}/\tau_t$, $\tau_t = 2.0\, ps$.
- Parrinello-Rahman barostat: $\tau_p = 10\, ps$ and compressibility $= 4.5 \cdot 10^{-5}\, bar^{-1}$.

Table 1: The 11 FreeSolv compounds with known experimental enthalpies from ORCHYD. All values are in kJ·mol^{-1}. Sources of experimental data are given in the FreeSolv database itself.

FreeSolv key	**CID**	**SMILES**	$\Delta G^{hyd}_{FreeSolv}$	ΔG^{hyd}_{expt}	$\Delta H^{hyd}_{FreeSolv}$	ΔH^{hyd}_{expt}
mobley_2689721	8078	C1CCCCC1	6.3 ± 0.1	5 ± 3	-34 ± 3	-33.2 ± 0.3
mobley_2784376	6351	C1CC1	10.40 ± 0.07	3 ± 3	-19 ± 3	-23.3 ± 0.2
mobley_3053621	241	c1ccccc1	-3.4 ± 0.1	-3.8 ± 0.8	-46 ± 3	-31.8 ± 0.2
mobley_3183805	7247	Cc1ccc(c(c1)C)C	-3.3 ± 0.1	-4 ± 3	-55 ± 3	-42 ± 2
mobley_3211679	8079	C1CCC=CC1	4.9 ± 0.1	0.6 ± 0.4	-38 ± 3	-35 ± 3
mobley_3452749	10686	Cc1cccc(c1C)C	-3.7± 0.1	-5 ± 3	-56 ± 3	-42 ± 2
mobley_7010316	7966	C1CCC(CC1)O	-17.5 ± 0.1	-23 ± 3	-71 ± 3	-70.6 ± 0.4
mobley_8006582	9253	C1CCCC1	6.90 ± 0.09	5 ± 3	-35 ± 3	-31 ± 1
mobley_8127829	7500	CCc1ccccc1	-2.5 ± 0.1	-3 ± 3	-52 ± 3	-40.3 ± 0.4
mobley_8885088	8882	C1CC=CC1	5.15 ± 0.08	2 ± 3	-33 ± 3	-26 ± 3
mobley_9100956	7962	CC1CCCCC1	7.0 ± 0.1	7 ± 3	-33 ± 3	-37 ± 2

Electrostatics

- PME cut-off: 1.2 nm.
- PME order: 6
- Fourier spacing = 0.10 nm
- additional details can be found in the MDP files deposited with this paper and on GitHub at `github.com/mobleylab/freesolv`.

vdW interactions

- Cut-off: 1.0 nm
- Switch at 0.9 nm
- DispCorr = AllEnerPres

Free Energy calculation control parameters

- vdW lambda schedule: 0.0, 0.0, 0.0, 0.0, 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
- FEP lambda schedule (all non-specified lambdas use this schedule): 0.0, 0.25, 0.5, 0.75, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0
- soft-core $\alpha = 0.5$
- soft-core power (m in Equation 8) $m = 1$
- additional details can be found in the MDP files deposited with this paper and on GitHub at `github.com/mobleylab/freesolv`.

All input files were generated (as noted above) via the `rebuild_freesolv.py` script deposited in the FreeSolv GitHub repository. This relies on `openmoltools`; we used version 0.6.7. As noted in the main body of the text, AM1-BCC charges were assigned with OpenEye's quacpac python module; we used `openmoltools` to drive this process. Specific source code used for charging is available at `https://github.com/choderalab/openmoltools/blob/v0.6.7/openmoltools/openeye.py#L13` and generates molecular conformations prior to charging, as was recommended at `http://docs.eyesopen.com/toolkits/cookbook/python/modeling/am1-bcc.html`. We have found this procedure considerably more robust than the Antechamer AM1-BCC procedure used in earlier versions of the database, in part because it removes the dependence of charges on the input conformation.

For solvated systems, all solutes were placed in cubic boxes with at least 1.5 nm from the solute to the nearest box edge, and then solvated with TIP3P water using the gromacs tool `genbox`, so the number of water molecules used varied depending on the solute (but can be obtained from the topology and coordinate files deposited in the database).

2.6.6 Absolute differences between old and new FreeSolv ΔG^{hyd} values

Table **2** shows the largest differences between the calculated values previously deposited in FreeSolv and those shown here. For *most* compounds in the set, differences are relatively modest, but for this particular group some of the changes are quite significant. Some of these compounds are carboxylic acids in their neutral form, which can suffer from slow sampling of the orientation of the hydroxyl proton [53] so that may be one possible explanation for some of the discrepancies.

Table 2: Fifteen biggest differences between old and new ΔG^{hyd} values, in kJ·mol^{-1}.

FreeSolv key	**name**	**old** ΔG^{hyd}	**new** ΔG^{hyd}	$\Delta\Delta G^{hyd}$
mobley_2099370	ketoprofen	-49.82	-72.19	22.37
mobley_1527293	flurbiprofen	-36.43	-58.42	21.99
mobley_820789	butyric acid	-22.86	-39.50	16.64
mobley_2078467	ibuprofen	-28.93	-45.46	16.53
mobley_2850833	2-hydroxybenzaldehyde	-20.47	-36.88	16.41
mobley_4792268	pentanoic acid	-22.48	-37.90	15.42
mobley_2929847	3-methylbutanoic acid	-23.07	-37.03	13.96
mobley_1735893	hexanoic acid	-21.27	-32.98	11.71
mobley_7758918	propionic acid	-26.84	-38.05	11.21
mobley_8207196	simazine	-36.13	-45.69	9.56
mobley_2913224	acetylsalicylic acid	-47.10	-39.35	7.75
mobley_8916409	malathion	-54.39	-46.88	7.52
mobley_1821184	3-methyl-1H-indole	-27.42	-34.17	6.74
mobley_7690440	methyldisulfanylmethane	6.20	-0.39	6.59
mobley_1792062	1,2-dibromoethane	0.80	-5.34	6.13

It is possible that other discrepancies could result from parameter differences, though we have not been able to identify any clear origins of differences. Lennard-Jones parameters seem to be identical between the (potentially different) GAFF versions used in these setups, though potentially there could be differences in bonded parameters (because of differences in how input files were generated between when the database was originally constructed and now, GROMACS topologies use different function types for these parameters so equivalent

parameters will not appear identical). However, a more likely origin of discrepancies is partial charges, as charging procedures for the studies originally used in constructing FreeSolv in some cases used Antechamber's AM1-BCC charging procedure on a database conformation of the molecule, rather than our current, more modern charging procedure which uses reasonable molecular conformations before assigning AM1-BCC charges with the OpenEye toolkits. However, we have not yet verified whether these issues can definitively be linked to the charging procedure.

Another possibility is simply protocol differences and differences in software versions. For example, some of our early work used constant volume simulations for our free energy calculations (after equilibration at constant pressure) which we later found could, in some cases, introduce additional noise to calculated hydration free energies due to artifactual densities at some λ values [87].

Chapter 3

Infinite Dilution Activity Coefficients as Constraints for Force Field Parameterization and Method Development

Abstract

Molecular simulations see widespread use in calculating various physical properties of interest, with a key goal being predictive molecular design. These simulations, including molecular dynamics (MD) simulations, begin with a underlying energy model or force field and then, based on this model, use simulations to compute properties of interest. However, one of the most significant challenges in molecular dynamics and modeling studies is ensuring that the force field is a good enough approximation of the underlying physics that computed quantities can be used to reproduce experimental properties with the desired level of accuracy. Parameterization of force fields depend on various experimental properties including as much of the chemistry of interest as

possible. Physicochemical properties measurable in a relatively straightforward manner are particularly interesting for developers. Such properties can be measured for a relatively diverse chemical set and used to expand the parameterization dataset as needed. Here, we examine infinite dilution activity coefficients (IDACs) which are experimental quantities that can play this role. We retrieved 237 empirical IDACs from NIST's ThermoML, a database of measured thermodynamic properties, and we estimated the corresponding values using solvation free energy calculations. We found that calculated IDAC values correlate strongly with experiment. Specifically, the natural logarithm of calculated and experimental IDAC values shows a Pearson correlation coefficient of 0.92 ± 0.01. The calculated IDAC values allow us to identify strengths and potential weaknesses of force field parameters for specific functional groups in solutes and solvents, suggesting these may be a valuable source of data for force field parameterization, capturing some of the same type of information as hydration and solvation free energies and thus potentially providing a useful new source of experimental data.

3.1 Introduction

Infinite dilution activity coefficients (IDACs) tell us how far an infinitely dilute mixture is from ideal solution conditions [170, 49, 50, 51, 171] and they are of considerable experimental and theoretical interest [172, 51, 95]. Deviations from ideality indicate whether a solute is particularly good or particularly poor for a given solvent. This means that activity coefficients have a variety of downstream applications, such as for input to chemical engineering models studying liquid-vapor coexistence [50, 171].

Ideal solutions are mixtures where the interactions between two solvent molecules are equal to the interactions between two solute molecules and to the interactions between a solvent and a solute molecule [170]. Real solutions, however, do not satisfy this condition; in most cases, solute and solvent molecules have self-interactions which are not identical to their

mutual interactions in solution. In this sense, activity coefficients can be interpreted as a measure of the propensity of a solute molecule to interact with the solvent.

Activity coefficients are also important because they help us determine the effective concentration of a particular component, or its propensity to react. Specifically, the activity of a component in a system, or its effective concentration, is the product between the component's activity coefficient and its real concentration and is related to the chemical potential of a component in a mixture (μ_i) (Eq. 3.1).

$$\mu_i = \mu_i^0 + RT \ln \gamma \cdot \frac{[i]}{[i_0]}, \tag{3.1}$$

where γ is the activity coefficient, μ_i^0 is the standard chemical potential of the component, $[i]$ and $[i_0]$ are its concentration and standard concentration, R is the ideal gas constant, and T is the absolute temperature. If $\gamma = 1$, the mixture is ideal; if $\gamma > 1$ or $0 < \gamma < 1$, the mixture behaves non-ideally.

There are different ways to express activity coefficients, including in terms of concentration, mole fraction, or partial pressure. Here, we focus on activity coefficients expressed in terms of mole fraction (χ) henceforth, but it is worth remembering the connections to concentration and other forms. The reference ideal state also plays an important part in the definition of activity coefficients. Here, we define activity coefficients with reference to an ideal solution in the sense of Raoult's law, where, for each component in a mixture, $\gamma_i \to 1$ as $\chi_i \to 1$ [173]. In this definition, the pure liquid is considered an ideal solution because all the interactions between its components have the same magnitude. With this reference state (called the Lewis-Randall standard state), the activity coefficient is 1 for the pure solution. Other reference states are also commonly employed. For example, a common textbook definition uses an ideal dilute solution as a reference state, where the activity coefficient is 1 for a

solute at infinite dilution. Here, however, we use the Lewis-Randall reference state as it is the state employed by the database of experimental values we will compare to.

An infinite dilution activity coefficient (IDAC) is the activity coefficient of a component when its concentration is infinitely small in a mixture. It is related to the slope of isothermal pressure-composition phase diagrams when the mole fraction tends to zero, and is proportional to the Henry's Law constant [174]. The measurement of IDACs depends on factors such as the volatility of solvent and solute [171]. Techniques such as gas-liquid chromatography[171, 175], high-performance liquid chromatography[175] and differential ebulliometry [171, 175, 176] are traditionally used to measure activities in extremely dilute systems at varying concentrations, leading to the infinite dilution activity coefficient by extrapolation [175]. There has been considerable interest in predicting these coefficients [172, 177, 95, 178, 179, 97, 180, 181, 182, 183, 184, 185, 176, 175, 171] due to their use in phase equilibria studies in chemical engineering applications [49, 50, 51, 179].

IDACs (also represented by γ^{∞}) are related to solvation free energies by the following equation [172, 95, 178]:

$$\gamma_i^{\infty} = \exp\left(\frac{\Delta G_i^{\text{solv}} - \Delta G_i^{\text{self solv}}}{k_B T}\right) \cdot \frac{\rho_{\text{molar}}^{\text{solvent}}}{\rho_{\text{molar}}^{\text{pure solute}}} \tag{3.2}$$

where ΔG_i^{solv} is the solvation free energy of a solute i, $\Delta G_i^{\text{self solv}}$ is the solvation free energy of a solute i in its bulk phase, k_B is the Boltzmann constant, T is the absolute temperature, and ρ_{molar} is the molar density of component of the mixture.

The solvation free energy is the free energy change of transferring a molecule from an ideal gas state to a solvent [33, 88, 53, 186, 187]. If the solvent is water, we call the solvation free energy a hydration free energy (ΔG^{hyd}). Solvation free energies tell us which phase a

given molecule prefers in a multiphasic system and also provide information on how a solute behaves in different environments [52, 53].

In this work, our interest is in physical models which can be used to calculate activity coefficients and related properties from molecular simulations. Particularly, molecular simulations begin with a description of the energy and forces in a physical system as a function of the coordinates — what is known as a "force field" — and allow calculation of numerous physical properties from simulations of such systems [188] . In addition to potentially providing predictions of various quantities like host-guest and protein-ligand binding affinities [189, 137, 190, 191, 192], distribution and partition coefficients [66, 40], solvation free energies [193, 194, 195] or other physical properties for design applications, comparison of such results to experiment provides a quantitative test of the underlying physical model or force field.

Hydration and solvation free energies have proved particularly valuable in quantitatively testing all-atom molecular simulations and force fields and in highlighting systematic errors, in part because ΔG^{solv} and ΔG^{hyd} for small molecules can be calculated to a precision better than 0.1 kcal/mol [68, 143]. The SAMPL series of blind challenges for computation have helped illustrate this, with several challenges focusing on solvation and hydration free energy prediction [52, 58, 59, 83, 196]. SAMPL1 through SAMPL4 featured blind prediction of hydration free energies [52, 58, 59, 83]. SAMPL5 included the prediction of partition coefficients, which also required the calculation of solvation free energies in water, cyclohexane, and octanol [197, 196]. Likewise, the FreeSolv database of calculated and experimental hydration free energies has been broadly useful for similar reasons [57, 187]. Work on FreeSolv has helped highlight and resolve various force field problems, such as with hydroxyl group parameters [73].

Even though there are experimental and calculated hydration free energy databases such as FreeSolv[57, 187] and ATB [142, 140], ΔG^{solv} and ΔG^{hyd} measurements are difficult

and require considerable expertise and specialized equipment, so few to no experimental measurements are presently made [52]. Infinite dilution activity coefficients, however, are a critical property of consideration when studying binary mixtures, and even more so as the number of components and/or the system size grows for industrial applications [50], meaning that they are subject to considerable experimental attention. Additionally, they are easier to measure than ΔG^{solv} and ΔG^{hyd} [50, 51, 101, 102], and can be calculated with similar precision as the aforementioned free energies. This means they are an ideal candidate as an alternative to hydration free energies for benchmarking computational chemistry methods and force fields.

3.2 Computational Methods

We obtained experimental activity coefficients at infinite dilution from ThermoML [198, 199, 200], an XML-based system for storage and exchange of thermochemical data. ThermoML was accessed on July 27, 2017 using `thermopyl` [201], a Python tool that allows interaction with the database and provides access via a Pandas `Dataframe`. We made a search for IDACs of organic compounds containing less than 40 heavy atoms at temperatures between 250 K and 400 K. All the activity coefficients were obtained approximately at 101 kPa. The search was restricted to molecules containing no elements other than C, O, N, F, P, S, and Cl. The heaviest solute molecule of the set was hexadecane (226.44 Da) and the lightest was methanol (32.04 Da). The heaviest solvent molecule was tetradecanoic acid (228.37 Da) and the lightest was water (18.02 Da). Most molecules were fairly rigid with less than three rotatable bonds, but a few, such as hexadecane and undecane, had up to 13 rotatable bonds.

We found 263 coefficients but limited our study to 237 coefficients. The reduced set size resulted from problems building the simulation boxes for some systems with solutes or solvents with long chains, as well as parameterization issues for some tertiary amines. The final set

contains a variety of combinations of a moderate number of different solvents and solutes. This allows us to look for trends in accuracy both as a function of solute and as function of solvent.

All solvation free energy calculations were performed using now relatively standard alchemical free energy calculations described further below, but automated via the OpenEye Orion cloud computing platform. The calculations could have been done on local computing resources using an identical protocol, but Orion allowed for higher throughput.

Setup of calculations began with processing the solute and solvent names from the data obtained from ThermoML. From names, SMILES strings were generated using OpenEye's `OEChem` toolkits, and stored as `OEMol` objects [202], with one `OEMol` for each solvation free energy calculation to be done (i.e. one for calculation of the solvation free energy of each solute in pure solute, and one for calculation of solvation free energy of the solute in pure solvent). In each case we attached the SMILES string of the solvent (generated with `OEChem`) to the `OEMol` for the solvent, along with the target temperature and pressure for each simulation (as these were required by the Orion workflow we constructed) and then output the resulting set of molecules to an OpenEye binary file (`.oeb`) for use on Orion.

The Orion workflow then conducted solvation free energy calculations from these input files in a straightforward manner, ultimately using Yank [203] to run free energy calculations as further detailed below. Before input into Yank, however, simulation boxes were built and parameters were assigned. Specifically, starting geometries for simulation boxes were built (in PDB format) from the solute and the specified solvent (as indicated by SMILES strings attached to the input molecule) using the OpenEye toolkits to generate molecular structures and conformers, and `PACKMOL` (version 17.221) to build boxes consisting of the solvated systems[204]. Force field parameters were then assigned via Antechamber and Ambertools (version 16.16.0), using the GAFF 1.8 small molecule force field [67] and AM1-BCC charges [157, 158] (the latter as assigned by the OpenEye toolkits, version 2018.2.1) to describe

solvents and solutes, with the exception of water, which was modeled by using TIP3P [159]. Once parameterized, the resulting systems were stored as `ParmEd` [205] (version 2.7.3) objects and attached to the OpenEye data record to progress through the workflow.

Following parameterization, equilibration stages were run using OpenMM [206, 207, 208] (version 7.1.1), followed by production free energy calculations done with Yank (version 0.20.1), using protocols that are now relatively standard (e.g. as in [187]). Nonbonded interactions were calculated for all inter-atomic distances under a cutoff of 9 . Electrostatic interactions were computed using particle mesh Ewald (PME) [209, 210]. Each Hamiltonian replica exchange simulation run using Yank had 1000 iterations of 500 MD steps of 2 femtoseconds each at each λ value, totaling one nanosecond per replica. Bonds involving hydrogen were constrained. All simulations were conducted at the target temperature and pressure associated with the ThermoML data for the experiment, as provided by our input files. Solvation free energies were estimated with the Multistate Bennett Acceptance Ratio [2] (MBAR), an extension of the Bennett Acceptance Ratio [1] that considers the overlap between a given state and all the others in the path between the end states, as provided by Yank. MBAR is the most consistently well-performing free energy estimator [7] and is the default free energy estimator in Yank. 5ns Langevin dynamics simulations were run using OpenMM [207, 208] to obtain densities of pure solvents and pure solutes to calculate the ratio in Eq. 3.2.

3.3 Results

IDACs (γ^{∞}), as defined in the Introduction, tell us how far from ideality a mixture is when the concentration of the solute is infinitely small. They are widely used as input for engineering models, such as for prediction of liquid-vapor equilibria, and they can be calculated from solvation free energies (Eq. 3.2) [172, 95, 178]. The natural logarithm of γ^{∞} is proportional

to the difference between the free energy of solvation of a solute i in a given solvent (ΔG_i^{solv}) and the free energy of solvation of the solute molecule in its pure bulk phase ($\Delta G_i^{\text{self solv}}$, free energy of "self-solvation") plus the ratio between the densities of the solvent and the pure solute:

$$k_B T \ln \gamma_i^{\infty} = \Delta G_i^{\text{solv}} - \Delta G_i^{\text{self solv}} + k_B T \ln \left(\frac{\rho_{\text{molar}}^{\text{solvent}}}{\rho_{\text{molar}}^{\text{pure solute}}} \right) \tag{3.3}$$

where k_B is the Boltzmann constant and T is the absolute temperature.

Here, we calculated the solvation free energies using MBAR [2] and compared to experimental values as shown in Figure **6**.

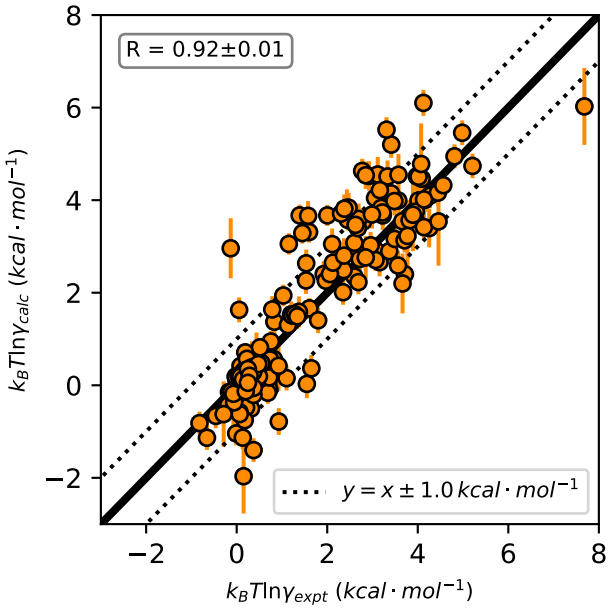

Figure 6: Calculated versus experimental $k_B T \ln \gamma^{\infty}$ for 237 solute – solvent pairs taken from ThermoML. Calculated values are on the vertical axis and experimental on the horizontal.

We found an average error (in free energy units, as in Equation 3.3) of 0.06 ± 0.05 kcal·mol^{-1}, a root-mean-square (RMS) error of 0.73 ± 0.05 kcal · mol^{-1}, an average unsigned error of

0.48 ± 0.04 kcal $\cdot$ mol^{-1}, a Kendall τ value of 0.67 ± 0.03, and a Pearson R value of 0.92 ± 0.01.

Given that each IDAC tells us how well a solute molecule interacts with the solvent with respect to how well it interacts with itself, Fig. **6** can potentially also give us some idea whether a given force field underestimates or overestimates the intermolecular forces between solvent and solute. The diagonal line in Fig. **6** corresponds to the cases where the simulation agreed with the experiment. If a point is located below the diagonal line, the force field potentially underestimates solute – solvent interactions relative to solute – solute interactions. On the other hand, if a point is located above the diagonal line, the force field potentially overestimates solute – solvent interactions relative to solute – solute interactions.

Having an extensive set of IDAC values allows us to look for systematic errors in the force field and how it describes particular functional groups and solvents, as has been done previously in studies with hydration free energies [143, 211]. Here, in order to detect possible issues with force field parameters, we partitioned our dataset by functional groups and by solvents. The absolute value of the average error of the free energy differences for functional groups with more than five occurrences in the set can be seen in Figure **7**.

Here (Fig. **7**), analysis of errors by functional group is slightly complicated by the fact that errors could depend on the identity of the solvent or the identity of the solute. In the limit of very large datasets this should be easily surmountable, because a large number of samples would ensure that analysis by solute would involve averaging over a large number of solutes, and analysis by solvent would involve averaging over a large number of solutes. Here, however, our set is relatively small, so it is important to not place too much confidence in any analysis of systematic errors. Still, such analysis can suggest likely targets needed for follow up studies to confirm potential problems, and some trends seem clear.

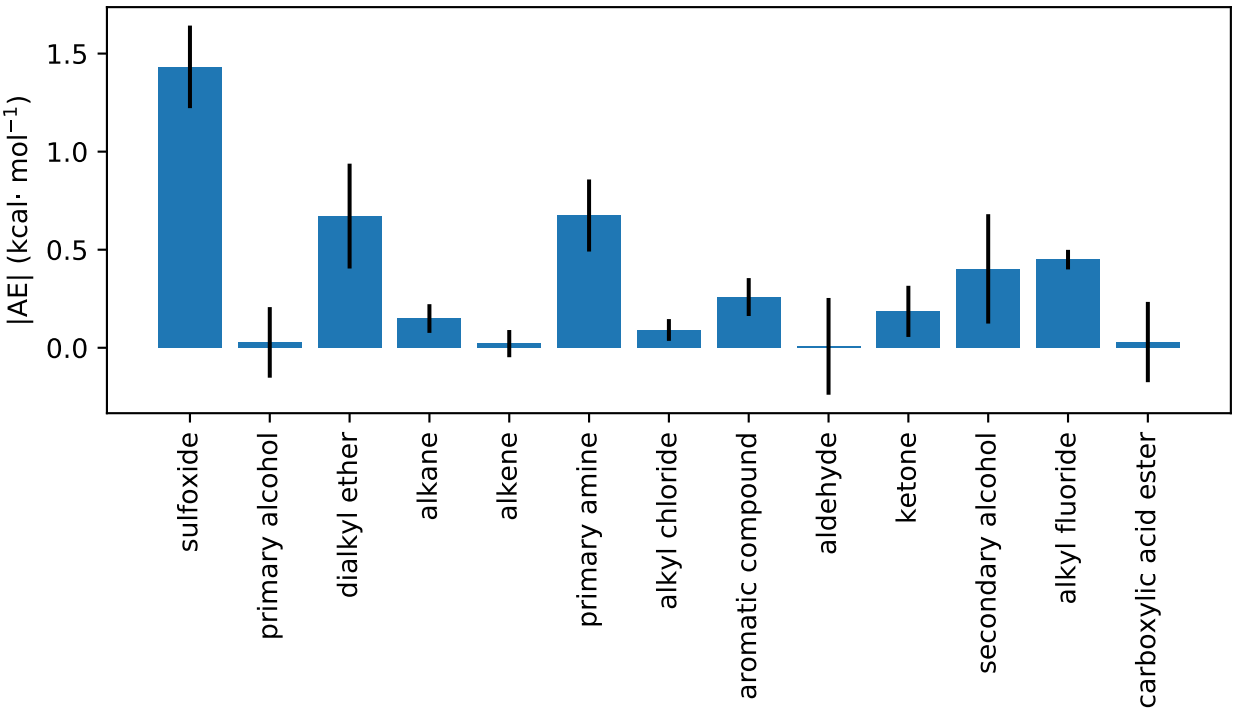

Figure 7: Absolute values of the average errors (AE) for functional groups with more than five occurrences in the set. Error bars denote the standard error in the mean of the quantity on the vertical axis.

Here, based on our analysis of errors in activity coefficients broken down by the functional group observed in the solute, we found that sulfoxide had the largest absolute value of the average error of the set (Figure **7**). In fact, the only sulfoxide present in the set was dimethylsulfoxide (DMSO), and all IDACs involving DMSO were for this molecule in different solvents (DMSO was never present as a solvent in the simulations). The average error of $+1.4 \pm 0.2$ kcal/mol suggests a systematic error in the GAFF description of DMSO (Fig. **8**).

We did a similar analysis of IDAC values broken down by functional groups appearing in the solvent. The absolute average error of the free energy differences by solvents with more than five occurrences can be seen in Figure **9**.

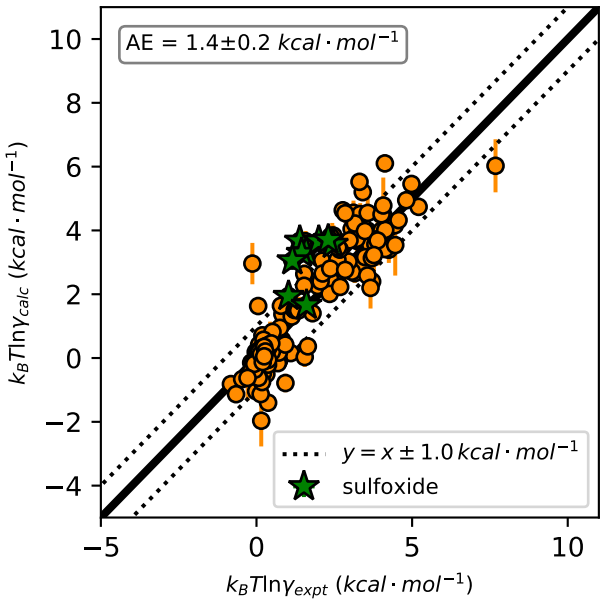

Figure 8: As a solute, dimethylsulfoxide (DMSO) shows a positive shift (average error of 1.7 ± 0.2 kcal·mol^{-1} for DMSO) with respect to the $y = x$ line, suggesting a potential systematic error in the force field. The set contained no measurements where DMSO was a solvent.

Given that solvents tended to occur many times in IDAC measurements, our analysis by solvent provided more data concerning potential systematic errors than did our analysis by solute. Methanol, formamide, and ethylene glycol were the solvents whose IDACs showed the largest average absolute errors of the set (Figure **10**).

Figure **10** singles out solvents containing four different functional groups for particular analysis, and highlights several potentially important trends. For instance, when examining water as a solvent, IDAC values are nearly evenly spread around the $x = y$ line for IDACs in water (Fig. **10**(d)), which suggests that the differences with respect to the experimental IDACs are random in nature or are caused by solute parameters. This is perhaps expected, given that water models are typically given special attention and parameterized quite carefully. In contrast, Fig. **10** (a), (b), and (c) show systematic shifts away from the diagonal line, suggesting potential systematic errors for these solvents. Average errors were 0.9 ± 0.1 kcal·mol^{-1}, -0.7 ± 0.2 kcal·mol^{-1} and -0.6 ± 0.3 kcal·mol^{-1} for $k_BT \log \gamma$ for solutes in diethylene glycol,

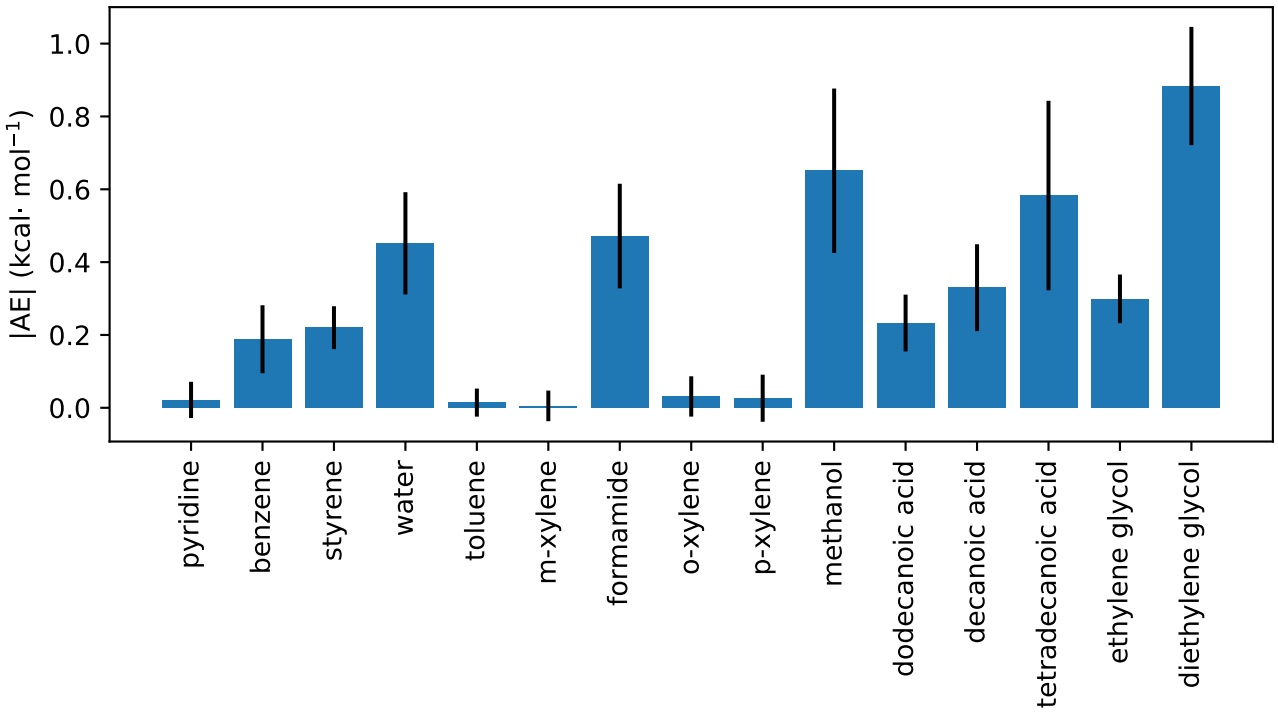

Figure 9: Absolute values of the average errors (AE) for solvents with more than five occurrences in the set. Error bars denote the standard error in the mean of the quantity on the vertical axis.

methanol and tetradecanoic acid, respectively. The plots in Fig. **10**(a), (b) and (c) suggest the potential presence of systematic errors, but the size of our sample (11, 7 and 11 IDACs per solvent in (a), (b), and (c)) limits our ability to investigate in much detail. We believe, however, that the expansion of the data set can help confirm our analysis.

3.4 Discussion

Here, we calculated a large number of infinite dilution activity coefficients and compared with experimental values extracted automatically from NIST's ThermoML database. We used relatively standard (if computationally demanding), easily-automated approaches for calculation of solvation free energies, and performed the calculations in a high-throughput

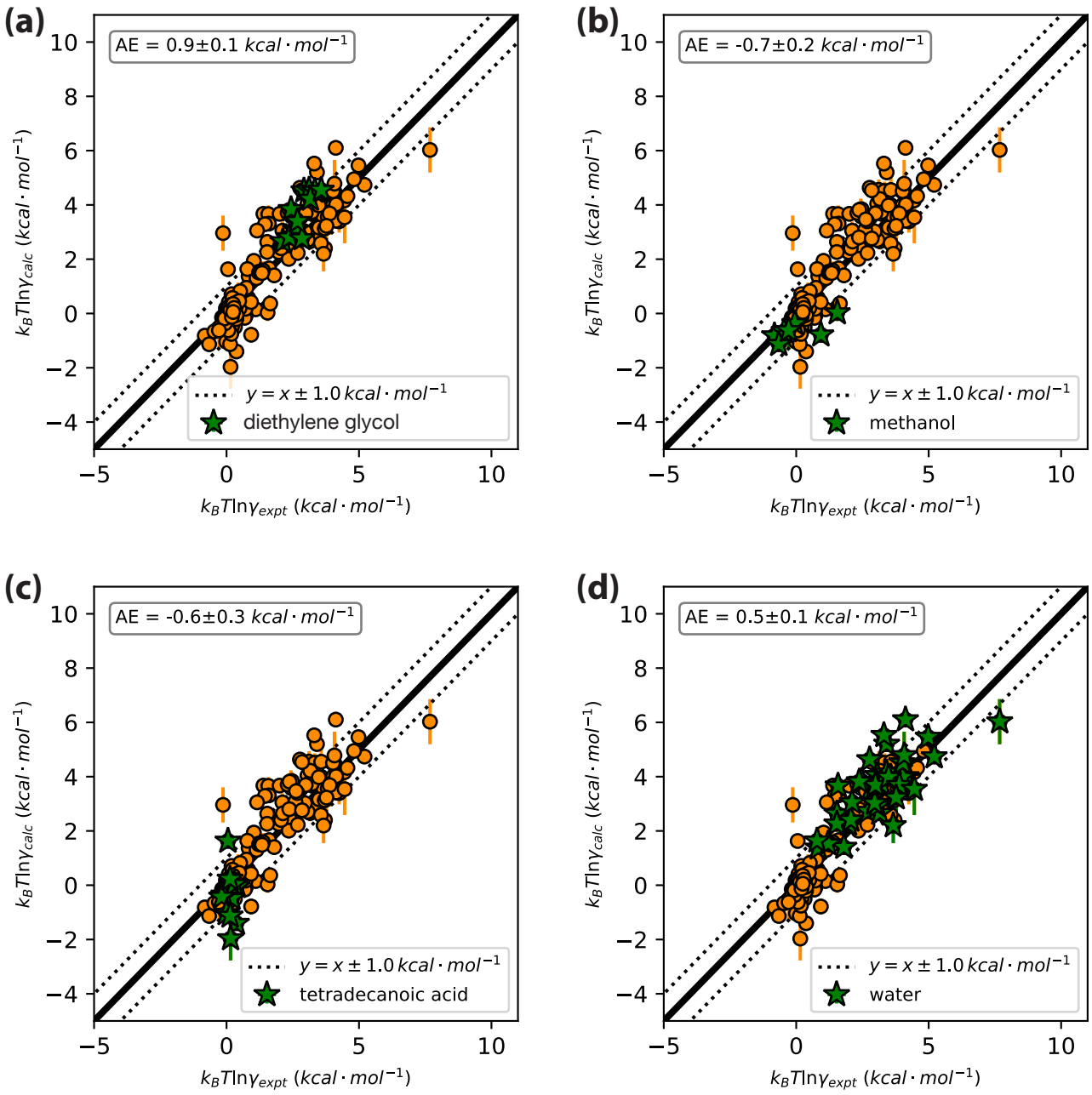

Figure 10: Plots highlighting the IDACs for solutes in diethylene glycol (a), methanol (b), tetradecanoic acid (c), and water (d) simulations. While points in (d) are consistently spread around the $y = x$ line, the remaining plots suggest systematic errors in the description of diethylene glycol, methanol and tetradecanoic acid. Additional data can be found in the Supporting Information. AE stands for the average error of the green star-shaped points.

manner on OpenEye's Orion cloud computing platform. This allowed the calculations to run in a highly parallel and relatively efficient manner and complete on an overnight timeframe. Interestingly, agreement between calculated and experimental values is actually quite good, and these calculations are also able to highlight clear systematic issues for particular functional groups or types of solutes/solvents, suggesting promising areas for investigation of possible force field deficiencies. The fact that infinite dilution activity coefficients can also be measured in a relatively straightforward manner means these will likely serve as a valuable source of data for future tests of computational methods.

Our calculations were kept rather short for computational efficiency, resulting in somewhat high statistical errors. For computational efficiency, we ran only 1 ns per lambda value, allowing each $k_B T \log \gamma^{\infty}$ to be computed quite quickly. With additional sampling at each lambda value we could further reduce the statistical error and better ensure that sampling is adequate, so extending the simulations may be something to explore in future work. Advances in hardware have already provided considerable gains in this area, already making it possible to perform the large number of calculations reported here in a relatively short amount of time, in part due to the availability of GPUs [212].

Each calculated $k_B T \log \gamma^{\infty}$ value requires two solvation free energy calculations, which modestly increases the computational cost in comparison to our traditional approach of using hydration free energies, which requires a single free energy calculation [187] . This small increase is worthwhile given that hydration free energies involve gas-to-water transfers while many events computational chemistry seeks to predict (e.g. binding, solubility, partitioning, permeation, etc.) involve transfer between condensed phases. Thus IDACs may be particularly appealing for force field parameterization since IDAC calculations involve the transfer of molecules between condensed phases, similar to biological and pharmacological events which typically involve transfer from one condensed phase environment to another.

We also believe that there are two other advantages of using IDACs: First, IDACs for a given solute of interest can be obtained in different solvents, allowing the potential to explore how well a force field represents molecules both as solutes and as solvents. Second, molecules become polarized when transferred from gas phase to water and hydration free energies with conventional force fields (excepting polarizable force fields) might not be able to describe this phenomenon well; parameterization to hydration free energies could even build in systematic error resulting from lack of treatment of polarization. This may be particularly important; while an IDAC calculation also involves a transfer between environments, it is a transfer between two condensed phases, which usually is associated to a much smaller change in polarization of the solute molecule in comparison to transfer from the gas phase, as in the case of hydration free energies. IDAC values thus could potentially be an even better way to test how a force field represents a condensed-phase environment than hydration free energies, and potentially a better source of parameterization data.

The abundance of $k_B T \ln \gamma^\infty$ values around zero (Figure **6**) is noteworthy and is potentially an artifact of the type of data which is available in ThermoML. Recall that a value of 0 here corresponds to an activity coefficient of 1 (see Section 4.1). Specifically, a large portion of the available data is for transfer of solute molecules to solvents of similar polarity — for example, transfer of a polar solute to a polar solvent, or (more commonly) transfer of a nonpolar solute to a nonpolar solvent. If the dataset contained more cases of transfer of a nonpolar solute to a polar solvent, or a polar solute to a nonpolar solvent, we would expect to see more values substantially different from 0. Thus, we believe that the IDAC data should be expanded to include more activity coefficients for compounds of very different polarity than the solvent, to capture more features of transfers between nonpolar (or weakly polar) environments to very polar environments.

3.5 Conclusion

Here we reported our results calculating some 237 different infinite dilution activity coefficients (IDACs) for small molecules in various solvents, and comparing to experiment. In general, results were quite promising and showed considerable predictive power over a range of 6 $kcal/mol$ in free energy units.

Our results suggest that IDAC values can potentially play an important role in testing force fields and assisting with force field parameterization. They are frequently measured for applications in engineering while other commonly used quantities, such as hydration free energies (ΔG^{hyd}), are not routinely measured. Furthermore, since IDACs can be calculated in a straightforward manner using solvation free energy calculations, they can be calculated with essentially the same degree of precision as solvation free energies, and with the same procedures. IDACs actually could be even better than solvation free energies, which involve transfer between gas and liquid phases, since they are related to transfer between two condensed phases – pure solute and pure solvent – which makes them an ideal candidate to test how a force field represents condensed-phase environments.

Not only are IDAC calculations appealing in principle, but our results suggest that these calculations can indeed be helpful in identifying force field issues needing attention. Specifically, graphic analysis of experimental and calculated $k_B T \ln \gamma^{\infty}$ values enabled the identification of possible systematic errors in the force field used in this study. We hope that the evidence shown in this work drives future research in expanding the number of experimental activity coefficients at infinite dilution in the literature, and in using γ^{∞} as a new source of constraints for force field parameterization and method development.

Chapter 4

Reproducibility of Free Energy Calculations Across Different Molecular Simulation Software

Abstract

Alchemical free energy calculations are an increasingly important modern simulation technique to calculate free energy changes on binding or solvation. Contemporary molecular simulation software such as AMBER, CHARMM, GROMACS and SOMD include support for the method. Implementation details vary among those codes but users expect reliability and reproducibility, i.e. for a given molecular model and set of forcefield parameters, comparable free energy differences should be obtained within statistical bounds regardless of the code used. *Relative* alchemical free energy (RAFE) simulation is increasingly used to support molecule discovery projects, yet the reproducibility of the methodology has been less well tested than its absolute counterpart. Here we present RAFE calculations of hydration free energies for a set of small organic molecules and demonstrate that free energies can be reproduced to within about

0.2 kcal/mol with aforementioned codes. Absolute alchemical free energy simulations have been carried out as a reference. Achieving this level of reproducibility requires considerable attention to detail and package–specific simulation protocols, and no universally applicable protocol emerges. The benchmarks and protocols reported here should be useful for the community to validate new and future versions of software for free energy calculations.

4.1 Introduction

The free energy is a fundamental function of thermodynamics as it explains how processes in nature evolve. The equilibrium balance of products and reactants in a hypothetical chemical reaction can be immediately determined from the knowledge of the free energy difference of reactants and products and their concentrations. The free energy landscape of a given system, however, can be very complicated and rugged with barriers which impose limits on how fast the process can take place. It is therefore of little surprise that the determination of free energy changes is of utmost importance in the natural sciences, e.g. for binding and molecular association, solvation and solubility, protein folding and stability, partition and transfer, and design and improvement of force fields.

The calculation of free energies via molecular simulations [137, 194, 213, 214, 215] has been particularly attractive as it promises to circumvent certain limitations of experimental approaches. Specifically, processes can be understood at the atomic level and there is the potential that computational techniques can be more cost and time effective, especially if they can predict the properties of new molecules before their synthesis. Thus, a multitude of methods have been devised to make reversible work estimates accessible through computation [137, 194, 213, 214, 215]. However, the reliability of estimates is still very much a matter of concern [194, 216].

Here we are interested in *alchemical* free energy methods because they are firmly rooted in statistical thermodynamics and should give asymptotically correct free energy estimates, i.e. they are correct for a given potential energy function in the limit of sufficient simulation time [217, 218, 219, 137]. The method has been applied in various forms for several decades now since the early days of computer simulation [220, 1, 30, 31, 34, 15]. The method is also increasingly referred as free energy perturbatioBhati:2017:JChemTheoryComputn (FEP) in the literature, even though different techniques may have actually been used to estimate free energy changes. The method has gained renewed attention in recent years — concomitant with improvements in computer hardware design — within the traditional equilibrium framework [35, 221, 118] and also increasingly in combination with non-equilibrium techniques [5, 222, 223]. The name "alchemical" comes from the nonphysical intermediates that often need to be created to obtain reliable estimates of free energy differences between physical end states, and because parts or all of a molecule may effectively appear or disappear in a transformation. In the context of force field methods the transformation takes place in parameter space, i.e. the various force field parameters are varied by scaling. This can be a particularly efficient approach compared to methods involving physical transition pathways or order parameters, as it does not require sampling of diffusive motions, avoids crossing prohibitively large energy barriers if transition pathways are not well chosen, and is easier to automate.

Alchemical free energy simulations rely on the concept of thermodynamic cycles [34]. As the free energy is a state function, the sum of free energy changes computed around any closed cycle must be zero. This also implies that the reversible work can be computed along conveniently chosen legs of the cycle, even if the cycle is artificial. For example, in Fig. **11** the relative free energy of hydration can be computed along the vertical legs, that is, following the physical process of moving a molecule from the gas phase to the liquid phase, or along the horizontal legs in a non-physical but computationally more efficient alchemical calculation.

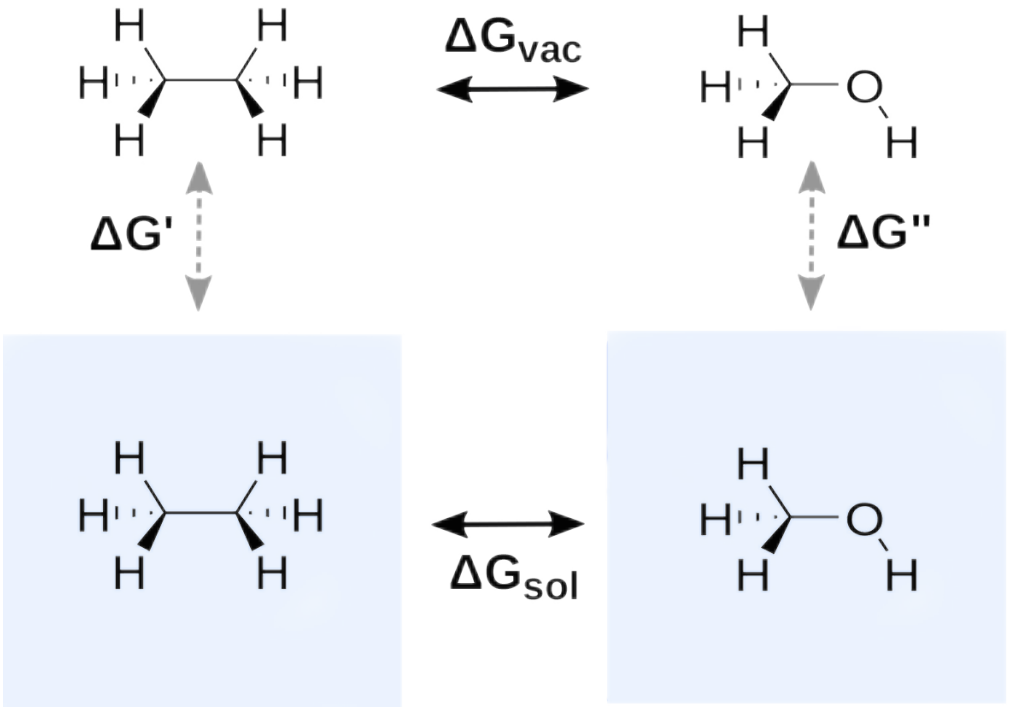

Figure 11: The thermodynamic cycle to compute the relative free energy of hydration $\Delta\Delta G_{\text{hydr}} = \Delta G_{\text{sol}} - \Delta G_{\text{vac}} = \Delta G'' - \Delta G'$. The example is for the ethanol $\leftrightarrow$ methanol transformation. A blue background indicates water and a white background indicates gas phase. Alchemical simulations are performed along the non-physical horizontal legs while vertical legs illustrate the physical process of moving a molecule from the vacuum to the solution. The latter is also accessible through absolute alchemical free energy simulation, see e.g. Ref. 187.

Absolute (standard) alchemical free energy calculation has been of particular interest for many years [35, 221, 118, 5, 223, 224]. *Absolute* here really means that the equilibrium constant of a physical reaction, e.g. binding and dissociation, can be calculated directly by completely decoupling or annihilating a whole molecule from its environment. This term is mostly used to distinguish it from techniques usually referred to as *relative* (see below). It should be emphasized that the "absolute" approach still results in a *relative* free energy between the state where the solute fully interacts with its environment and the state where it does not. The term *decoupling* here is taken as meaning the scaling of the non–bonded *inter*–molecular interactions between the perturbed group (all atoms that differ in at least one force field parameter between the end states) and its environment. We distinguish decoupling from *annihilation*, as the latter also includes a scaling of the *intra*–molecular non-bonded

interactions in addition to the inter–molecular interactions. [85][1] Torsional interactions may also be scaled in an annihilation protocol, but bond and angle terms are usually not scaled as this leads to poorly converging free energy estimates. [225] These schemes may require two simulations along the opposite edges of a quadrilateral thermodynamic cycle but approaches that produce the reversible work directly in one simulation have been proposed as well [226, 227].

Relative alchemical free energy (RAFE) calculations transform or mutate one molecule into another. An appealing aspect of RAFE calculations is the hope that they may be somewhat less demanding computationally or converge better than the more ambitious approaches that require a complete decoupling or annihilation of a ligand from its environment. RAFEs have proven useful for instance to rank sets of related molecules according to their binding affinity for a given receptor. This approach has recently gained increased traction in the context of relative free binding energies between small molecules, e.g. drug or lead like molecules and biomolecules [136, 228, 191, 229].

RAFEs can be calculated by making use of either the so–called single or dual topology method. Dual topology means that groups of atoms of the end states are duplicated and thus both sets are present at all times but do not interact with each other [230, 225]. The atom types are not changed, and, in principle, the groups of both states would need to have the same total charge to avoid partially charged intermediates. In practice this could require, depending on force field, to duplicate all atoms of the end states. Only non–bonded interactions need to be scaled such that the disappearing end state is fully decoupled from its environment [225]. The dual topology method is the most straightforward approach to compute RAFEs when the two molecules are structurally dissimilar. In situations where all atoms in a perturbed molecule are duplicated a dual topology calculation is the technically

[1]It is worth noting that the terms "double decoupling method" and "double annihilation method" also employ the words "decoupling" and "annihilation" but used in an entirely different sense in the context of standard binding free energy calculations.

same as two absolute calculations executed simultaneously in opposite directions. This, however, comes with additional complications as the two independent molecules can drift apart and sample completely different environments (e.g. binding site versus bulk solution). It has been shown though that with the introduction of special restraints or constraints this can be a viable option [231, 232, 233]. Restraints between corresponding atoms can also be used without affecting the free energy [233]. A recent alternative considered molecules with a common core where all atom types are the same. [234] The charges that would be typically different in individual parameterization due to the local chemistry were made equal. This means that the core does not need to be duplicated and thus is not included in the mutation.

Single topology means that the alchemical transformation of one molecule into another molecule is handled via a single set of connected atoms. Atoms of a given type are directly transformed, typically by linearly scaling the force field parameters, into atoms of a different type. The single topology method offers a straightforward route to implement RAFE calculations.[230, 119, 15, 225]

An example of single topology is the simple transformation of methane (CH4) into tetrafluoromethane (CF4), in which real atoms are mapped into real atoms with different force field parameters at the two end states. However, in most typical implementations, a certain number of non-interacting "dummy" atoms must hold the place of disappearing/appearing atoms in order to balance the number of atoms in both end states. For example, one dummy atom is necessary to match a methane molecule (5 atoms) transforming into methanol (6 atoms). Dummy atoms have no non–bonded interactions in the end state but normally retain the bonded terms of the original atom to avoid complications with unbound atoms [225]. Some practitioners stress that a given dummy atom should retain at most only one angle term with respect to "real" non-dummy atoms to yield correct results [225, 235, 236]. For example, if the angle term Atom2–Atom1–Dummy1 is included then additional angle terms such as Atom3–Atom1–Dummy1 should not be included. Likewise, if the dihedral term

Atom3–Atom2–Atom1–Dummy1 is included then additional terms such as Atom4–Atom2–Atom1–Dummy1 should not be included.

The single topology approach seeks to exploit the topological and structural similarity of the two end states. [230] Chemical similarity is also of importance; e.g. chirality and binding modes where the relative three dimensional arrangement of groups in space must be taken into account. These considerations notwithstanding, the single topology approach is broadly applicable to a wide range of transformations. For example, ring breaking is technically challenging, [229] but it has been shown this can be done in certain circumstances [237, 235]. Generally, modern MD software (e.g. AMBER, [238] CHARMM, [239] GROMACS, [150] GROMOS, [240] and SOMD. [241, 208]) support a hybrid approach that combines aspects of single and dual topology [235].

Another algorithmic decision for single topology is whether the implementation scales force field parameters ("parameter scaling") and/or energy components ("energy scaling") [225]. In the former case each parameter is scaled individually, e.g. in the case of a harmonic bond or angle term, the force constant and the equilibrium distance/angle are scaled individually. In the latter case, the total energy is scaled, all at once, or, equivalently for each individual force field contribution. While free energy is a state function that depends only on the end points, the pathways taken by the two methods through state space or alchemical space are different.

As alluded to above, consistency and reliability are the principal matter of concern. In particular, we need to ensure reproducibility of free energy results among computer codes. To the best of our knowledge this has not been systematically tested yet for a set of different MD packages. However, there have been some recent efforts to test *energy* reproducibility across packages [242] — a necessary but not sufficient prerequisite. Another study went further and also compared liquid densities across packages, revealing a variety of issues [243]. For free energies, given a predefined force field and run–time parameters we ought to be able to

obtain comparable free energy results within the limits of statistical convergence. Prior work has successfully compared calculated absolute hydration free energies across GROMACS and DESMOND codes. [53] This comparison has not yet been carried out for relative free energies.

Nevertheless, it is critical that free energy changes computed with different simulation software should be reproducible within statistical error, as this otherwise limits the transferability of potential energy functions, and the relevance of properties computed from a molecular simulation to a given package. This is especially important as the community increasingly combines or swaps different simulation packages within workflows aimed at addressing challenging scientific problems [244, 245, 246, 247, 248].

In this work we compute the relative hydration free energies of a set of small organic molecules using several software and protocols (see Fig. **12**). Solvation free energies have a wide range of uses and various methods exist to compute them [86]. They are also needed for calculations of a variety of important physical properties, and to calculate binding free energies where the solution simulation (see Fig. **11**) is combined with a mutation of the molecule bound to a partner [86]. A large database of hydration free energies computed from alchemical free energy (AFE) simulations, FreeSolv, has been presented recently. [57, 187] Here, we focus on the reproducibility of RAFE with the simulation programs AMBER, CHARMM, GROMACS and SOMD. We will discuss the reversible work results obtained with these packages and make observations regarding simulation protocols, setup procedures and analysis techniques. We will also deliberate on what needs to be done to progress the field, both from a usability perspective as well as from the view point of code development.

4.2 Methods

One practical challenge is that the free energy methodologies used in one MD program are not always available in another package, or the same functionality is provided via different algorithms (e.g. algorithms for pressure and temperature scaling, integrators, cufoffs for Coulomb and vdW interactions, etc). We also note that the implementation of alchemical free energy calculations is very different among the simulation codes (see Section 4.2.1 for details). This implies that using the same free energy parameters across all codes, and especially using the same lambda schedule, will not automatically lead to equivalent free energies. In Fig. **13** we show various plots of the free energy derivative versus lambda to demonstrate this. Hence, the protocol and especially the choice of lambda values was adjusted individually for each code based on previous experience of the researchers involved. In addition there may be differences in the choice of physical constants used for evaluating potential energies. A previous study noted that variations in the hardcoded values of Coulomb's constant lead to detectable differences in single point energies calculated by CHARMM, AMBER or GROMACS [242].

To circumvent some of these practical problems, we will compare relative free energies calculated via three protocols. In the "unified protocol" we calculate relative free energies by scaling together all force field parameters simultaneously along the alchemical path — i.e. partial charges, van der Waals parameters, and bonded parameters. In the "split protocol" we calculate relative free energies by scaling separately the van der Waals parameters and the partial charges parameters. The order in which this has to be done is detailed in section 4.5.3. The scaling of the bonded terms can be combined with either transformation. In the "absolute protocol" we calculate relative hydration free energies as the difference between two calculated absolute hydration free energies.

4.2.1 Alchemical Free Energy Implementations

We begin by examining the differences in the alchemical free energy implementations of the four MD codes we consider — AMBER, CHARMM, GROMACS and SOMD. One key difference is in the softcore functions implemented in each code as summarized in section 4.5.1. [10, 11] Softcore functions are used to avoid the numerical stability problems of the conventional Lennard-Jones (LJ) and Coulombic inverse power law potentials [249, 130], as they display singularities at zero distance (vertical asymptotes). Attempting to modify interactions by linearly scaling back the LJ potential as a function of an interaction parameter, λ, causes the r^{-12} term to increasingly behave as a sharp repulsive singularity as $\lambda \to 0$ [249]. This means that there is an unbounded discontinuous change between $\lambda = 0$ where particles can overlap, and $\lambda = \delta$, even as $\delta \to 0$, where particles still behave like minuscule hard spheres. This can lead to strongly fluctuating forces/energies and to severe instabilities in the integrator, as well as numerical errors in post processing analyses even when simulations do terminate normally [10, 11, 130].

One additional important issue is whether the code allows holonomic constraints to be applied to bonds, which may change bond lengths in some transformations, e.g. C–H to C–C. Changes in bond length need to account for the associated change in the free energy. These and other details will be outlined below.

AMBER. This code uses a hybrid dual/single topology approach. All terms are energy scaled. The perturbed group must be entirely duplicated, i.e. for `sander` this means two topology files with one end state each, and for `pmemd` both end states in one topology file. In AMBER16 `sander` and `pmemd` implement free energy simulations in an equivalent fashion. However, pmemd does not support vacuum free energy simulations in that version. Hence, all vacuum simulations needed to be run with `sander` while all solution runs were done with `pmemd`.

The code loads two separate input topologies that describe the end states of interest and allows users to map atoms between the two end–states that will share the same coordinates for the free energy calculation. Evaluation of the interactions involving these atoms as a function of the coupling parameter is done by default via linear scaling of the energy and forces of the end–states. Alternatively the user can request that a softcore potential be used. Atoms that are not paired between the end–states are are effectively treated as dummy atoms in one of the two end–states. Bonded terms involving different unpaired atoms are ignored and their non–bonded interactions are handled with a softcore potential. We call this the "implicit dummy protocol" since the procedure is handled automatically by the software through analysis of the end–state topologies rather than via explicit introduction of dummy atoms that is required in computations based on a single topology framework.

The code cannot handle bond length changes involving a constraint. There is only one global λ for parameter transformation. Protocols that couple only some parameters (split protocols, see below) must be emulated through careful construction of topologies. For instance one can keep the LJ and bonded terms fixed at the initial state for a charge transformation. The setup for the two end–states must therefore use identical atom types with only the charges varying.

Alternatively it is possible for the user to construct an input topology of a single molecule that explicitly contains dummy atoms such that the desired end–states can be simulated. This is a similar approach to that employed by SOMD and GROMACS, and we call this the "explicit dummy protocol".

CHARMM. The PERT module duplicates the topology similarly to `sander` but mapped atoms are given in the topology only once. The module requires balancing with explicit dummy atoms. All energy terms are linearly scaled by the coupling parameter λ. The softcore potential (activated with the PSSP keyword and used here as identifier in the further

discussion, see the SI for implementation details) is applied to *all* atoms in the perturbed group (see section 4.5.1). The code can handle constraints of changing bond lengths in the perturbed group but this may cause incorrect results with PSSP softcores (Stefan Boresch, private communication). There is only one global λ for parameter transformation, however, the scripting facilities in CHARMM allow run time modification of topologies e.g. by setting charges or LJ parameters to arbitrary values.

GROMACS. This code uses a single topology description. Bonded terms are strictly parameter–scaled, which requires proper balancing of multi–term dihedrals, i.e. each individual term in the Fourier series must have an equivalent in both end states. If the term does not exist it must be created with parameters zeroing its energy. The softcore potential applies to dummy atoms only determined from atoms having zero LJ parameters in the end states. The code allows changing bond lengths involving constraints within the perturbed group but this can lead to instabilities and wrong results (Michael Shirts, private communication). There are separate lambda values for LJ, Coulomb and bonded parameters (and some other possible terms in the potential) which allows easy implementation of split protocols.

SOMD. SOMD is a software built by linking Sire and OpenMM molecular simulation libraries. [241, 208] This code uses a single topology description. The alchemical state is constructed at run time from an input topology together with a "patch" (list of force field parameters to be modified). All dummy atoms needed to describe the transformation must be present in the initial state. Bond and angle terms are parameter–scaled while the dihedral term is energy–scaled. The softcore potential applies to atoms that become dummy atoms in one end–state. Dummy atoms are specified by a keyword in the patch file. The code cannot handle constraints of changing bond lengths in the perturbed group. There is only one global λ for parameter scaling. Separated protocols (see below) must be emulated through careful construction of the patch file.

4.2.2 RAFE Setup

The setup for all relative free energy simulations has been carried out with the tool FESetup (version 1.2). [247] FESetup is a perturbed topology writer for AMBER, CHARMM, GROMACS, SOMD and NAMD [250] (NAMD is currently a purely dual–topolgy code and has thus not been considered in this study due to the technical differences as explained in the Introduction). The tool makes use of a maximum common substructure search algorithm to automatically compute atoms that can be mapped, i.e. atoms that have a direct relationship to an equivalent atom in the other state – atoms undergoing atom type conversion or modification. The only current limit is that rings are required to be preserved [237]. With this strategy, a single topology description is achieved: any atom that does not match is made a dummy atom. FESetup allows equilibration of the solvated simulation systems and ensures that "forward" and "backward" simulations have the same number of total atoms. With SOMD the mass of each perturbed atom is taken as the mass of the heavier end–state atom (e.g. a hydrogen atom that is perturbed to a carbon atom has an atomic mass of 12 amu at all lambda values). The masses of perturbed atoms are set to the mass of the heavier atom description they are being perturbed to for SOMD. The other codes use the atom masses of the initial state (AMBER, CHARMM) or allow the user to define how masses vary as a function of lambda (GROMACS). The tool creates all input files with control parameters, topologies and coordinates as required for RAFE simulations. Full details on FESetup can be found in Ref. 247.

Fig. **12** shows all 9 transformations, run in forward and backward directions, considered in the present study. In the limit of sufficient sampling, RAFE simulations should not depend on the "forward" and "backward" direction of change with respect to the coupling parameter λ. However to test for possible discrepancies, we have run simulations in both directions. As we shall discuss in the Results section, we do see differences in some cases.

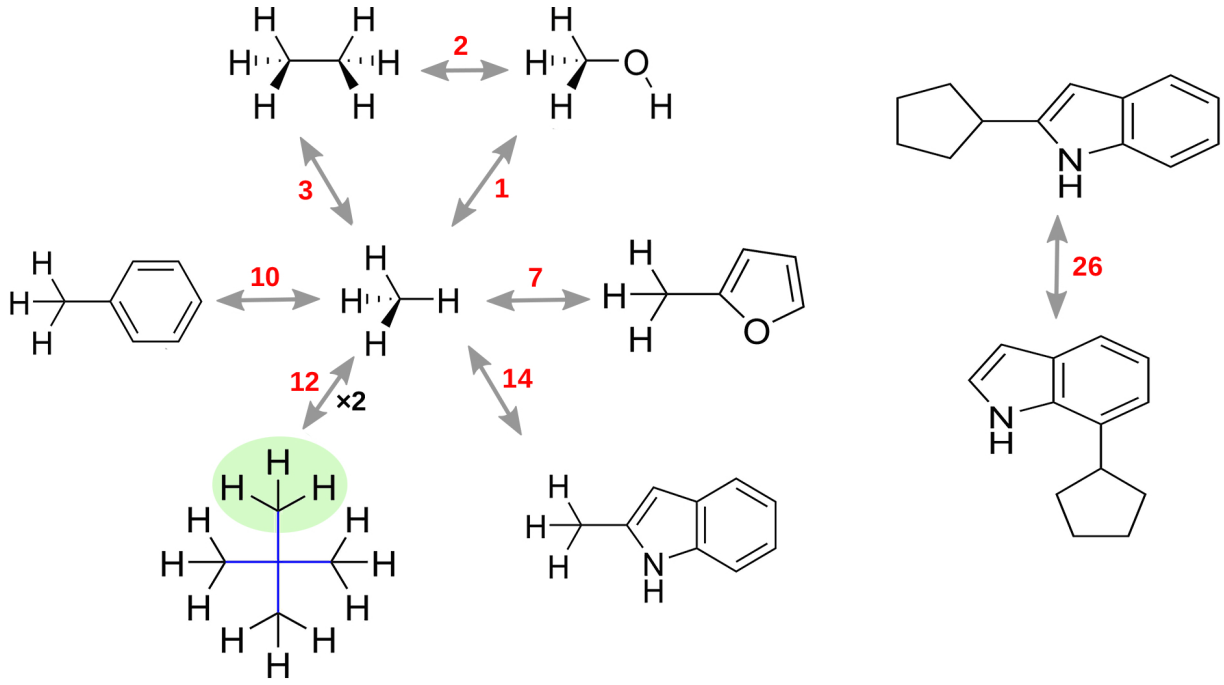

Figure 12: The thermodynamic cycles considered in this study. To compute the free energy of hydration, all pair–wise transformations have to be carried out once in solution and once in vacuum. Green and blue colours in neopentane show two alternative mappings for methane. The numbers in red denote the number of dummy atoms.

The ethane → methanol transformation is traditionally regarded as a standard test for RAFE simulations [15, 251]. The other transformations are centered around mutations from and to methane, and are meant to mimic components of typical transformations that could be attempted in the context of e.g. protein–ligand binding calculations. The 2–cyclopentanylindole to 7–cyclopentanylindole (2–CPI to 7–CPI in our notation) transformation has been added to include both deletion as well as insertion of sub–parts of the perturbed group in one transformation, an aspect not tested by the other transformations. For neopentane → methane two alternative mappings have been considered, see Fig. **12**. One mapping has methane matched to a terminal methyl (green) and the other one has the methane carbon matched with the central carbon in neopentane (blue). The first approach will be called "terminally mapped" and the second one "centrally mapped".

4.2.3 Free Energy Simulation Protocols

One of the major goals of the present study is to ensure consistency and reproducibility from the computational protocols. This is complicated by the fact that a given MD software may employ a range of methods and algorithms that one may not be able to duplicate exactly with other MD software. In particular, how the alchemical transformation is controlled via the coupling parameter may be very different. At the most basic level, important differences could even result from pressure and temperature scaling, integrators and other algorithms. It is unclear if and how any of these implementation details affect results. The implementation details of alchemical free energy simulation in code are discussed in subsection 4.2.1.

In this study we consider a set of simple organic molecules (see Fig. **12**). As the focus here is on probing for reproducibility among various MD packages, we chose fairly small, rigid and neutral molecules to minimize statistical sampling errors, and avoid difficulties with charged particles [232, 252]. The force field was chosen to be GAFF (version 1.8), [67] utilizing AM1-BCC charges for the solute,[157, 158] and TIP3P for the solvent. [159] Charges were computed with the `antechamber` program and missing bonded and vdW terms were generated with the `parmchk2` program, both from the AmberTools16 distribution. All parameters and input files are available at https://github.com/halx/relative-solvation-inputs. The quality of free energies estimated using various small molecule force fields has been discussed elsewhere and is not a focus of this work; here we focus on reproducibility given a particular force field. [253]

While the MD packages employed principally allow a "one–step" transformation [13], that is, with both LJ and Coulombic parameters varied simultaneously (what we call a unified protocol), it has also been proposed that carrying out a split protocol may be more efficient. [254, 135, 131] In such a protocol the charges are transformed linearly between the end states followed by a mutation of the van der Waals parameters using a softcore potential (see

section 4.5.1) on the LJ term only. [10, 11] It is important to note that in the split protocol, charges have to be switched off before LJ parameters (and vice versa for the transformation in opposite direction) to avoid collapse of other atoms, e.g. solvents, onto a "naked" charge,[255, 256, 13] see section 4.5.3.

All simulations were started from simulation boxes prepared by FESetup [247]. During construction of the perturbed systems, steric overlaps between the solute and the solvent may happen. This is because each unperturbed solute is independently equilibrated but the final perturbed system is a composite of coordinates from those potentially differently sized solutes. To make the number of atoms the same for forward and backward setups, the water coordinates of the larger of the two boxes are chosen. Thus, in transformations from a smaller to a larger solute, water molecules may be in close proximity to the solute. At the end of the construction process, FESetup performs a minimization onto the system. In addition, some simulation protocols started with an additional (redundant) minimization step. All production simulations were run at 298 K and 1.0 bar in the NPT ensemble. Water molecules were constrained. Atomic masses were not changed along the alchemical transformations as this would affect only the kinetic energy, and would not contribute to the free energy change. A summary of the main algorithmic differences between each simulation package is given in Tab. **3**.

AMBER. The AMBER16 program was used for this set of free energy calculations. Typically 11 windows were used for charge mutations and 21 windows for VdW mutations. In some instances, steep variations in TI gradients were observed by visual inspection with this protocol and additional windows were added to obtain smoother integration profiles. The starting coordinates were usually taken directly from the pre–equilibrated setup step but no further λ specific equilibration was carried out, i.e. RAFE MD simulations were started with new velocities appropriate for the final simulation temperature. In a very few cases it was

necessary to use coordinates from the end of the simulation at a nearby λ state because of simulation instabilities. This happened in transformations with a larger number of dummy atoms. Absolute transformations were carried out using a one step protocol featuring 21 windows initially. For some perturbations additional windows were run in regions where the free energy gradients varied sharply. Each window was simulated for 2.5 ns, with the first 0.2 ns discarded prior to analysis. Water hydrogens (TIP3P) were constrained with SHAKE. None of the atoms in the perturbed group were constrained and hence the time step was set to 1 fs. An alternative protocol with SHAKE on bonds that do not change during transformation and a time step of 2 fs was also tested (see SOMD protocol below). The temperature was controlled through a Langevin thermostat with a friction constant of $2.0\,\mathrm{ps}^{-1}$ and pressure rescaling through a Monte Carlo barostat with 100 steps between isotropic volume change attempts. Long–range electrostatics in solution was handled with Particle Mesh Eward (PME) and an atom–based cutoff of 8.0 Å for the real-space Coulomb and vdW interactions. No cutoff was used for the vacuum simulations. A Long Range Correction (LRC) term for truncated VdW interactions was applied during the MD simulations.

CHARMM. The version c40b1 was used for this set of free energy calculations. The PERT module was used to handle the alchemical transformations. Three different approaches were used to calculate the relative Gibbs free energy: (i) RAFE simulation where electrostatic and VdW interactions were changed separately (split-protocol), (ii) RAFE simulation where electrostatic and VdW interactions were changed together (unified-protocol), and (iii) difference between free energies from two AFE simulations where AFE simulations followed unified-protocol. In total, 21 evenly spaced windows were used and all windows were run for 1.5 ns with a timestep of 1 fs. Most windows used the same pre-equilibrated configuration. A few windows at the end-points (involving hydrogen being transformed to heavy atom or vice versa) were unstable due to steric clashes with starting coordinates and were equilibrated using 0.1 fs to 0.5 fs. Only water hydrogens (TIP3P) were constrained with SHAKE. Con-

ditions of constant temperature and pressure control were maintained using the Berendsen weak coupling method, with a compressibility of $4.63 \times 10^{-5}\,\mathrm{atm}^{-1}$ and temperature and pressure coupling constants of $5.0\,\mathrm{ps}^{-1}$. Long–range electrostatics in solution was handled with PME to order 6 with a cutoff of 12.0 Å for the real-space Coulomb and vdW interactions. No cutoff was used for the vacuum simulations. No LRC term was applied during the alchemical MD simulations but a solute-solvent LRC term was included in post-processing to calculate the final free energy. The PSSP softcore potential function was used for the perturbed atoms. The PERT module currently does not currently support the force switching (option `VFSwitch`) for LJ potentials with softcores. The CHARMM PARAM27 force fields, however, is parameterized to use force switching [239]. Accordingly, we used the potential switching only (option `VSwitch`) with an inner cutoff of 10 Å and outer cutoff of 12 Å.

GROMACS. GROMACS version 4.6.7 was used to carry out this set of free energy calculations. Each transformation had its Gibbs free energy calculated: (i) in a single topology approach in which LJ energy terms were changed separately from the electrostatic and bonded components; (ii) in a single topology approach in which bonded, LJ, and electrostatic terms are changed together; and (iii) via the difference between two absolute calculations. In the first two cases, each alchemical transformation was described by 31 and 16 states, respectively, and simulated for 4.2 ns with time steps of 1.0 fs in water and vacuum. We used a 20-window alchemical protocol with five windows for charge coupling and fifteen windows for LJ coupling [57, 187]. Our choice allows soft core potentials to be used only when changing nonpolar interactions, and allows electrostatic interactions to be changed linearly [131]. The free energies were calculated from 5 ns Langevin dynamics at 298 K. A friction coefficient of $1.0\,\mathrm{ps}/\mathrm{m}_{\mathrm{atom}}$ was used, where m_{atom} is the the mass of the atom. No holonomic bond or angle constraints for the solutes were used. Waters were constrained with LINCS. A Parrinello–Rahman barostat with $\tau_p = 10\,\mathrm{ps}$ and compressibility equal to $4.5 \times 10^{-5}\,\mathrm{bar}^{-1}$ was used. Two methods were used to calculate electrostatic interactions: Particle Mesh

Ewald (PME) and charge group-based Reaction Field with a dielectric of 78.3, as implemented in the software. PME calculations were of order 6 and had a tolerance of 1.0×10^{-6}, with a grid spacing of 1.0 Å. We set the real-space electrostatic and VdW cutoffs to 10.0 Å; a switch was applied to the latter starting at 9.0 Å. A cutoff of 50.0 Å was used for the vacuum simulations. A Long Range Correction (LRC) term for truncated VdW interactions was applied during the MD simulations. All transformations required the use of softcore potentials to avoid numerical problems in the free energy calculation. We chose the 1–1–6 softcore potential for LJ terms (α=0.5 and σ=0.3) for atoms whose parameters were being perturbed and used the default softcore Coulomb implementation in paths where charges, LJ, and bonded terms were modified together, but no soft core potentials were applied to Coulomb interactions when electrostatic interactions were modified separately.

SOMD. This set of free energy calculations was carried out with SOMD from the Sire 2016.1 release. [241, 208] Each alchemical transformation was divided into 17 evenly spaced windows and simulated for 2 ns each both in water and in vacuum. The absolute hydration free energies were computed by annihilating non-bonded interactions of the solute in two steps. In the first step the free energy change for discharging the solute was computed. In the second step the free energy change for turning off the Lennard-Jones terms of the discharged solute was computed. Each step was carried out using 17 evenly spaced windows. The starting coordinates for each window were obtained by an additional energy minimization of the same pre-equilibrated and minimized configuration generated by FESetup. A velocity-Verlet integrator was employed with a 2 fs time step. Water hydrogens (TIP3P) were constrained with SHAKE. For the alchemical solute, only bonds involving hydrogens which are not alchemically transformed were constrained. This approach is referred as the "unperturbed H bond constraint protocol". Given the number of the perturbed hydrogen bonds in the solutes (Fig. **12**, this constraint allows to use a 2 fs time-step through use of higher atomic masses for perturbed hydrogen atoms (see discussion below). Temperature

Table 3: Summary of the technical details for the relative hydration free energy calculations carried out with the various codes.

	AMBER	CHARMM	GROMACS	SOMD
Version	AMBER16	c40b1	4.6.7	2016.1
Module	`pmemd, sander`	`PERT`	`gmx`	`somd-freenrg`
Protocol	Split protocol	Unified protocol	Split protocol	Unified protocol
Number of λ windows	11 (charge mutations) 21 (vdW mutations)	21 evenly spaced	31 (charge mutations) 31 (vdW mutations)	17 evenly spaced
Starting coordinates	FESetup pre-equilibration	FESetup pre-equilibration	FESetup pre-equilibration	FESetup pre-equilibration
Simulation length per window	2.5 ns	1.5 ns	4.2 ns	2 ns
Timestep	1 fs	1 fs	1fs	2fs
Electrostatic method	PME	PME	PME	atom-based RF
Solvated phase cutoff	8 Å	12 Å	10 Å	10 Å
Vacuum phase cutoff	no cutoff	no cutoff	50 Å	no cutoff
Constraint	none	none	none	H-bonds not perturbed
LRC corrections	during MD	post-processing	during MD	post-processing
Barostat	Monte Carlo	Berendsen	Parrinello-Rahman	Monte Carlo
Thermostat	Langevin	Berendsen	Langevin	Andersen
Soft core parameters	$r_{LJ} = (2\sigma_{ij}^6\lambda + r_{ij}^6)^{1/6}$ $r_{Coul} = (\beta\lambda + r_{ij}^p)^{1/p}$ $n = 1$	$r_{LJ} = (2\lambda + r_{ij}^2)^{1/2}$ $r_{Coul} = (\beta\lambda + r_{ij}^2)^{1/2}$ $n = 1$	$r_{LJ} = (2\sigma_{ij}^6\lambda + r_{ij}^6)^{1/6}$ $r_{coul} = r_{LJ}$ $n = 1$	$r_{LJ} = (2\sigma_{ij}\lambda + r_{ij}^2)^{1/2}$ $r_{Coul} = (\lambda + r_{ij}^2)^{1/2}$ $n = 1$

control was achieved with the Andersen thermostat, [257] with a stochastic collision frequency of $10\,\mathrm{ps}^{-1}$. A Monte Carlo barostat assured pressure control, with isotropic box edge scaling moves attempted every 25 time steps. A shifted atom–based Barker–Watts reaction field, [258] with a dielectric constant of 78.3 was adopted for the solution phase simulations with a cutoff of 10 Å. A similar cutoff was used for LJ interactions. The reaction field was not employed in the vacuum legs, where a Coulombic potential without cutoff was used. A protocol to account for the different treatment of intramolecular electrostatics in vacuum and solution is described in the supporting information. The softcore parameters (Eq. 4.5) were set to default values for all the transformations, specifically $n = 0$ for Coulombic interactions and $\alpha = 2.0$ for the LJ potential [231]. Additionally, an end-point correction for truncated VdW potentials was applied by post-processing of end-state trajectories as described previously elsewhere. [151, 259]

4.2.4 Free Energy Estimations

In this work we primarily focus on TI as this is supported by all the tested MD packages "out–of–the–box". Equation 4.1 computes the free energy as

$$\Delta G = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\mathscr{H}(\mathbf{q}, \mathbf{p}; \lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \tag{4.1}$$

where $\mathscr{H}(\mathbf{q}, \mathbf{p}; \lambda)$ is the Hamiltonian as a function of the coordinate vectors $\mathbf{q}$ and the momentum vectors $\mathbf{p}$, and parametric dependence on the coupling parameter λ is explicit. The angle brackets denote the ensemble average of the gradient of the Hamiltonian with respect to λ, at a given λ value. Results from additional estimators will be given where available. We have used the `alchemical analysis tool` [128] for all analyses. This tool provides various estimators such as TI, TI with cubic splines, BAR and MBAR. We have used the cubic splines method to integrate the free energy. All data was sub–sampled to eliminate correlated data [260].

All RAFE simulations were run in triplicate in forward as well as backward direction for a total of 6 simulations per mutation. The final hydration free energy $\Delta\Delta G_{\mathrm{hydr}}$ was computed as the average for each direction separately. For comparison we have also calculated the absolute (standard) hydration free energies for all molecules in Fig. **12**.

To estimate the reliability and convergence of the results, the standard error of the mean (SEM) has been calculated. The SEM is defined as

$$\mathrm{err}(\Delta\Delta G_{\mathrm{hydr}}) = \frac{\sigma}{\sqrt{n}} \tag{4.2}$$

where σ is the sample standard deviation of the three $\Delta\Delta G_{\mathrm{hydr}}$ values, and $n = 3$. For each free energy change the SEM was evaluated as:

$$\mathrm{err(combined)} = \sqrt{\sum_i \sigma_i^2}. \tag{4.3}$$

We also make use of the mean absolute error MAE (also called mean unsigned error, MUE) to compare data sets.

$$\mathrm{MAE} = \frac{1}{N} \sum_{\mathrm{i}=1}^{N} |y_{\mathrm{i}} - x_{\mathrm{i}}| \tag{4.4}$$

where N is the total number of samples, y_{i} and x_{i} are the i–th datum to be compared.

4.3 Results

4.3.1 Overall comparison

In the following we will present our RAFE results for the thermodynamic cycles shown in Fig. **11**. We will use absolute hydration free energies here as our standard point of comparison because for the present dataset they can be calculated with high precision [187], and are simpler to set up and implement than relative calculations.

Tab. **4** summarizes results for the absolute hydration free energies. The table shows the data from simulations with the protocol our groups considered most trustworthy for the respective MD code used, as discussed in detail in the following subsections. The precision of the calculated free energies is similar between AMBER, CHARMM and GROMACS, whereas the SOMD free energies are less precise. This may reflect differences in the lambda schedules and length of trajectories between the different codes. Nonetheless the standard

Table 4: Absolute hydration free energies (in kcal/mol) and end-state densities (in g/cm3) as obtained from AFE calculations. Uncertainties on the last decimal are given in parenthesis.

Solute	AMBER		CHARMM		GROMACS		SOMD	
	Free energy (kcal/mol)	Density (g/cm^3)	Free energy (kcal/mol)	Density (g/cm^3)	Free energy (kcal/mol)	Density (g/cm^3)	Free energy (kcal/mol)	Density (g/cm^3)
methane	2.47(1)	0.986(1)	2.48(1)	0.977(1)	2.44(1)	0.987(1)	2.52(2)	0.982(1)
methanol	-3.73(1)	0.988(1)	-3.72(1)	0.980(1)	-3.51(1)	0.988(1)	-3.70(5)	0.987(1)
ethane	2.50(1)	0.988(1)	2.50(1)	0.979(1)	2.48(1)	0.988(1)	2.56(1)	0.984(1)
toluene	-0.72(1)	0.991(1)	-0.64(1)	0.983(1)	-0.72(1)	0.991(1)	-0.55(2)	0.989(1)
neopentane	2.61(1)	0.990(1)	2.58(2)	0.981(1)	2.58(1)	0.990(1)	2.71(6)	0.987(1)
2-methylfuran	-0.49(2)	0.991(1)	-0.42(1)	0.983(1)	-0.51(1)	0.991(1)	-0.39(2)	0.989(1)
2-methylindole	-6.24(1)	0.993(1)	-6.06(1)	0.984(1)	-6.35(1)	0.993(1)	-6.06(4)	0.990(1)
2-CPI	-6.05(2)	0.995(1)	-6.18(4)	0.992(1)	-6.54(1)	0.994(1)	-6.14(9)	0.991(1)
7-CPI	-5.66(3)	0.995(1)	-6.28(3)	0.982(1)	-6.52(2)	0.995(1)	-6.1(1)	0.992(1)

errors are typically well under 0.1 kcal/mol, thus it becomes meaningful to investigate small differences of a few tenths of kcal/mol between codes.

The ΔG_{hydr} obtained with the various MD packages in this way agree quite well given statistical errors, although some larger deviations are apparent as well. GROMACS predicts a smaller ΔG_{hydr} for methanol by about $0.2\,\mathrm{kcal\,mol^{-1}}$. The largest deviation can be found for one of the largest molecules (7–CPI) with the AMBER result being less negative than with the other MD packages by 0.4–$0.8\,\mathrm{kcal\,mol^{-1}}$. This particular discrepancy does not correlate with significant variations in density between AMBER and other codes.

As an additional check we computed densities in the fully decoupled states and compared the results to reported densities for a pure TIP3P water box. The average densities across all simulations are $(0.980 \pm 0.002)\,\mathrm{g/cm^3}$, $(0.973 \pm 0.002)\,\mathrm{g/cm^3}$, $(0.979 \pm 0.002)\,\mathrm{g/cm^3}$, $(0.976 \pm 0.003)\,\mathrm{g/cm^3}$ for AMBER, CHARMM, GROMACS and SOMD respectively. AMBER and GROMACS show higher densities presumably because a LRC term was applied during the MD simulations, whereas LRC terms for SOMD and CHARMM are only applied via post-processing of trajectories. For reference, a recent study from Wang et al. reports a TIP3P water density of $0.980\,\mathrm{g/cm^3}$ [261].

Table 5: Mean Absolute Error (MAE) (kcal mol^{-1}) between relative free energies obtained with the absolute protocol for the SOMD, GROMACS, AMBER and CHARMM packages.

Package	GROMACS	AMBER	CHARMM
SOMD	0.20 ± 0.03	0.13 ± 0.04	0.08 ± 0.02
GROMACS		0.19 ± 0.01	0.15 ± 0.01
AMBER			0.12 ± 0.01

Tab. **5** shows the MAE between SOMD, GROMACS, AMBER and CHARMM. CHARMM produces figures that agree the most with other MD packages. The largest difference reaches 0.2 kcal mol^{-1} for SOMD and GROMACS. Variabilities between the codes may be partly explained by differences in densities due to different treatments of long range electrostatics and vdW interactions.

Having established the predictive value from absolute transformations we now turn to computing $\Delta\Delta G_{\text{hydr}}$ from relative mutations. Tab. **6** summarizes the results for the four MD packages. Again the data is from the recommended protocol for each package (see detailed discussions in the following subsections).

We reviewed firstly internal consistency of the different codes with the computed absolute hydration free energies. For each implementation we counted the number of times a calculated relative free energy deviates from the difference in reference absolute hydration free energies by more than 0.1 kcal/mol. This is significantly above the estimated uncertainties in calculated free energies in most instances. According to this criterion, the AMBER explicit implementation is the least consistent (10 deviations), followed by AMBER implicit (6 deviations), SOMD (6 deviations), CHARMM (5 deviations), GROMACS (5 deviations). The perturbations that give a discrepancy are not the same across codes, for instance methane->toluene with AMBER explicit deviates from the reference absolute hydration free energies by 0.33 kcal/mol, but at most 0.04 kcal/mol with other codes. SOMD and GROMACS show deviations of ca. 0.25 kcal/mol for methanol->methane but this is not the case for AMBER (implicit or explicit) or CHARMM.

Table 6: Comparison of relative free energies of hydration for various MD packages as obtained from absolute (AFE) and relative (RAFE) transformations via unified or split protocols. The values deduced from AFE transformations (given in the first row) were taken from Tab. 1. Signs of the backward transformation have been reverted to correspond to the forward transformation.

Transformation[a]		AMBER[b]		CHARMM[c]	GROMACS[b]	SOMD[c]
		implicit[d]	explicit[d]			
ethane	methane		-0.02 ± 0.01	-0.03 ± 0.01	-0.04 ± 0.01	-0.05 ± 0.02
ethane	methane	0.02 ± 0.01	-0.13 ± 0.02	-0.09 ± 0.02	-0.04 ± 0.02	0.05 ± 0.02
methane	ethane	0.00 ± 0.03	-0.19 ± 0.03	-0.04 ± 0.01	-0.02 ± 0.01	0.01 ± 0.06
methanol	methane		6.20 ± 0.01	6.20 ± 0.02	5.95 ± 0.01	6.21 ± 0.06
methanol	methane	6.19 ± 0.01	6.20 ± 0.02	6.18 ± 0.01	6.20 ± 0.01	5.99 ± 0.05
methane	methanol	6.20 ± 0.03	6.15 ± 0.01	6.21 ± 0.01	6.20 ± 0.01	5.97 ± 0.04
ethane	methanol		-6.22 ± 0.01	-6.22 ± 0.02	-5.98 ± 0.01	-6.26 ± 0.05
ethane	methanol	-6.20 ± 0.01	-6.27 ± 0.01	-6.25 ± 0.01	-6.19 ± 0.01	-6.09 ± 0.03
methanol	ethane	-6.20 ± 0.01	-6.25 ± 0.01	-6.28 ± 0.01	-6.19 ± 0.01	-6.09 ± 0.02
toluene	methane		3.19 ± 0.01	3.12 ± 0.01	3.16 ± 0.01	3.07 ± 0.03
toluene	methane	3.24 ± 0.02	3.39 ± 0.02	3.04 ± 0.02	3.21 ± 0.01	2.89 ± 0.09
methane	toluene	3.42 ± 0.03	3.52 ± 0.03	3.09 ± 0.02	3.20 ± 0.01	3.06 ± 0.02
neopentane	methane		-0.13 ± 0.02	-0.11 ± 0.02	-0.14 ± 0.01	-0.19 ± 0.06
neopentane[e]	methane	0.32 ± 0.04	-0.03 ± 0.06	-0.35 ± 0.01	-0.15 ± 0.02	-0.20 ± 0.05
methane[a]	neopentane	0.25 ± 0.03	-0.07 ± 0.03	-0.24 ± 0.02	-0.16 ± 0.05	-0.13 ± 0.05
neopentane[f]	methane	-0.13 ± 0.01	-0.12 ± 0.02	-0.56 ± 0.02	-0.14 ± 0.01	-0.11 ± 0.01
methane[b]	neopentane	-0.13 ± 0.03	-0.12 ± 0.03	-0.40 ± 0.02	-0.18 ± 0.03	-0.10 ± 0.06
2–methylfuran	methane		2.96 ± 0.02	2.90 ± 0.01	2.95 ± 0.01	2.90 ± 0.03
2–methylfuran	methane	3.09 ± 0.01	3.10 ± 0.01	2.84 ± 0.03	2.93 ± 0.05	2.92 ± 0.05
methane	2-methyfuran	3.10 ± 0.03	3.15 ± 0.03	2.84 ± 0.02	2.96 ± 0.01	2.83 ± 0.03
2–methylindole	methane		8.72 ± 0.01	8.53 ± 0.02	8.79 ± 0.02	8.57 ± 0.03
2–methylindole	methane	8.78 ± 0.03	8.78 ± 0.04	8.49 ± 0.01	8.73 ± 0.03	8.64 ± 0.06
methane	2-methylindole	9.14 ± 0.02	9.13 ± 0.03	8.56 ± 0.02	8.74 ± 0.01	8.67 ± 0.08
2–CPI	7–CPI		0.39 ± 0.04	-0.11 ± 0.04	0.02 ± 0.05	0.08 ± 0.14
2–CPI[g]	7–CPI	0.36 ± 0.03	0.63 ± 0.06	-0.01 ± 0.01	-0.01 ± 0.03	-0.11 ± 0.07
7–CPI[g]	2–CPI	0.34 ± 0.05	0.50 ± 0.03	0.04 ± 0.01	-0.20 ± 0.04	-0.01 ± 0.08

[a]The values deduced from the AFE absolute of Table 1 are given first.
[b]split protocol.
[c]unified protocol.
[d]using either the implicit or the explicit dummy atom approach.
[e]central mapping.
[f]terminal mapping.
[g]partial re/discharge i.e. only the charges of the appearing and the disappearing 5–rings are switched.

Table 7: MAE (in kcal mol^{-1}) comparing relative free energies from relative simulations between SOMD, GROMACS, AMBER and CHARMM.

Package	GROMACS	AMBER	CHARMM
SOMD	0.11 ± 0.01	0.23 ± 0.01	0.15 ± 0.01
GROMACS		0.16 ± 0.01	0.13 ± 0.01
AMBER			0.26 ± 0.01

We next reviewed consistency between forwards and backwards relative hydration free energies. Again counting the number of deviations that exceed 0.1 kcal/mol indicates that AMBER explicit is the least consistent (3 deviations), followed by AMBER implicit (2 deviations), CHARMM (2 deviations), GROMACS (1 deviation), SOMD (1 deviation). The largest deviation is observed with AMBER implicit for 2-methylindole <-> methane (0.36 kcal/mol).

Next we compared relative free energies across packages. CHARMM tends to show relative free energies with smaller values for a number of transformations: neopentane, 2–methylfuran and 2–methylindole. SOMD displays smaller values $\Delta\Delta G_{\mathrm{hydr}}$ for the methanol and toluene transformations. The largest discrepancy, however, is in the neopentane transformation with central mapping where AMBER with implicit dummy atoms is about 0.5 kcal mol^{-1} higher and CHARMM about 0.2 kcal mol^{-1} lower than the other two codes. The terminal mapped neopentane case reveals AMBER to be in line with GROMACS and SOMD while CHARMM's results deviate further. AMBER deviates also quite strongly from the other codes in the cyclopentanyl indole cases. It is possible that the discrepancies observed with AMBER are partly due to inconsistencies in the end point geometries (see section 4.3.2).

The MAEs of the relative free energy simulations are presented in Tab. **7**. They are on average slightly larger than the MAEs from the absolute simulations (Tab. **5**) and reach 0.26 kcal mol^{-1} for AMBER compared with CHARMM.

Table 8: Cycle closure errors in kcal mol^{-1}) for ethane→ methanol. → methane → ethane. Uncertainties denote a 95% confidence interval.

Package and Protocol	Closure Error
AMBER implicit	0.07 ± 0.08
AMBER explicit	0.02 ± 0.10
GROMACS split reaction field	0.05 ± 0.04
GROMACS unified reaction field	0.13 ± 0.06
GROMACS split PME	0.04 ± 0.02
GROMACS unified PME	0.18 ± 0.06
CHARMM	0.01 ± 0.06
SOMD	-0.11 ± 0.16

We also computed cycle closure errors from Tab. **8** for the closed cycle ethane→ methanol → methane → ethane (see Fig. **12**). The results are shown in Tab. **8**. Uncertainties were estimated by propagating uncertainties from the individual perturbations. The AMBER protocols, CHARMM and SOMD are consistent within uncertainty estimates, but the deviations observed with the GROMACS protocols are small. The largest discrepancy is observed with the GROMACS unified PME protocol, with the error just under 0.2 kcal/mol.

Finally we also examined whether the codes reproduced consistent changes in mean box volumes between forward and backward transformations. We find that the codes are generally consistent with GROMACS giving the most precise volume changes, whereas SOMD gives the least precise volume changes (See Tab. S1 in the SI). This indicates that the barostats used by the different simulation packages relax volume fluctuations with different efficiency, or that they sample different volume fluctuations.

We now turn to considerations for individual packages.

4.3.2 AMBER

Using AMBER for RAFE simulations has revealed several problems with the implementation. Some issues were identified and the developers have fixed those for AMBER16, e.g. energy minimization in `sander` led to diverged coordinates for mapped atoms. For a single topology description, however, it is necessary to have the same coordinates. Other issues are that vacuum simulations can only be carried out with the `sander` program because `pmemd` cannot handle AFE simulations in vacuum as of this writing. This will, however, be rectified in future versions [262]. A disadvantage of `sander` is that it cannot be used to simulate the λ end points, [263] such that the TI gradients need to be extrapolated (minimum and maximum allowed λs are 0.005 and 0.995). Also, `sander` considers the whole system as the perturbed region while `pmemd` restricts this to a user chosen atom selection. This has obvious implications for performance [263].

We also found that, in contrast to the other three codes, AMBER does not yield correct relative free energies with the unified protocol, i.e. when all force field parameters are scaled simultaneously (see Tab. **14**). The issue becomes apparent when more than a few dummy atoms are involved, while the unified protocol works for the smaller transformations (refer to Fig. **12**). The split RAFE protocol and absolute free energies, however, are very close to the other MD packages as demonstrated in Tab. **9** below.

End point geometries appear to be another issue with AMBER simulations both in solution and vacuum. This is most obvious in the neopentane $\rightarrow$ methane test case with central mapping (see RAFE Setup and Fig. **11**). As shown in Fig. **16**, the methane end state exhibits incorrect distances between the carbon and the four attached hydrogens of approximately 1.23 Å. This value is about 1.12 Å for the terminal dummy atoms in the other test cases but still higher than the expected 1.09 Å on average. Fig. **16** demonstrates how this depends on the number of dummy atoms immediately surrounding the central atom.

Table 9: Comparing AMBER results for simulations with various split protocols. The emphasis is here on the data with SHAKE enabled and a time step of 2 fs (last column). Implicit, explicit and absolute protocols had SHAKE disabled and a time step of 1 fs. Signs of the backward transformation have been reverted to correspond to the forward transformation.

transformation		implicit $\Delta\Delta G$	explicit $\Delta\Delta G$	absolute ΔG	SHAKE[a] $\Delta\Delta G$
ethane	methanol	-6.20 ± 0.01	-6.27 ± 0.01	-6.22 ± 0.01	-6.18 ± 0.01
methanol	ethane	-6.20 ± 0.01	-6.25 ± 0.01		
toluene	methane	3.24 ± 0.02	3.39 ± 0.02	3.19 ± 0.01	3.27 ± 0.03
methane	toluene	3.42 ± 0.03	3.52 ± 0.03		
neopentane[b]	methane	0.32 ± 0.04	-0.03 ± 0.06		0.35 ± 0.02
methane[b]	neopentane	0.25 ± 0.03	-0.07 ± 0.03	-0.13 ± 0.02	
neopentane[c]	methane	-0.13 ± 0.01	-0.12 ± 0.02		
methane[c]	neopentane	-0.13 ± 0.03	-0.12 ± 0.03		

[a]implicit dummy atom protocol with $\delta t = 2$ fs and SHAKE on all H–bonds except perturbed bonds.
[b]central mapping.
[c]terminal mapping.

We also compare free energies obtained from the implicit dummy approach in AMBER with results from explicit dummy atom simulations and results from absolute transformations described in Tab. **4** and **6**. The relative simulations have been carried out with the split protocol while the absolute simulations used a unified protocol throughout. SHAKE was explicitly deactivated for all bonds in the perturbed region in these protocols. Tab. **9** shows selected results for transformations with SHAKE enabled for all bonds to hydrogens except those bonds that change bond length during transformation.

The time step has been increased from 1 fs as used in the other three protocols to 2 fs. As the results are essentially the same as the non–SHAKE simulations, this SHAKE protocol appears to be a viable solution to increase the performance of RAFE simulations. We have repeated this protocol with AMBER in response to the results obtained with SOMD using this implementation. From a practical point of view, AMBER uses an *atom* based mask for application of bond constraints such that the mask must be set for the hydrogens in question

while the same is not possible for their non–H counter–part in the other state because *all* bonds emanating from the atom would be affected.

In general, the free energies computed with each approach are in good agreement with each other and with the results of the other MD packages (Tab. **4** and **6**). There are, however, a few notable deviations. Neopentane → methane with central mapping differs from the result with terminal mapping by about $0.4\,\mathrm{kcal\,mol^{-1}}$. The terminal mapping and the free energies from the explicit dummy simulations are, however, consistent with the absolute transformations (Tab. **4**). We also observe a systematic deviation between forward and backward vacuum transformations in the 2–methylindole simulation (see Tab. **15**). The gradient is consistently shifted by 0.2–$0.4\,\mathrm{kcal\,mol^{-1}}$ for each λ step of the vdW plus bonded transformation with both implicit and explicit dummy atoms.

4.3.3 CHARMM

CHARMM for alchemical free energy calculation (AFE) has been widely used with the PERT module, but few issues not previously reported in CHARMM c40b1 were found and careful AFE setup is needed to produce robust and accurate results. Bugs regarding TI gradient accumulation in the parallel version were identified and fixed by Dr. Stefan Boresch. The PERT module does not allow a hydrogen bond constraint (SHAKE) to be applied on the perturbed region, and this requires end point lambdas to be equilibrated carefully. These windows at end-point lambda were started with their own equilibration using timesteps of 0.1 fs to 0.5 fs before the production run. The VSwitch option was used to apply a switching function to the potential since that option cannot be applied to forces for calculations run with the PERT module.

The PSSP softcore potential function cannot handle Long-Range Correction (LRC) correctly. This effect is not clearly shown when the initial and final states are comparable in size, but

Table 10: Comparing CHARMM results for simulations with various split protocols. Signs of the backward transformation have been reverted to correspond to the forward transformation.

transformation		split $\Delta\Delta G$	unified $\Delta\Delta G$	absolute(unified) $\Delta\Delta G$
ethane	methane	-0.09 ± 0.01	-0.09 ± 0.02	-0.03 ± 0.01
methane	ethane	-0.04 ± 0.01	-0.04 ± 0.01	
methanol	methane	6.20 ± 0.01	6.18 ± 0.01	6.20 ± 0.01
methane	methanol	6.30 ± 0.01	6.21 ± 0.01	
ethane	methanol	-6.21 ± 0.01	-6.25 ± 0.01	-6.22 ± 0.02
methanol	ethane	-6.25 ± 0.01	-6.28 ± 0.01	
toluene	methane	3.22 ± 0.01	3.04 ± 0.02	3.12 ± 0.01
methane	toluene	3.28 ± 0.01	3.09 ± 0.02	
neopentane[a]	methane	-0.29 ± 0.01	-0.35 ± 0.01	-0.11 ± 0.02
methane[a]	neopentane	-0.15 ± 0.01	-0.24 ± 0.02	
neopentane[b]	methane	-0.42 ± 0.01	-0.56 ± 0.02	
methane[b]	neopentane	-0.31 ± 0.01	-0.40 ± 0.02	
2-methylfuran	methane	2.87 ± 0.01	2.84 ± 0.03	2.90 ± 0.01
methane	2-methylfuran	2.93 ± 0.01	2.84 ± 0.02	
2-methylindole	methane	8.88 ± 0.01	8.49 ± 0.01	8.53 ± 0.02
methane	2-methylindole	8.81 ± 0.01	8.56 ± 0.02	
2-CPI	7-CPI	-0.02 ± 0.01	-0.01 ± 0.01	-0.11 ± 0.04
7-CPI	2-CPI	-0.01 ± 0.01	0.04 ± 0.01	

[a]central mapping.
[b]terminal mapping.

the deviation becomes larger for perturbations that involve large changes in solute size, or for absolute alchemical free energy calculations. It is necessary to disable the LRC to obtain consistent free energies from relative and absolute alchemical free energy calculation protocols (see SI for details).

Tab. **10** shows the relative free energies obtained from CHARMM simulations. While results from all three protocols (split, unified, absolute) seem to be in good agreement with each other, the split-protocol results are more precise due to the additional amount of data generated. It is notable that the split-protocol results are more similar to the ones obtained by other MD packages (i.e. neopentane and toluene), but the relative-unified results are

more consistent with the CHARMM absolute simulations (e.g. 2-methylindole). Overall, the relative free energies obtained by these three different protocols are in good agreement with those reported for the other MD packages (Tab. 1 and 3).

4.3.4 GROMACS

GROMACS has some run input options which can simplify the procedure for setting up free energy calculations. Specifically, `couple-moltype` implicitly defines the initial and final states by giving a special tag to a molecule and controls whether intramolecular interactions of the tagged molecule are retained or not along the alchemical path. It should be used in absolute free energy calculations to tag the molecule which will be decoupled from the rest of the system. Using this in relative calculations is possible, but will result in unintended behavior and errors. The keywords `couple-lambda0` and `couple-lambda1` control the interactions of the molecule specified by `couple-moltype` with its surroundings. The entries `vdw-lambdas` and `fep-lambdas` define the lambda schedule. The former indicates the value of the λ vector component that modifies van der Waals interactions for each state, while the latter changes all λ vector components that are not specified in the `.mdp` file.

Here, we use these options to simplify our setup. For instance, in split protocol simulations, these entries are sets such that the components of the energy are modified in different stages. If the transformation involves particle deletion ("forward process"), `fep-lambdas` is set to change charges and bonds before `vdw-lambdas` changes van de Waals components. If the process involves particle insertion ("backward process") we reverse the roles. In this work, `mass-lambdas` were all set to zero to avoid mass changes during the the free energy calculations. Unified protocols set all λ vectors the same.

Tab. **11** lists the relative free energies obtained from GROMACS simulations. Relative free energies are in good agreement with each other and with $\Delta\Delta G_{\text{hydr}}$ obtained from the

Table 11: Relative hydration free energies obtained from GROMACS simulations in $kcal \cdot mol^{-1}$. Signs of the backward transformation have been reverted to correspond to the forward transformation.

transformation		split[a] RF $\Delta\Delta G$	split[a] PME $\Delta\Delta G$	unified[b] RF $\Delta\Delta G$	unified[b] PME $\Delta\Delta G$	absolute[c] RF $\Delta\Delta G$	absolute[c] PME $\Delta\Delta G$
ethane	methane	-0.025 ± 0.005	-0.035 ± 0.020	-0.017 ± 0.003	-0.030 ± 0.001	-0.06 ± 0.01	-0.04 ± 0.01
methane	ethane	-0.01 ± 0.02	-0.02 ± 0.01	0.046 ± 0.020[d]	0.01 ± 0.02		
methanol	methane	6.163 ± 0.006	6.197 ± 0.004	7.30 ± 0.02	7.380 ± 0.007	5.77 ± 0.01	5.95 ± 0.01
methane	methanol	6.168 ± 0.005	6.199 ± 0.008	7.09 ± 0.02	7.17 ± 0.02		
ethane	methanol	-6.123 ± 0.007	-6.185 ± 0.006	-7.117 ± 0.005	-7.21 ± 0.02	-5.83 ± 0.01	-5.98 ± 0.01
methanol	ethane	-6.124 ± 0.005	-6.193 ± 0.004	-7.338 ± 0.004	-7.404 ± 0.004		
toluene	methane	3.22 ± 0.01	3.211 ± 0.006	3.229 ± 0.008	3.22 ± 0.01	2.97 ± 0.01	3.16 ± 0.01
methane	toluene	3.25 ± 0.01	3.20 ± 0.01	3.22 ± 0.01	3.211 ± 0.001		
neopentane[e]	methane	-0.103 ± 0.008	-0.15 ± 0.02	-0.08 ± 0.02	-0.18 ± 0.03	-0.18 ± 0.01	-0.14 ± 0.01
methane[a]	neopentane	-0.11 ± 0.02	-0.16 ± 0.05	0.00 ± 0.03	-0.18 ± 0.03		
neopentane[f]	methane	-0.116 ± 0.007	-0.13 ± 0.01	-0.14 ± 0.01	-0.14 ± 0.01		
methane[b]	neopentane	-0.10 ± 0.03	-0.18 ± 0.03	-0.089 ± 0.007	-0.15 ± 0.02		
2-methylfuran	methane	2.986 ± 0.006	2.930 ± 0.050	3.05 ± 0.01	3.00 ± 0.01	2.87 ± 0.01	2.95 ± 0.01
methane	2-methylfuran	3.007 ± 0.004	2.96 ± 0.01	3.056 ± 0.006	3.01 ± 0.01		
2-methylindole	methane	8.71 ± 0.02	8.73 ± 0.03	8.73 ± 0.01	8.80 ± 0.03	8.44 ± 0.02	8.79 ± 0.02
methane	2-methylindole	8.73 ± 0.03	8.74 ± 0.01	8.30 ± 0.02	8.77 ± 0.04		
2-CPI	7-CPI	-0.07 ± 0.02	-0.03 ± 0.03	-0.10 ± 0.05	-0.2 ± 0.1	-0.02 ± 0.05	0.02 ± 0.02
7-CPI	2-CPI	-0.12 ± 0.06	-0.20 ± 0.04	-0.04 ± 0.06	-0.14 ± 0.09		

[a]results obtained from alchemical transformations with electrostatic and bonded scaling separate from vdW parameter change.
[b]results obtained from alchemical transformation with all parameters scaling together.
[c]results obtained from absolute free energy calculations.
[d] inverted sign
[e]central mapping
[f]terminal mapping

Table 12: Relative hydration free energies of methanol → methane and methane → methanol transformations without and with the use of Coulomb softcore potentials from GROMACS. Signs of the backward transformation have been reverted to correspond to the forward transformation. The complete version of this table is in the SI.

		split		split+sc		absolute	
		RF	PME	RF	PME	RF	PME
transformation		$\Delta\Delta G$	$\Delta\Delta G$	$\Delta\Delta G$	$\Delta\Delta G$	$\Delta\Delta G$	$\Delta\Delta G$
methanol	methane	6.163 ± 0.006	6.197 ± 0.004	7.32 ± 0.03	7.42 ± 0.04	5.77 ± 0.01	5.95 ± 0.01
methane	methanol	6.168 ± 0.005	6.199 ± 0.008	7.14 ± 0.03	7.21 ± 0.03		

other software used in this study (Tab. **4** and **6**). A noteworthy exception is the difference between the unified and split results of methane → methanol and its reverse process. This was investigated further with additional split protocol simulations using Coulomb softcore potentials (Tab. **12**).

We noticed a difference of approximately $1.5\,\mathrm{kcal\,mol^{-1}}$ between the split protocol without Coulomb softcore potentials and both protocols that use it. The data shown in Fig. **20** suggests that softening of the electrostatic interactions requires adjustments in the λ-distance between states in the rapidly varying part of the $\partial\mathcal{H}/\partial\lambda$. A variant that combined the bonded terms with the vdW transformation did not change this result. Thus, we find that the split protocol without Coulomb softcore potentials is the most effective way to calculate relative free energies with the current GROMACS implementation.

Additionally it is worth mentioning that relative free energy simulations that feature alchemical transformations of a hydrogen atom into a heavy atom will crash if the bond involving the hydrogen atom is constrained with algorithms such as SHAKE or LINCS. Successful simulations require turning off the bond constraint and decreasing the time step to 1 fs. Alternative protocols that require some scripting and changes in the topology file could be pursued in the future. For instance 2 fs constraints protocols similar to those used in SOMD or AMBER in this study could be implemented via the definition of a new atom type for alchemically perturbed hydrogen atoms.

4.3.5 SOMD

Fig. **17**x compares relative free energy of hydration $\Delta\Delta G$ according to the protocol with unperturbed H bond constraints, with relative $\Delta\Delta G$ obtained from two absolute free energy calculations. Tab. **6** summarizes all the computed relative free energy of hydration for the dataset in Fig. **12**. A very good agreement is observed between both methodologies (R^2=0.99 ± 0.01 and MAE = $(0.10 \pm 0.03)\,\mathrm{kcal\,mol^{-1}}$), highlighting internal consistency within SOMD.

To achieve this level of reproducibility within SOMD it was crucial to pay close attention to constraints. Specifically, bonds that involve unperturbed hydrogen atoms are constrained. Bonds involving hydrogen atoms that are perturbed to a heavy element are unconstrained. Additionally the atomic mass of any perturbed hydrogen atom is set to the mass of the heavy atom it is being perturbed to. Bonds involving hydrogen atoms that are perturbed to another hydrogen atom type are constrained. We stress that it is acceptable to artificially increase the atomic mass of hydrogen atoms because the calculated excess free energy changes do not depend on atomic masses.

This protocol suppresses high frequency vibrations in flexible bonds involving hydrogen atoms, thus enabling a time step of 2 fs, whilst giving essentially negligible errors due to the use of constraints for perturbed bonds. This is apparent from the comparison with the absolute hydration free energy calculations. Additionally, the protocol yields relative hydration free energy very similar (MAE = $0.09\,\mathrm{kcal\,mol^{-1}}$) to those computed from simulations where no constraints are applied for solutes and a timestep of 1 fs is used (See Fig. **19**).

By contrast, a protocol that constrains all bonds in a solute leads to significant differences with the absolute hydration free energies. For instance neopentane $\rightarrow$ methane (centrally mapped) gives a RAFE $\Delta\Delta G$=$(2.04 \pm 0.01)\,\mathrm{kcal\,mol^{-1}}$ whereas the absolute hydration free energy calculations give $\Delta\Delta G$=$(-0.19 \pm 0.06)\,\mathrm{kcal\,mol^{-1}}$ as shown in Tab. **18** and fig. **19**.

This discrepancy occurs because in the SOMD implementation, the energies of constrained bonds are not evaluated, but the calculation of the energies of the solute at perturbed λ values is carried out using the coordinates of the reference λ trajectory. This leads to a neglect of contributions of the bonded term (and associated coupled terms) to the free energy change. The effect is more pronounced for perturbations that feature a large change in equilibrium bond lengths, such as those where a hydrogen atom is perturbed to/from a heavy atom.

The reaction fields implemented in SOMD and GROMACS differ somewhat (atom-based shifted Barker Watts, [258] versus group-based switched Barker Watts), but nevertheless SOMD and GROMACS RF produce comparable results with a MAE of $0.18\,\mathrm{kcal\,mol^{-1}}$. Overall, the SOMD free energy estimations are in good agreement with the other MD packages, as the MAE suggests (see Tab. **7**). For the methane $\rightarrow$ neopentane transformations SOMD yields consistent results between central and terminal mappings, as shown in Tab. **17**. Reaction field and PME results are in good agreement. All SOMD RAFE simulations were carried out with simultaneous transformation of Lennard-Jones, charges, and bonded terms. This suggests that the failure of the GROMACS "unified protocol" in some instances may be due to differences in the softcore Coulomb implementations.

4.4 Discussion and Conclusions

This study addressed whether contemporary MD packages such as AMBER, CHARMM, GROMACS and SOMD are able to reproduce relative alchemical free energies of hydration for a set of neutral small organic molecules, given a pre–defined force field. We have found that establishing a simulation protocol that leads to consistent results across codes has been cumbersome due to technical difficulties encountered with every code. This was the case despite our best efforts to maintain fairly consistent protocols for settings which were expected to significantly impact results. For example, we used nominally the same form

of soft core potentials in most of the codes compared, and implementations of many other algorithms which should be the same or are thought to be equivalent. Still, we encountered numerous difficulties. Overall, the MD codes have a wide range of options and setup features which makes it difficult for the inexperienced user to decide on the most appropriate ones.

The free energies we have computed appear to be in reasonably good agreement with each other (see Tab. **4** and **6**). The average MAE between all codes 0.14 kcal/mol for absolute free energies and 0.17 kcal/mol for relative free energies. This can be interpreted as the current "limit of reproducibility" for the field. We have found viable protocols for each MD code to achieve this level of reproducibility. There is some doubt, however, over the AMBER results because the particular version of the software we tested cannot reproduce the correct end-point geometries. This is particularly evident in the case of the transformation from neopentane to methane with central mapping, where the relative free energies are clearly different from the other packages. We suspect these reflect issues in the AMBER package but have been unable to isolate it; we have reported the issue to the AMBER developers.

We were unable to define a *universal* protocol that could be recommended for use with all four codes. Unified protocols do not appear to work adequately with AMBER and GROMACS while SOMD and CHARMM had no problem in this regard. We cannot rule out that the problem may lie e.g. only with the vacuum leg of the thermodynamic cycle. In the case of AMBER the vacuum simulation has currently been done with the separately developed `sander` module. The problem may be a consequence of the different softcore functions (see Eq. S**??**x) used in these MD packages but further investigations are needed to resolve this issue.

The unperturbed H bond protocol is an interesting alternative which applies constraints to all non–transforming bonds and thus allowed us to increase the time step to 2 fs. The split protocol was found to work well for all codes. It appears to be the most effective approach for GROMACS as shown with the methanol to methane case because the unified

protocol produces a less smooth function [84]. A complete separation of lambdas may not be necessary though as a certain degree of overlap between vdW and Coulumb λ may be a viable solution [264] for equilibrium AFEs.

Comparison between codes is hampered by several factors. Firstly, the codes use different simulation algorithms — e.g. electrostatics are handled differently in vacuum i.e. infinite cutoff vs. reaction field. Temperature and pressure control, time step, integrators, etc. are other examples. But the data here suggest that, if there are any systematic errors introduced through these algorithms, then they are small. It is reassuring that AFEs for the systems tested here show only a small dependence on MD protocol decisions (provided a correct implementation).

Some of the differences between protocols used in this comparison could have been avoided, and it may be worth pursuing further harmonizing the protocols in follow up work. For example, the number of lambda values, length of simulations, and choice of Lennard-Jones cutoff were varied across packages in some cases. While previous studies have suggested results are relatively insensitive to these choices, it may be worth further exploring these issues in follow-up work to ensure results are robust with respect to these settings. This could also allow for a direct comparison of efficiency across codes.

To aid with follow-up studies, we make our input data and protocols available. We recommend using this dataset to test and benchmark future RAFE implementations to validate reproducibility against other simulation packages. Where possible, we recommend comparing results from both absolute and relative transformations to verify internal consistency. The relative transformation should be run in both forward and backward directions, even if the free energy estimator is agnostic to this decision, as other implementation details (e.g. parameters, atomic masses, use of bond constraints, or details of the number of dummy atoms and how they are handled) may lead to inconsistent results depending on the transformation direction.

More specifically, various issues with current code bases have been revealed through this work. We have found that constraints in connection with varying bond length can cause errors with GROMACS, just as masses (in many codes) must not be allowed to vary in RAFE simulations, both to avoid crashes and incorrect results from the software. CHARMM has problems handling constraints and the PSSP softcores, and the PERT module cannot make use of the force switch as is now standard for CHARMM force fields. Care must be taken when using the LRC long range correction keyword to avoid producing inconsistent results. AMBER's problem with end point geometries and unified protocols has been pointed out above.

Another question is the ease of use of the different software. For example, when a mutation entails both appearing and disappearing parts in split protocols there is the problem of intermediates having a non–integral total charge on the molecule. An alternative would be to totally discharge and then recharge the whole molecule which would have the advantage of eliminating one additional evaluation of the reciprocal sum in PME [263]. However this is not attractive as this could significantly increase the sampling needed to obtain converged free energy changes.

In general we found that split protocols perform well, but these can necessitate complex steps to set up the calculations. For instance, in GROMACS it is necessary to carry out two separate simulations per lambda because discharging and recharging groups cannot be selected separately. Lambda paths as implemented in GROMACS could also be beneficial for other codes as they make the setup of split protocols easier. The alternative we have used in codes lacking this feature is to mimic this protocol through careful construction of topologies via scripting.

It may seem remarkable that of the computed free energies, the absolute hydration free energies seem to be more reproducible across codes than relative free energies. Conventional wisdom is that relative free energy calculations are computationally less demanding than

absolute free energy calculations, which tend to lead to the opposite result [193, 119, 265]. However, absolute calculations are considerably simpler to implement and deploy correctly as they do not involve as many challenging technical issues such as atom mapping [265, 266], and the lambda protocols which must be employed have been optimized fairly well, since such calculations always involve either removal or insertion of atoms but never both simultaneously. This has made absolute calculations valuable as large-scale tests of free energy methods and force fields (e.g. [57, 253, 187] and others), and in the SAMPL series of blind challenges (e.g. [73, 59, 57]). Thus absolute calculations are already well automated, robust across codes (e.g. [53]), and well-performing protocols are available. Apparently similar is yet needed for relative free energy calculations.

The primary focus of this work was to achieve low statistical errors to establish if codes are able to reproduce free energies. We have not investigated the efficiency of the respective protocols as this would require further, complex investigations. As noted, the work reported here used different protocols and in some cases even different numbers of simulation steps for each code. Thus a direct comparison of efficiency is outside the scope of this work. However, it is worth briefly noting the number of steps employed in each study. For absolute calculations the most demanding and (perhaps not surprisingly) precise protocol is GROMACS (200 million aggregate time–steps per solute, average SEM 0.011 kcal/mol). The least demanding protocol is CHARMM (31.5 million time–steps per solute, average SEM 0.015 kcal/mol). SOMD's aggregate time–steps is comparable to CHARMM (34 million time–steps) but the free energies are less precise (average SEM 0.045 kcal/mol). For relative calculations, the least demanding protocol is SOMD (17 million time–steps), and this is also the least precise (average SEM 0.048 kcal/mol). The most demanding protocol (GROMACS 197.4 million time–steps, average SEM 0.020 kcal/mol) is less precise than the CHARMM protocol which used fewer time–steps (31.5 million time-steps, average SEM 0.015 kcal/mol). Further work should be pursued to understand what algorithmic details in the various implementations are

important for the efficiency of the free energy calculations. In particular it may be interesting to apply overlap measures to explore the relative efficiency of the different protocols. [267]

Beyond careful protocol validation, further automation of alchemical free energy studies will also decrease user errors, and thus increases reproducibility. Various attempts in this direction are currently underway for both absolute and relative setups [136, 222, 265, 246, 268, 269, 247]. To conclude, we hope this study will stimulate the field to improve the transferability of alchemical free energy calculation protocols across software. Reproducibility is crucial to enable robust use of alchemical free energy methods in molecular design.

4.5 Supporting Information

4.5.1 Softcore Functions

We describe here the softcore functions [10, 11] as implemented in the MD packages AMBER, CHARMM, Gromacs and SOMD. Both the van der Waals, V_{LJ} (Lennard–Jones potential) and the electrostatic interactions, V_{Coul} (Coulomb potential) as a function of the order parameter λ are given for the disappearing atoms only. For the appearing atoms replace λ with $1-\lambda$ and *vice versa*. Eq. (4.5) is the generalized form for all codes while the specific distance dependent functions are outlined in eq. (4.6) for SOMD, eq. (4.7) for AMBER, eq. (4.8) for Gromacs and eq. (4.9) for CHARMM.

$$V = V_{\mathrm{LJ}} + V_{\mathrm{Coul}} = 4\epsilon_{\mathrm{ij}}(1-\lambda)\left[\left(\frac{\sigma_{ij}}{r_{\mathrm{LJ}}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{\mathrm{LJ}}}\right)^{6}\right] + (1-\lambda)^{n}\frac{q_i q_j}{4\pi\varepsilon_0 r_{\mathrm{Coul}}} \tag{4.5}$$

For SOMD

$$\begin{aligned} r_{\mathrm{LJ}} &= (\alpha\sigma_{ij}\lambda + r_{ij}^2)^{\frac{1}{2}} \\ r_{\mathrm{Coul}} &= (\lambda + r_{ij}^2)^{\frac{1}{2}} \end{aligned} \tag{4.6}$$

For AMBER

$$\begin{aligned} r_{\mathrm{LJ}} &= (\alpha\sigma_{ij}^6\lambda + r_{ij}^6)^{\frac{1}{6}} \\ r_{\mathrm{Coul}} &= (\beta\lambda + r_{ij}^p)^{\frac{1}{p}} \\ n &= 1 \end{aligned} \tag{4.7}$$

For Gromacs

$$\begin{aligned} r_{\mathrm{LJ}} &= (\alpha\sigma_{ij}^w\lambda^p + r_{ij}^w)^{\frac{1}{w}} \\ p &= 1, 2; w = 6, 48; \\ r_{\mathrm{Coul}} &= r_{\mathrm{LJ}} \\ \alpha_{\mathrm{Coul}} &= 0, \alpha_{\mathrm{LJ}} \\ n &= 1 \end{aligned} \tag{4.8}$$

For CHARMM (PSSP), applied to all "reactant" and all "product" atoms

$$\begin{aligned} r_{\mathrm{LJ}} &= (\alpha\lambda + r_{ij}^2)^{\frac{1}{2}} \\ r_{\mathrm{Coul}} &= (\beta\lambda + r_{ij}^2)^{\frac{1}{2}} \\ n &= 1 \end{aligned} \tag{4.9}$$

r_{vdW}, r_{Coul} and r_{Coul} are distance dependent functions, ϵ_{ij} and σ_{ij} are the Lennard-Jones parameters, q_i and q_j are the charges and ε_0 is the vacuum permittivity, α and β are the

softcore tuning parameters determining the softness of the potential, and r_{ij} the distance between atoms i and j.

The exponent n is only used in the Coulomb softcore funtion of SOMD. Gromacs allows additional exponents for λ ($p = 1$ or 2) and w for the distance dependency with values of either 6 or 48. AMBER allows an exponent p (namelist option SCEEORDER) for the Coulomb softcore. The Coulomb softcore parameter α_{Coul} in Gromacs is the same as for the Lennard–Jones parameter α_{LJ} unless the Coulomb softcore function is requested not to be used and thus $\alpha_{\mathrm{Coul}} = 0$. The CHARMM softcore function (PSSP) is applied to *all* atoms in the perturbed group and not only to dummy atoms as in the other codes. The perturbed group comprises of all atoms that need to be transformed, i.e. any atom that differs in at least one force field parameter in the other end state. "Dummy" atom is used here as a shorthand notation for any atom that appears or disappears during the course of the transformation.

4.5.2 TI gradients

The free energy derivative versus lambda from thermodynamic integration simulations (TI gradients) for all four codes are summarized in Fig. **13** by example of the absolute transformation of methanol. The derivatives demonstrate how the different transformation functions especially the softcore potentials (see section 4.5.1) influence the TI gradients. As can be seen these are very different from each other implying that the lambda schedule must be appropriately planned for each simulation code and also for the particular transformation.

4.5.3 Split Protocols

When the AFE (alchemical free energy) simulation is separated (split) into van der Waals and Coulomb steps it must be ensured that charges of vanishing atoms are switched off

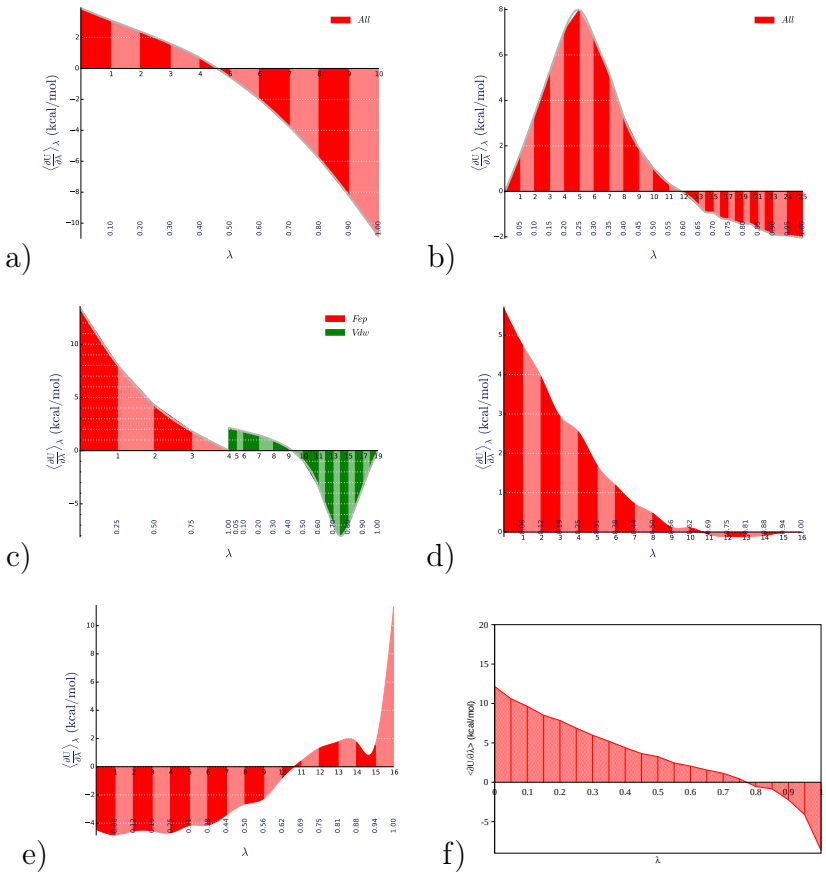

Figure 13: TI gradients obtained from various simulation codes for the absolute transformation of methanol. a) and b) AMBER charge and vdw transformations. c) Gromacs where green is from the vdw only transformation and red for the free energies from all other contributions, d) and e) SOMD charge and vdw transformations. f) CHARMM unified protocol.

before the vdW radius is scaled to zero. This is to avoid that other atoms e.g. from solvent come in close contact to a charged atom without the associated excluded volume from the van der Waals term as this could lead to large forces and thus instabilities in the integrator.

Fig. **14** depicts how force field parameters vary for a transformation carried out in the direction of disappearing atoms. The mutation is shown with the charge step first followed by the vdW step but each step can really be run independently. Please note that both charge and vdW step would be simulated at a range of individual λs. Typically the charge transfer is done with linear scaling while the vdW mutation is done with softcores (see above). The

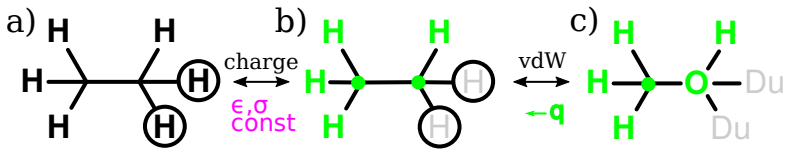

Figure 14: The mutation of ethane into methanol and explicit topologies for three states. a) The two circles denote atoms that have both vdW and Coulomb terms switched on with parameters for the respective hydrogen atom type. b) The two hydrogen atoms have their charges switched to zero (gray symbols in black circle). All other charges are the ones from the methanol end state c (green) to ensures charge neutrality at each step. VdW parameters are constant in the charge transfer step (see annotations in magenta). c) vdW and Coulomb parameters as for methanol while dummy atoms (gray Du) have those parameters all set to zero.

transformation is fully symmetrical that is the parameters must be switched on in opposite order if atoms are to be "created". The intermediate state b has the vdW parameters from state a but the charges from state c.

Fig. **14** shows how topology files may be created in cases where the MD software does not allow independent λs for electrostatic and vdW mutations. With Gromacs, for instance, the transformation only requires a single topology file with both A and B states (in single topology fashion, see main text) and a single simulation control file with separate λ vectors for charge and vdW transformations. Any intermediate state from Fig. **14** is thus created "on–the-fly" i.e. implicitly during the simulation run. With AMBER (up to version 16 as of this writing), however, three explicit topology files (with sander, two with pmemd) and two control files would need to be created. The state b in Fig. **14** would be created from state a with the charges from state c. The bonded terms can be combined with either mutation step or run separately. For AMBER the easiest way is to combine vdW with bonded terms because charges are independent of atom types.

Fig. **15** illustrates explicit topologies for transformations with both appearing and disappearing atoms in one simulation. The principle is essentially the same as in Fig. **14**: charges of dummy atoms must be switched off before vdW parameters are set to zero to avoid interactions of "naked" charge sites with other atoms possibly leading to very close contact,

large energies and forces, and thus to unstable simulations and/or noisy statistics. However, charge neutrality at every λ step is not supported in most MD codes i.e. the total system charge varies with λ unless the charges of *all* atoms are switched off. Possible strategies would be to explicitly create topology files for each intermediate λ state and distribute the diminished charges from the dummy atoms over to the non–dummy atoms. MD software like CHARMM allows to do this through internal scripting although this would be just as extensive as external scripting the aforementioned strategy.

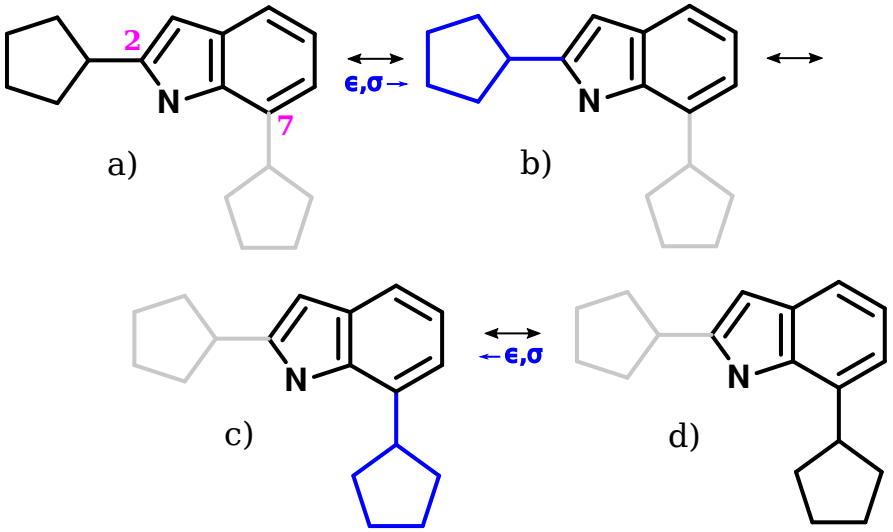

Figure 15: Explicit topologies involved in a mutation with both appearing and disappearing atoms by example of the cyclopentanyl transfer from the 2 position in indole to the 7 position. Blue lines denote atoms which have their charges switched off but with vdW parameters from the left (state b) or right (state c). Gray lines are dummy atoms with Coulomb and vdW parameters all zero. Note, the hydrogen bound to the 2 (state d) and 7 positions (state a) can be directly mutated from the respective carbon atom type without ring breaking [237].

With the MD packages tested in this study the number of input files are as follows. With Gromacs this can be done with only two topology and two control files where one charge transfer can be combined with a vdW on/off step. Gromacs' λ vectors only apply to the perturbed group as a whole and so it is not possible to define a λ vector for only a subset. AMBER requires two such files with sander and three topology/two control files with pmemd for the three steps charge off, vdw on/off and charge on. This is possible because with AMBER a subset of the perturbed group can be chosen to have zero charges (namelist

Table 13: Changes in volumes (in cubic angstrom) for selected perturbations across packages

transformation		AMBER	CHARMM	GROMACS	SOMD
methane	ethane	38±6	24±3	31±3	61±25
ethane	methane	-29±7	-30±4	-28±3	-50±30
ethane	methanol	-33±9	-36±2	-36±5	-66±36
methanol	ethane	38±8	37±3	36±5	46±31
methane	methanol	-5±11	-9±2	-3±5	11±20
methanol	methane	8±10	1±2	4±3	27±16
methane	toluene	93±11	89±3	110±2	120±56
toluene	methane	-106±8	-89±8	-113±3	-145±23
methane	2-methylindole	164±11	64±3	142±3	140±22
2-methylindole	methane	-138±8	-120±7	-139±5	-166±45
methane	neopentane	99±9	90±2	115±4	117±50
neopentane	methane	-105±11	-100±2	-114±4	-68±43

option CRGMASK; but AMBER does not have λ vectors). CHARMM has scripting facilities that let the user manipulate force field parameters of any arbitrary subset of the system such that intermediate states can be defined "on–the-fly" with only one control script and one topology file. The tool FESetup [247] automates most of these setup steps for all these MD packages.

4.5.4 Detailed Results

AMBER

Tab. **14** compares the split protocol with the unified protocol. The split protocol transforms Coulomb force field parameters separately from the Lennard–Jones and all bonded parameters. The unified protocol transforms all force field parameters simultaneously and thus invokes both Coulomb and vdW softcore functions.

The split protocol produces consistent results in both solution and in vacuum. The values are in line with the free energies obtained with the other MD packages (see main text). Each

Table 14: Comparison between split and unified protocol in **AMBER**. The data for the unified protocol highlights inconsistencies in the code in red. ΔG_{sol} has been computed with pmemd. ΔG_{vac} has been computed with sander.

		split protocol			unified protocol		
transformation		ΔG_{sol}	ΔG_{vac}	$\Delta\Delta G$	ΔG_{sol}	ΔG_{vac}	$\Delta\Delta G$
ethane	methane	1.79	1.77	0.02	2.78	2.85	-0.07
methane	ethane	-1.80	-1.80	0.01	-2.87	-2.86	-0.01
methanol	methane	2.74	-3.44	6.19	7.10	0.87	6.23
methane	methanol	-2.75	3.45	-6.19	-7.18	-0.86	-6.32
ethane	methanol	-2.84	3.36	-6.20	-2.25	3.99	-6.25
methanol	ethane	2.83	-3.36	6.19	2.20	-3.99	6.20
toluene	methane	9.22	5.98	3.24	6.09	0.45	5.64
methane	toluene	-9.29	-5.86	-3.42	-6.15	-0.54	-5.61
neopentane[a]	methane	70.16	69.85	0.31	65.76	58.50	7.28
methane[a]	neopentane	-70.17	-69.92	-0.25	-65.82	-58.78	-7.04
neopentane[b]	methane	11.41	11.54	-0.13	4.42	3.48	0.94
methane[b]	neopentane	-11.43	-11.55	0.12	-4.45	-3.49	-0.96
2-methylfuran	methane	14.62	11.53	3.09	2.20	-0.94	3.15
methane	2-methylfuran	-14.60	-11.50	-3.10	-2.22	-0.06	-2.15
2-methylindole	methane	24.25	15.47	8.78	7.11	-4.02	11.13
methane	2-methylindole	-24.31	-15.17	-9.14	-7.13	1.86	-8.97

[a]central mapping.
[b]terminal mapping.

ΔG is the sum of the charge and vdW plus bonded contributions. The unified protocol on the other hand displays various problems. While the smaller systems with only a few dummy atoms show ΔG and $\Delta\Delta G$ consistent with the split protocol, the larger transformations show, in part, striking differences and even inconsistencies in forward and backward vacuum simulations. It is not clear, however, if the inconsistencies can be attributed to the vacuum transformations only.

Fig. **16** shows a problem with end point geometries. This is demonstrated with the average distance between the carbon atom and the four attached hydrogens atoms in the neopentane to methane case. The methane carbon atom is mapped here to the central atom of neopentane. The distances are recorded for the vdW plus bonded transformation i.e. the charges correspond to the methane end state.

The geometrical variation along λ for the data in the main text is shown in the black and red graphs. The initial distance is slightly smaller than what is expected from a C–H distance for the particular atom types at $\lambda = 0$. The final distance is about 1.23 Å which is in contrast to the 1.09 Å of the c3–hc bond of the GAFF force field. The crosses in violet mark the geometries of the "pure" (non–perturbed) end points and are connected with a straight line. The other crosses denote test cases which successively replace the methyls on neopentane with hydrogens. The C–H distance decreases in correlation with the number of the methyl groups i.e. tert-butane, propane, ethane. This seems to suggest that a "crowding" of dummy atoms around a central atom compounds the problem of a too long C–H distance. Neither of these three test cases, however, displays the expected end point distance.

To further test this hypothesis methyl and ethyl groups are added to all terminal methyl groups of neopentane, see cyan and green lines in Fig. **16**. In both cases the end point distance is about 1.12 Å with a slightly lower value for the ethyl substitution but which are still too high.

Table 15: **AMBER**: Free energy components for 2–methylindole computed from implicit dummy RAFE simulations. The data are averages over three runs.

transformation		$\Delta G^{\mathrm{vdW}}_{\mathrm{sol}}$	SEM	$\Delta G^{\mathrm{vdW}}_{\mathrm{vac}}$	SEM	$\Delta G^{\mathrm{elec}}_{\mathrm{sol}}$	SEM	$\Delta G^{\mathrm{elec}}_{\mathrm{vac}}$	SEM
2-methylindole	methane	4.83	0.02	3.48	0.01	19.41	0.01	11.99	0.01
methane	2-methyindole	-4.90	0.02	-3.18	0.01	-19.41	0.01	-11.90	0.01

As Gromacs conveniently allows us to use a separate λ for bonded terms we tested this on the neopentane case. After the charges were transformed to the methane end state (dummy atoms have zero charges), the bonded terms were mutated from neopentane to methane while the vdW parameters were kept constant at the neopentane initial state. The observed end distance was about 1.23 Å which is to be expected given that the symmetrically arranged methyl groups will repel each other and thus not allow to reach the final distance. Only after the final vdW (only) mutation had been carried out, were the final distances of 1.09 Å reached.

Tab. **15** summarizes the free energy components for the 2–methylindole to methane case for both forward and backward transformations. The electrostatic contributions display a very small SEM and the averages from both directions agree with each other up to the second digit after the comma. The van der Waals contributions show a higher SEM and the averages from the solution simulations agree well with each other. However, the van der Waals contributions from the vacuum transformation are apart by 0.3 kcal mol^{-1} (highlighted in red).

In sum, this indicates a problem of the RAFE code in AMBER. Whether that is a bug or a conceptual issue with the particular implementation can not be explained at the moment.

Tab. **16** summarizes the free energies obtained from forward and backward simulations of the cyclopentanyl transfer from position 2 to position 7 on indole and *vice versa*. Results from three different protocols are shown: 1) implicit dummy atoms and partial re/discharge of the 5–ring only; 2) implicit dummy atoms and full re/discharge of all atoms; 3) explicit dummy

Table 16: **AMBER**: Free energies of hydration for the 2–cyclopentanylindole to 7–cyclopentanylindole case with three different protocols. The data are averages over three runs.

transformation		ΔG_{sol}	SEM	ΔG_{vac}	SEM	$\Delta\Delta G_{\mathrm{hydr}}$	SEM
implicit, partial							
2–cyclopentanylindole	7–cyclopentanylindole	4.19	0.03	3.83	0.01	0.36	0.03
7–cyclopentanylindole	2–cyclopentanylindole	-4.30	0.04	-3.96	0.01	-0.33	0.04
implicit, full							
2–cyclopentanylindole	7–cyclopentanylindole	4.28	0.06	3.92	0.02	0.36	0.07
7–cyclopentanylindole	2–cyclopentanylindole	-4.41	0.03	-4.10	0.01	-0.32	0.03
explicit, partial							
2–cyclopentanylindole	7–cyclopentanylindole	4.14	0.04	3.51	0.04	0.63	0.06
7–cyclopentanylindole	2–cyclopentanylindole	-4.25	0.03	-3.76	0.01	-0.50	0.03

atoms and partial re/discharge. The free energies from the implicit dummy simulations agree very well with each other while the explicit dummy atom results are about $0.2\,\mathrm{kcal\,mol^{-1}}$ higher and forward and backward simulations have a larger hystersis of $0.1\,\mathrm{kcal\,mol^{-1}}$.

SOMD

As discussed in the main tex, the presend SOMD calculations used a shifted atom-based Barker-Watts reaction field cutoff to handle electrostatic interactions in solutions and a Coulomb laws without cutoffs to handle electrostatic interactions in vacuum. This leads to an inconsistent description of the intramolecular electrostatic interactions of the solute in the solvated and vacuum phases. To maintain a consistent description of intramolecular energetics across vacuum and water legs, a free energy correction term ΔG_{c} was evaluated

as detailed in Ref. [259]. The ΔG_{c} term was obtained via post-processing of the end state trajectories of each water phase simulation, using the Zwanzig relationship [8]:

$$\Delta G_{\mathrm{c}} = -\beta^{-1} \ln\langle \exp\left[-\beta(U_{\mathrm{ic,nc}}(\mathbf{r}) - U_{\mathrm{ic,sim}}(\mathbf{r}))\right]\rangle_{\mathrm{sim}} \tag{4.10}$$

where $U_{\mathrm{ic,nc}}(\mathbf{r})$ is the solute intramolecular electrostatic-no cutoff potential that depends on the coordinates $\mathbf{r}$ of the solute and is given by Coulomb's law computed without cutoffs. $U_{\mathrm{ic,sim}}(\mathbf{r})$ is the intramolecular electrostatic potential term as computed in the simulation with the shifted atom-based Barker-Watts Reaction Field cutoff.

Fig. **17** compare relative free energy of hydration $\Delta\Delta G$ with relative $\Delta\Delta G$ estimations from absolute free energy calculations for all the transformation of the dataset.

Table 17: **SOMD**: Final relative free energy of hydration $\Delta\Delta G$ estimations and standard error SEM with Sire/SOMD unperturbed hydrogen bonds protocol, RAFE, compared with relative free energy of hydration computed from absolute free energy simulations, RAFE-absolute. Signs in backward transformation are reverted for better comparison.

		RAFE		RAFE-absolute	
transformation		$\Delta\Delta G$	SEM	$\Delta\Delta G$	SEM
ethane	methane	-0.01	0.05	0.04	0.02
methane	ethane	-0.04	0.02	0.04	0.02
methanol	methane	5.99	0.05	6.21	0.05
methane	methanol	5.96	0.04	6.21	0.05
ethane	methanol	-6.09	0.03	-6.26	0.05
methanol	ethane	-6.09	0.02	-6.26	0.05
toluene	methane	2.89	0.03	3.06	0.03
methane	toluene	3.06	0.02	3.06	0.03
neopentane[a]	methane	-0.20	0.054	-0.19	0.06
methane[a]	neopentane	-0.13	0.055	-0.190	0.060
neopentane[b]	methane	-0.11	0.01	-0.19	0.06
methane[b]	neopentane	-0.10	0.06	-0.19	0.06
2-methylfuran	methane	2.92	0.05	2.90	0.03
methane	2-methyfuran	2.83	0.03	2.90	0.03
2-methylindole	methane	8.64	0.06	8.57	0.03
methane	2-methylindole	8.67	0.08	8.57	0.03
2–cyclopentanylindole	7–cyclopentanylindole	0.11	0.077	0.08	0.14
7–cyclopentanylindole	2–cyclopentanylindole	0.01	0.081	0.08	0.14

[a]central mapping.
[b]terminal mapping.

Tab. **17** shows relative free energy of hydration $\Delta\Delta G$ compared to $\Delta\Delta G$ values extracted from absolute free energy calculations, depicted in Fig. **17**.

Initially, SOMD RAFE protocols used all bonds constraint algorithm. In this way all the solute bonds are constrained, which results in a systematic offset for each RAFE predictions, compared to the RAFE from absolute free energy calculations. Fig. **18** shows the discrepancy between RAFE computed with all bond constraints and RAFE from absolute free energy calculations.

Table 18: **SOMD**: Relative free energy of hydration $\Delta\Delta G$ computed with all bond constraints, *All bonds*, no constraints, *None*, and unperturbed hydrogen bond constraint, *unpert H bonds*

		All bonds		None		unpert H bonds	
transformation		$\Delta\Delta G$	SEM	$\Delta\Delta G$	SEM	$\Delta\Delta G$	SEM
ethane	methane	-0.48	0.01	-0.18	0.04	-0.01	0.05
methane	ethane	-0.49	0.01	-0.01	0.02	-0.04	0.02
methanol	methane	6.06	0.01	6.49	0.01	5.99	0.05
methane	methanol	6.08	0.01	6.15	0.01	5.96	0.04
ethane	methanol	-6.22	0.01	-6.14	0.03	-6.09	0.03
methanol	ethane	-6.23	0.01	-6.09	0.01	-6.09	0.02
toluene	methane	3.73	0.27	3.09	0.06	2.89	0.09
methane	toluene	3.79	0.03	3.07	0.06	3.06	0.02
neopentane[a]	methane	-2.09	0.01	-0.14	0.14	-0.20	0.05
methane[a]	neopentane	-2.04	0.01	-0.018	0.06	-0.13	0.05
neopentane[b]	methane	-0.48	0.01	-0.14	0.06	-0.11	0.01
methane[b]	neopentane	-0.59	0.02	-0.14	0.060	-0.10	0.06
2-methylfuran	methane	3.38	0.02	2.81	0.03	2.92	0.05
methane	2-methyfuran	3.40	0.03	2.89	0.06	2.83	0.03
2-methylindole	methane	9.29	0.06	8.72	0.05	8.63	0.06
methane	2-methylindole	9.10	0.04	8.61	0.04	8.67	0.08

[a]central mapping.
[b]terminal mapping.

Finally, Fig. **19** and Tab. **18** compare relative free energy of hydration $\Delta\Delta G$ estimated with RAFE using all bonds constraints, no constraints and unperturbed hydrogen bond constraints.

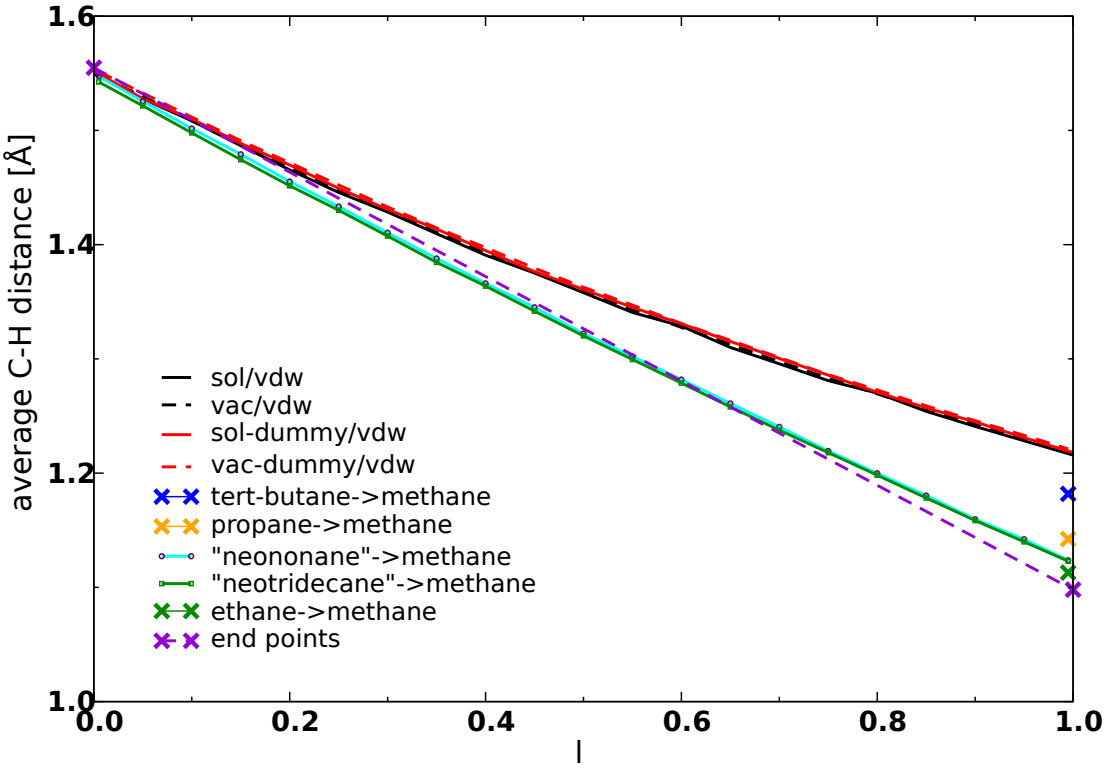

Figure 16: **AMBER** : The average C–H distance shown as a function of λ for the neopentane to methane and related cases. The black and red lines display how the distance changes in solution and the vacuum phase, and with and without explicit dummy atoms. The other test systems are designed to reduce the number of dummy atoms that surround the central carbon atom to show whether "crowding" is the cause of the issue. The crosses denote end points only, in particular the violet crosses represent the non–perturbed end point distances. For details see the text.

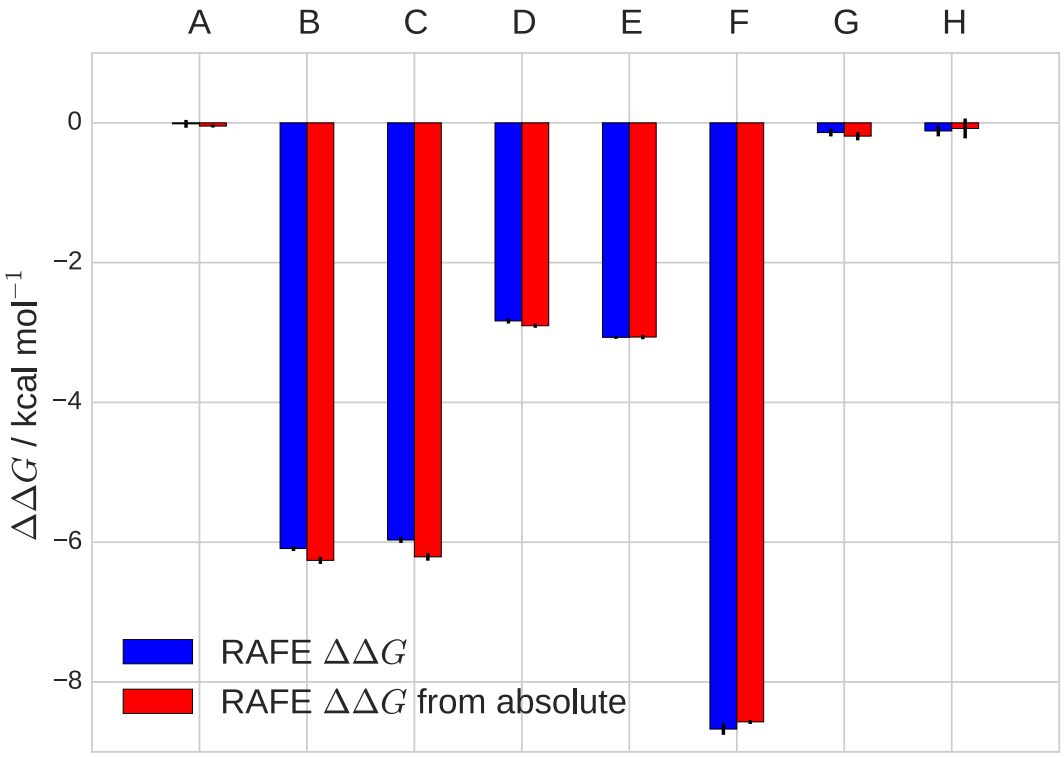

Figure 17: **SOMD**: Relative free energy of hydration $\Delta\Delta G$, computed with RAFE calculations, compared with $\Delta\Delta G$ derived from absolute free energy calculations for A: methane to ethane, B: ethane to methanol, C: methane to methanol, D: methane to 2-methylfuran, E: methane to toluene, F: methane to 2-methylindole, G: methane to neopentane, H: 2-cyclopentanylindole to 7-cyclopentanylindole.

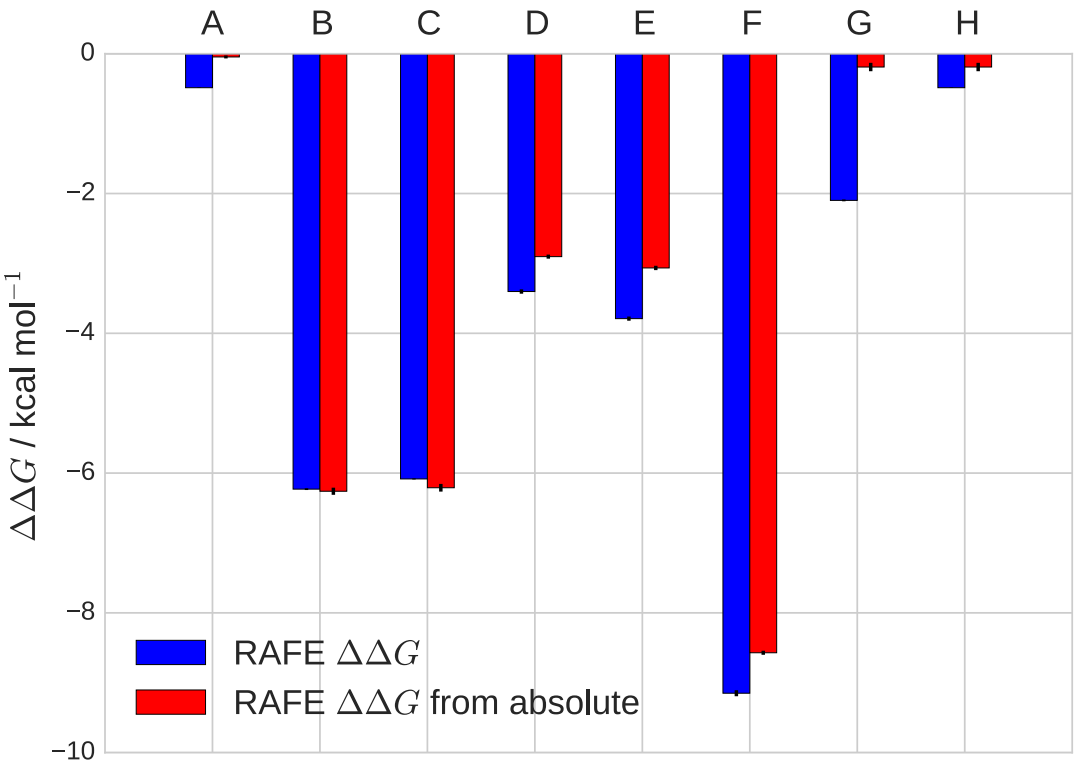

Figure 18: **SOMD**: Comparison between RAFE of hydration computed with all bond constraints and RAFE computed from absolute free energy calculations

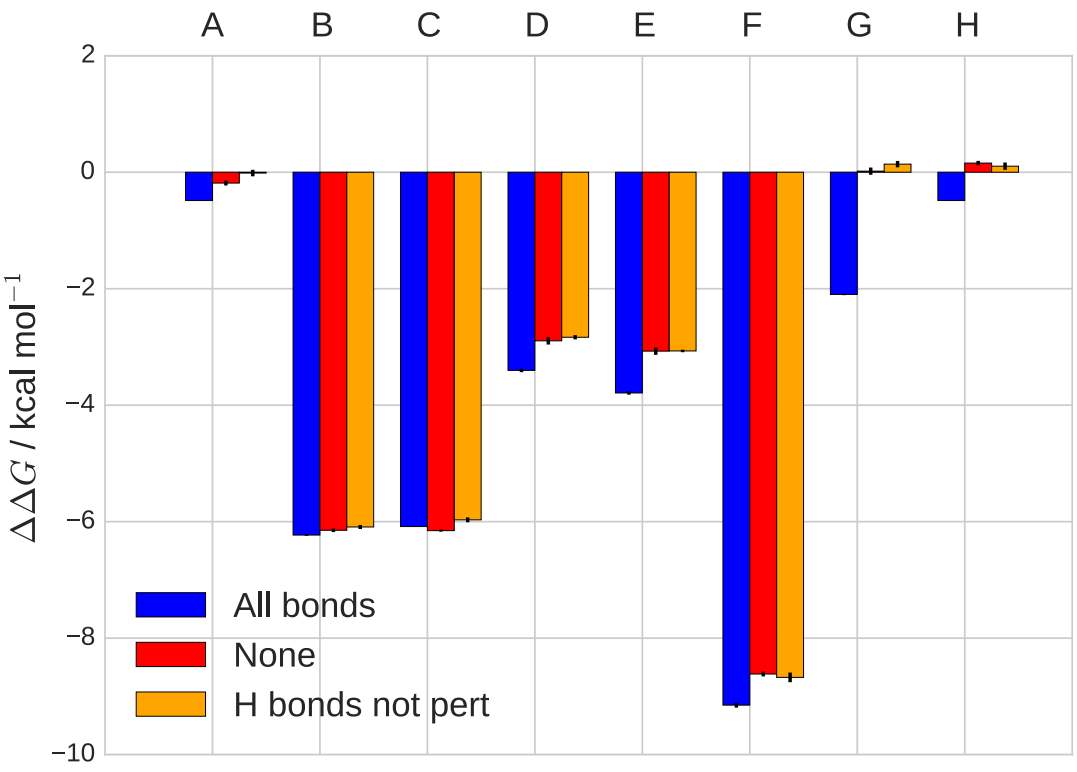

Figure 19: **SOMD**: Comparison between RAFE of hydration computed with all bond constraints, no constraints and unperturbed hydrogen bonds constraint

Table 19: **GROMACS**: $\Delta\Delta G_{hydr}$ results in different scenarios with or without Coulomb softcore potentials, in kcal $\cdot$ mol^{-1}.

		without Coulomb softcore		with Coulomb softcore		absolute	
		RF	PME	RF	PME	RF	PME
Transformations		$\Delta\Delta G$	$\Delta\Delta G$	$\Delta\Delta G$	$\Delta\Delta G$	$\Delta\Delta G$	$\Delta\Delta G$
ethane	methane	-0.02 ± 0.01	-0.03 ± 0.02	-0.03 ± 0.04	-0.02 ± 0.04	-0.06 ± 0.01	-0.04 ± 0.01
methane	ethane	-0.01 ± 0.02	-0.02 ± 0.01	-0.01 ± 0.04	-0.02 ± 0.04		
methanol	methane	6.16 ± 0.01	6.20 ± 0.01	7.32 ± 0.03	7.42 ± 0.04	5.77 ± 0.01	5.95 ± 0.01
methane	methanol	6.17 ± 0.01	6.20 ± 0.01	7.14 ± 0.03	7.21 ± 0.03		
ethane	methanol	-6.12 ± 0.01	-6.19 ± 0.01	-6.15 ± 0.02	-6.21 ± 0.02	-5.83 ± 0.01	-5.98 ± 0.01
methanol	ethane	-6.12 ± 0.01	-6.19 ± 0.01	-6.15 ± 0.02	-6.21 ± 0.02		
toluene	methane	3.22 ± 0.01	3.211 ± 0.010	3.22 ± 0.04	3.21 ± 0.04	2.97 ± 0.01	3.16 ± 0.01
methane	toluene	3.25 ± 0.01	3.20 ± 0.01	3.27 ± 0.04	3.22 ± 0.04		
neopentane[a]	methane	-0.10 ± 0.01	-0.15 ± 0.02	-0.13 ± 0.08	-0.13 ± 0.08	-0.18 ± 0.01	-0.14 ± 0.01
methane[a]	neopentane	-0.11 ± 0.02	-0.16 ± 0.05	-0.12 ± 0.08	-0.15 ± 0.08		
neopentane[b]	methane	-0.12 ± 0.01	-0.13 ± 0.01	-0.10 ± 0.04	-0.13 ± 0.04		
methane[b]	neopentane2	-0.10 ± 0.03	-0.18 ± 0.03	-0.08 ± 0.06	0.15 ± 0.06		
2-methylfuran	methane	2.97 ± 0.01	2.93 ± 0.05	3.07 ± 0.03	3.02 ± 0.04	2.87 ± 0.01	2.95 ± 0.01
methane	2-methylfuran	3.00 ± 0.01	2.96 ± 0.01	3.08 ± 0.03	3.02 ± 0.04		
2-methylindole	methane	8.71 ± 0.02	8.73 ± 0.03	8.79 ± 0.04	8.82 ± 0.05	8.44 ± 0.02	8.79 ± 0.02
methane	2-methylindole	8.73 ± 0.03	8.74 ± 0.01	8.79 ± 0.05	8.81 ± 0.06		
2-cyclopentanylindole	7-cyclopentanylindole	-0.07 ± 0.02	-0.03 ± 0.03	-0.12 ± 0.03	-0.14 ± 0.05	-0.02 ± 0.05	0.02 ± 0.02
7-cyclopentanylindole	2-cyclopentanylindole	-0.12 ± 0.06	-0.20 ± 0.04	1.20 ± 0.20[c]	1.50 ± 0.10[c]		

[a]central mapping
[b]terminal mapping
[c] inverted sign

GROMACS

Tab. **19** compares RAFE results subject to the use of Coulomb softcore potentials. In principle, the use of softcore functions is redundant in the split protocol because charges are changed while van der Waals parameters are fully tuned to the transformation's final state parameters. SEM values tend to be larger when they are used.

The effect of the Coulomb softcore potential can be seen in Fig. **20**.

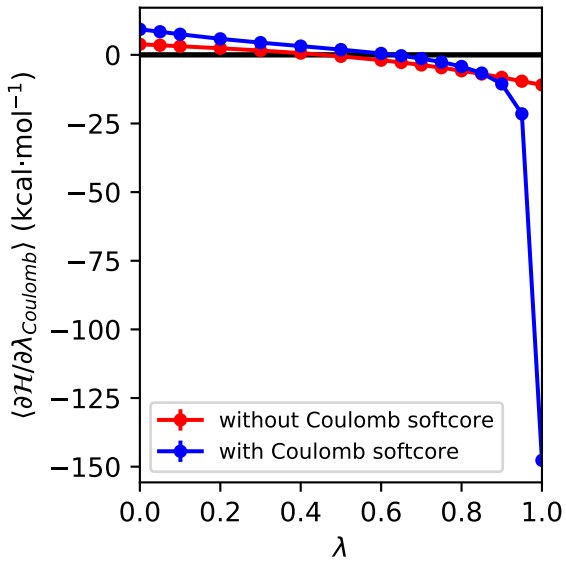

Figure 20: **GROMACS**: $\langle \partial \mathscr{H} / \partial \lambda \rangle$ plot for the change in electrostatic terms in methane $\rightarrow$ methanol.

CHARMM

Fig. **21** shows the effect of turning long range corrections (LRC) on/off as function of the cut–off values on the absolute solvation free energy of methane. The switching function starts 2 Å before the cut-off value. The figure shows that cutoff values greater than 10 Å give consistent results, if the LRC has been disabled. The bottom graph shows the relative solvation free energy of 2–methylindole to methane from direct RAFE and from AAFEs of methane and 2–methylindole with LRC on/off. This shows that the protocols will be inconsistent by ca. 0.8 $kcal.mol^{-1}$ if the LRC is not disabled. By post-processing LRC, RAFE result was more comparable to the results from other MD packages.

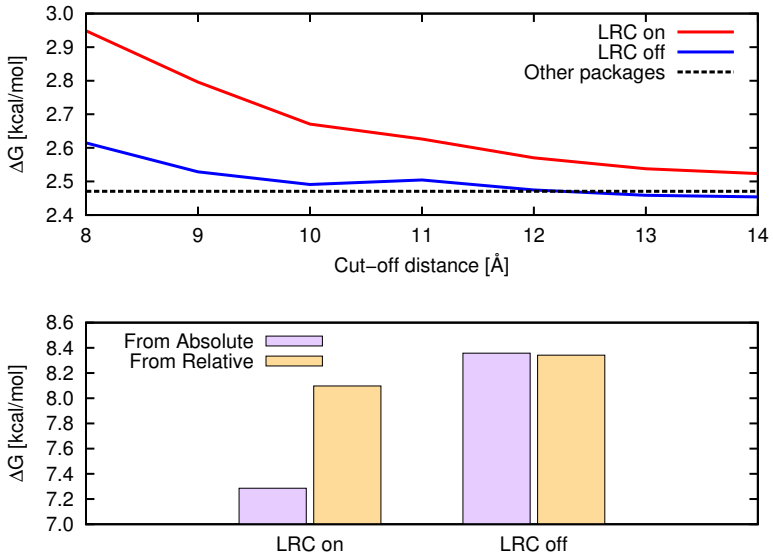

Figure 21: **CHARMM**: (top) Absolute solvation free energy of methane as a function of different cut–off values and presence or absence of a Long Range Correction term. (bottom) Relative solvation free energy between 2–methylindole and methane from relative or absolute alchemical free energy protocols and with/out a LRC term.

Chapter 5

Challenges of the Use of Atomistic Simulations to Predict Solubilities of Drug-like Molecules

Abstract

Background. Solubility is a physical property of extreme importance to the Pharmaceutical industry whose prediction for potential drugs has so far been a hard task. We attempted to predict the solubility of acetylsalicylic acid (ASA) by estimating absolute chemical potentials of its most stable polymorph and of solutions with different concentrations of the drug molecule.

Methods. Chemical potentials were estimated from all-atom molecular dynamics simulations. We used the Einstein Molecule Method to predict the absolute chemical potential of the solid and solvation free energy calculations to predict the excess chemical potentials of the liquid phase systems.

Results. Reliable estimations of the chemical potentials for the solid and for a single ASA molecule using the Einstein Molecule Method required an extremely large number

of intermediate states for the free energy calculations, meaning that the calculations were extremely demanding computationally. Despite the computational cost, however, the computed value did not agree well with experiment, potentially due to limitations with the underlying energy model. Perhaps better values could be obtained with a better energy model; however, it seems likely computational cost may remain a limiting factor for use of this particular approach to solubility estimation.

Conclusions. Solubility prediction of drug-like solids still is a challenge on the computational side, and it appears that both the underlying energy model and the computational approach applied may need improvement before the approach is suitable for routine use.

5.1 Introduction

Solubility is a critical property for pharmaceutical drug discovery, and problems with solubility can frustrate drug discovery efforts and block treatments. The bioavailability of a drug depends on the solubility difference between different crystal structures (polymorphs), dose, drug permeability, and formulation [270], so solubility plays a key role. Solubility problems can be unexpected and pose crucial obstacles that even threaten the administration of care. For example, a well-documented case occurred in the late 90's when ritonavir, an HIV-protease inhibitor marketed as Norvir, failed dissolution requirements [271]. Since ritonavir is not bioavailable in its solid form, it was administrated in capsules containing solutions designed not to be saturated with respect to the originally known molecular crystal (form I) [271]. The newly identified polymorph, form II, was unusually stable and unusually hard to crystallize; the preparation protocol of Norvir was inadequate to make capsules from the new polymorph, which severely threatened the supply of the drug in the market and endangered the lives of many HIV+ patients [271]. Considerable effort has already been devoted to the methods to predict crystal polymorphs [272, 273, 274, 275, 276, 277, 278], but much less attention has been given to methods to predict solubilities, with or without likely polymorphs as input.

Due to the importance of aqueous solubilities in different industrial processes and environmental applications, a scientific challenge consisting of the prediction of 32 solubilities given a database of 100 reliable measurements [45, 279] was created with the goal of comparing the outcomes of different solubility prediction techniques. Participants employed methods such as artificial neural networks [280], quantitative structure-property relationship (QSPR) [281], and deep learning [282] to predict the aqueous solubilities of drug-like molecules. All of the employed methods were empirical and trained on existing measurements. The limitation of these methods, however, is the dependence of a training set of data that limits

their applicability to compounds similar to those in the training of the set and impairs its transferability.

Some newer methods attempt to predict solubilities based on a physical description of the interactions in solution and in the solid state, yielding results that are in principle rigorous given an accurate energy model and an adequate method. In these approaches, molecular systems are described by force fields, i.e, potential energy functions that contain parameters describing bonds, atoms, electrostatic and non-electrostatic interactions. Molecular dynamics (MD) or Monte Carlo (MC) simulations are commonly used to sample different configurations of the system described by an energy model called a force field. The simulations then allow the estimation of physical properties such as internal energy, free energy, and enthalpy under different conditions. The quality of the results of such methods depends on how well the force field describes the system under study and how good the sampling method is. Thus, some researchers have recently estimated aqueous solubilities using simulations of thermodynamic cycles encompassing the crystal, the ideal gas, and an infinitely dilute solution of a given molecule [103, 283]. When the structure of the solid is unknown, some studies have substituted simulations of solid melts in place of a structure of the solid [106, 107, 108, 109].

While these physical methods for predicting solubilities have received some attention in the literature, most are still in their infancy with only a handful studies applying them and it is not yet clear how broadly applicable they will be [106, 107, 108, 109], and others have only been suggested or demonstrated in proof-of-principle tests [42, 44, 283, 284]. Our view is that the time is ripe for physical methods to predict solubility, especially given the routine nature of solvation free energy calculations at present [53, 285, 253, 86, 187, 286] which comprise essentially half of the solubility problem (see the Theory section). Polymorph and crystal structure prediction successes also mean that we may often have a suitable crystal structure of the compound as input [272, 287, 288, 289, 290, 273, 291, 274, 277, 278, 292], so what

remains is to predict the solubility given a crystal structure and simulations of the relevant phases.

Here, we focus on adapting and testing an existing approach for solubility prediction in the hope that it will prove to be a generally applicable method for solubility prediction that can be applied routinely. This method uses all-atom molecular dynamics simulations to estimate absolute chemical potentials and predict aqueous solubilities of molecular solids, given the crystal structure (or an estimate thereof) as input.

5.2 Theory

5.2.1 The solubility of a molecular solid is related to the chemical potentials of each phase

Solubility is defined as the maximum concentration of solute that can be dissolved in a selected bulk solvent. Chemical potentials (μ) of the solid-state solute and the solution are by definition equal at the solubility point, when the solution is in equilibrium with the solid.

$$\mu_{solute}^{solid} = \mu_{solute}^{solution} \tag{5.1}$$

Solid particles precipitate in concentrations higher than the solubility point because the solid phase becomes more stable in these conditions. In principle, we can predict at which concentration a molecule precipitates in solution if we calculate the chemical potentials of the components:

$$\mu_i = \left(\frac{\partial A}{\partial N_i}\right)_{V,T,N_{j,\,j\neq i}} = \left(\frac{\partial G}{\partial N_i}\right)_{P,T,N_{j,\,j\neq i}} \tag{5.2}$$

where μ_i is the chemical potential of component i; A is the Helmholtz free energy; G is the Gibbs free energy; $N_{j,\,j\neq i}$ is the number of molecules of each component in the mixture; V is the volume of the system; T its temperature; and P its pressure. Calculations from systems under constant V and T yield A; G is obtained from simulations under constant P and T conditions. In order to estimate the chemical potential of one component in solution and in its molecular solid, however, we need to know the absolute free energy of the system in these states. We calculated absolute free energies using alchemical free energy calculations.

5.2.2 Alchemical free energy calculations can be used to calculate absolute free energies

The absolute free energy of a system can be determined if we know its partition function (Q), a function that connects microscopic properties of the system with macroscopic thermodynamic quantities. Unfortunately, it is very hard to calculate the absolute free energy of real systems because we don't know their partition functions. Free energy calculations allow us to bypass this problem, but require at least two states: a reference state whose free energy can be analytically or numerically found, and a final state of interest [3, 293]. We chose to calculate the free energy difference using alchemical free energy calculations, a method in which we simulate a series of nonphysical intermediates between the end states [84].

Each intermediate state in the alchemical path is described by a Hamiltonian $\mathcal{H}(\mathbf{q}, \mathbf{p}; \lambda)$, i.e, the energy of the state as a function of atomic positions ($\mathbf{q}$), momenta ($\mathbf{p}$) and a coupling parameter (λ):

$$\mathcal{H}(\mathbf{q}, \mathbf{p}; \lambda) = f(\lambda)\mathcal{H}_{initial}(\mathbf{q}, \mathbf{p}; \lambda) + g(\lambda)\mathcal{H}_{final}(\mathbf{q}, \mathbf{p}; \lambda) \tag{5.3}$$

where $\mathcal{H}_{initial}$ and $\mathcal{H}_{final}$ respectively are the Hamiltonians of the initial and the final state; and $f(\lambda)$ and $g(\lambda)$ are functions used to mix the Hamiltonians and are usually set such that $\mathcal{H} = \mathcal{H}_{initial}$ at $\lambda = 0$ and $\mathcal{H} = \mathcal{H}_{final}$ at $\lambda = 1$.

A variety of different estimators can be used to analyze alchemical free energy calculations, and have different strengths and weaknesses as well as different data requirements. Here, we employ several different estimators we introduce briefly in the following.

One way to calculate the free energy difference (ΔA) between the end states is Thermodynamic Integration (TI) [4]:

$$\Delta A = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial \mathcal{H}}{\partial \lambda} \right\rangle_{\lambda} \mathrm{d}\lambda \tag{5.4}$$

in which a set of discrete λ values correspond to states along the alchemical path. $\langle\rangle$ means that we are have to calculate the ensemble average of the derivative between the brackets. TI performs as well as more efficient methods if the integrand is smooth, but breaks down if this condition is not satisfied [5, 6, 7].

An alternate free energy estimation method computes ΔA directly via:

$$\Delta A = -\frac{1}{\beta} \ln \langle \mathrm{e}^{-\beta[\mathcal{H}_{final} - \mathcal{H}_{initial}]} \rangle_{initial} \tag{5.5}$$

where the ensemble average is calculated over the configurations of the initial state, and β is the reciprocal of $k_B T$, the product between the Boltzmann constant and the absolute temperature. We call this approach exponential averaging [8] (EXP).

Most free energy calculations involve many intermediates associated with the coupling parameter (λ), allowing simulation of intermediate states in between the two end states of interest. The free energy change between the end points of a path defined by N intermedi-

ates is:

$$\Delta A = \sum_{n=1}^{N-1} \Delta A_{n \to n+1} \tag{5.6}$$

where $\Delta A_{n \to n+1}$ is the free energy difference between $(n+1)$-th and the n-th intermediate states. Eq. 5.5 can be used to calculate the free energy difference between each adjacent pair of states and yields the exact result at the limit of very large samples, but it is inefficient for most applications [84].

The Bennett acceptance ratio[1] (BAR) provides an estimator which is superior for most purposes. It calculates the free energy difference between the n-th and the $(n+1)$-th states from the following relationship:

$$\left\langle \frac{1}{1 + \frac{N_n}{N_{n+1}} \mathrm{e}^{\beta(\Delta \mathcal{H}_{n \to n+1} - \Delta A)}} \right\rangle_n = \left\langle \frac{1}{1 + \frac{N_{n+1}}{N_n} \mathrm{e}^{\beta(\Delta \mathcal{H}_{n+1 \to n} + \Delta A)}} \right\rangle_{n+1} \tag{5.7}$$

where N_n and N_{n+1} respectively are the number of statistically independent samples in states n and $n+1$, and $\Delta \mathcal{H}_{n \to n+1} = -\Delta \mathcal{H}_{n+1 \to n}$ are the Hamiltonian differences between n and $n+1$. BAR is more efficient than EXP[123, 124] and minimizes the free energy uncertainty [1]. Multistate Bennett acceptance ratio [6] (MBAR) is an extension of BAR that takes in consideration the degree of configuration space overlap between a given state and all other states in the transformation, while BAR only uses the information of neighboring states. MBAR and BAR perform similarly when the spacing between the intermediate states is moderate, but MBAR is the most well-performing free energy estimator [7].

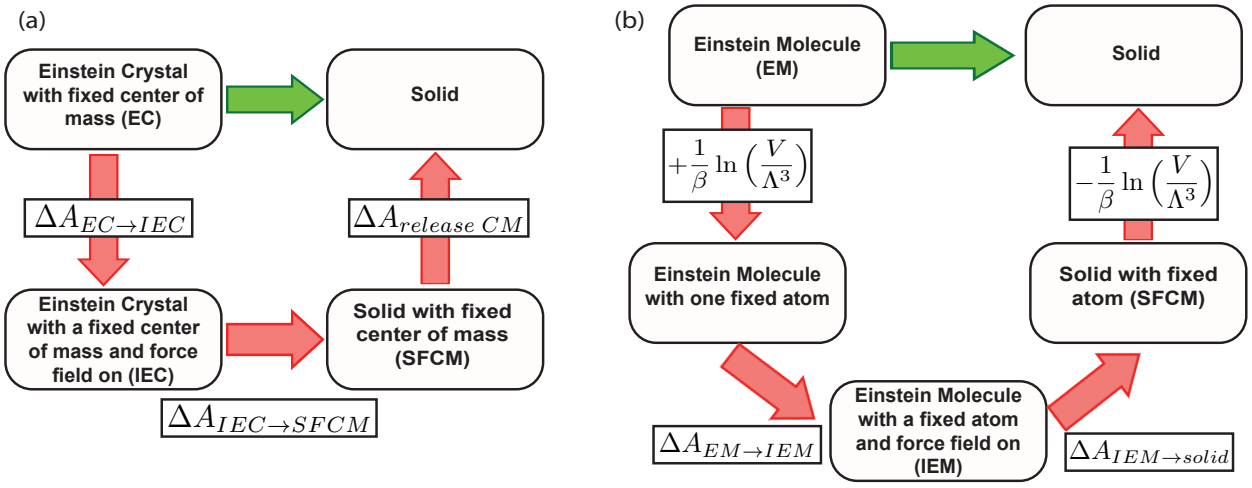

Figure 22: (a) Thermodynamic cycle representing the Einstein Crystal Method (ECM). (b) Thermodynamic cycle representing the Einstein Molecule Method (EMM). Notice that EMM requires only two free energy calculations despite being a bigger thermodynamic cycle. The canceling terms in (b) correspond to the free energies of fixing and releasing one atom in the crystal lattice [48].

5.2.3 The absolute free energy of a solid is calculated using an ideal system as reference

In this work, we seek to predict solubilities of molecular solids. Part of this problem requires predicting the free energy or chemical potential of the solid. One way this has been attempted in the past is via the Einstein Crystal Method (ECM) which calculates the absolute free energy of a solid using an Einstein crystal as a reference state. In this method, the crystal lattice is made of atoms restrained to their positions by a harmonic potential; additionally, the center of mass of the system is held fixed [294].

In the ECM, and in this work, the absolute free energy of the molecular solid is found by designing a path where force field terms are progressively turned on, and the harmonic potential position restraints are turned off. The fixed center of mass is important to avoid a quasi-divergence issue when calculating the free energy term of releasing the system from the harmonic position restraints, but the contribution of the fixed center of mass needs to be

included in the cycle to obtain the correct absolute free energy for the system (Fig. **22**(a)) [294, 48, 295].

In ECM, the free energy is calculated by:

$$A^{solid} = A^{EC}_{FCM} + \Delta A_{EC \rightarrow IEC} + \Delta A_{IEC \rightarrow SFCM} + \Delta A_{release\,CM} \tag{5.8}$$

where A^{EC}_{FCM} is the free energy of the Einstein crystal (EC) with a fixed center of mass (FCM); $\Delta A_{EC \rightarrow IEC}$ is the free energy difference between the Einstein crystal (EC) and the interacting Einstein crystal (IEC), i.e., the free energy difference in a transformation where the force field is progressively turned on throughout the calculation path. $\Delta A_{IEC \rightarrow SFCM}$ is the free energy difference between the IEC and the solid with a fixed center of mass (SFCM), i.e, turning off the harmonic restraints; and $\Delta A_{release\,CM}$ is the free energy of release of the center of mass (CM).

ECM can be difficult to implement because of the need for a fixed center of mass, so our work here is based on an alternative approach which is easier to implement. When particles move in ECM, the lattice needs to be moved because the center of mass is fixed [294, 296, 48]. Our method of choice, the Einstein Molecule Method (EMM, see Fig. **22**(b)), fixes a single atom in the lattice instead of the center of mass and is more easily implemented than ECM because of the relative difficulty of introducing center of mass restraints into existing simulation packages[296, 44, 48, 104, 297]. EMM has been used to predict phase diagrams of TIP4P and SPC/E water models [48], free energies of ice polymorphs, solid methanol and toy systems [295, 297], and the solubilities of potassium and sodium chlorides [44, 104].

In EMM, the free energy of a solid is:

$$A^{solid} = A^{EM} + \Delta A_{EM \rightarrow IEM} + \Delta A_{IEM \rightarrow solid} \tag{5.9}$$

where A^{EM} is the free energy of the ideal Einstein molecule; $\Delta A_{id \to IEM}$ is the free energy difference between the ideal Einstein molecule and the interacting Einstein molecule (i.e, turning on the force field); and $\Delta A_{IEM \to solid}$ is the free energy difference between the interacting Einstein molecule and the solid (i.e, turning off the harmonic restraints). The advantage of EMM over ECM is the absence of the need to calculate a free energy term associated with releasing the fixed reference point [48].

Here, as per equation 5.9, we compute the free energy of the solid by combining the absolute free energy of the ideal Einstein molecule with two terms that we calculate via alchemical free energy calculations — $\Delta A_{EM \to IEM}$ and $\Delta A_{IEM \to solid}$; these involve alchemically changing the interactions in the system. Numerical integration of Eq. 5.10 allows the calculation of the ideal term, A^{EM} [297]:

$$A^{EM} = -\frac{1}{\beta} \ln Q_{EM} = \frac{1}{\beta} \ln \frac{N\Lambda^3}{V} - \frac{1}{\beta} \ln \int \mathrm{e}^{-\beta U_{EM,1}(\Omega_1)\, \mathrm{d}\Omega_1} - \frac{(N-1)}{\beta} \ln \int \frac{1}{\Lambda^3} \mathrm{e}^{-\beta U_{EM,2}(r_2, \Omega_2)\, \mathrm{d}r_2\, \mathrm{d}\Omega_2} \quad (5.10)$$

where A^{EM} and Q_{EM} are the free energy of the Einstein molecule and its partition function; $U_{EM,1}(\Omega_1)$ is the potential energy of the fixed particle 1; $U_{EM,2}(r_2, \Omega_2)$ is the potential energy of a non-fixed particle at a distance r_2 of particle 1; Ω_1 and Ω_2 are all the possible orientations the molecules can have in the lattice; Λ, V, N, and β respectively are the de Broglie wavelength, the system's volume, its number of particles, and the reciprocal of $k_B T$, the product of the Boltzmann constant and the absolute temperature.

5.2.4 The chemical potential of a component of a solution can be calculated using free energy calculations.

Another critical component of computing the solubility of a compound is estimating the chemical potential of a solute in solution, since the solubility point is the concentration at which the chemical potentials of compound in the two phases are equal.

The chemical potential of a component i in solution, μ_i, has an ideal and and excess component:

$$\mu_i = -\frac{1}{\beta}\ln q_i + \frac{1}{\beta}\ln\frac{\Lambda_i^3 N_i}{V} - \frac{1}{\beta}\ln\langle \mathrm{e}^{-\beta[U(N_i+1)-U(N_i)]}\rangle_{initial} \tag{5.11}$$

where q_i is the internal partition function of a single molecule of the solute, $U(N_i)$ is the potential energy of the system with N_i particles, Λ is the de Broglie thermal wavelength, and V is the system's volume [170]. $\langle\rangle_{initial}$ means that the term was obtained from an ensemble average over the configurations of the initial state (see Eq. 5.5). The first two terms of the equation above correspond to the ideal component of μ_i; the last one, μ_i^{ex}, corresponds to the excess component of μ_i, and is associated with all non-ideal interactions of the extra component i with the solution (i.e. physical interactions that differ from those given by the ideal gas law). We obtained excess chemical potentials from solvation free energy calculations; the solute molecule is inserted in the solution by progressively turning on its interactions with the surrounding environment [53, 57, 187].

The challenge associated with the calculation of μ_i is the calculation of the standard chemical potential of i, μ_i^0, the first term of equation 5.11. q_i, the internal partition function, includes the rotation, vibrational, electronic and nuclear partition functions of a single molecule [170] and is unknown. Here, we found a way of calculating μ_i^0 without the knowledge of q_i by

alchemically transforming a single solute molecule into a single Einstein molecule, whose absolute free energy we know how to calculate [296, 48, 295].

5.2.5 Distinctives of this work

We are aware of three main approaches to compute the solubility of solids in solution using physical approaches: ECM-based methods [42, 284], EMM-based methods [44, 104, 105], and the approach of Michael Schnieders' and collaborators which computes sublimation and solvation free energies and uses these in an alternate thermodynamic cycle to obtain solubility estimates [103, 298].

Many of the applications of these approaches have been to the solubility of ionic solids, with both ECM [42] and EMM-based approaches [44, 104, 105] having some success. However, molecular solids introduce substantial additional complexities for both of these approaches.

The ECM *has* seen an initial test on solubility estimation. Li et al. [284] used the ECM to estimate the solubility of napthalene, but made several approximations such as assuming that the internal partition function component of the solute cancels between environments (perhaps justified given napthalene's low solubility).

We are not aware of any work applying the EMM to solubility estimation of molecular solids; to our knowledge our work is the first to make such an attempt, though EMM has been used before to estimate the free energy of simple molecular solids [295, 297] but not the solubility. This explains our need to find our own approach to estimate μ_i^0 for a single solute molecule.

The Schneiders approach is an orthogonal one that we do not examine here.

5.3 Methods

5.3.1 Systems under study

Here, we chose three systems to study: An argon crystal for some small initial tests, α-methanol to help establish our protocol, and acetylsalicylic acid (ASA) as our main object of study. ASA is a known anti-inflammatory whose most stable polymorph, form I [299], has an aqueous solubility of approximately 0.038 % mole fraction at 298 K [300]. We also used α-methanol at 150 K and a toy face-centered cubic (fcc) argon crystal [301] to help us find an optimal protocol to calculate the absolute free energy of a molecular solid. α-methanol was chosen because it had been used before in a study which applied the EMM to calculate the absolute free energy of the solid [297].

All simulations were run in GROMACS 4.6.7 [145, 147, 302, 149]. With one exception, all simulations used the General Amber force field (GAFF) version 1.7 with AM1-BCC charges [157, 158]; the exception was α-methanol, because we ran these simulations using the input files – coordinates and force field parameters – provided by Aragonès *et al.*, who used an united atom version of the OPLS force field [297].

We simulated all solids and liquids using 5 ns Langevin dynamics simulations. ASA, α-methanol, and argon were simulated respectively at 298.15 K, 150.0 K, and 4.0 K. Our simulations had the same length as the simulations run by Aragonès *et al.* All solid state simulations were run in NVT conditions. Liquid state simulations were run in NPT conditions; pressure was kept constant at 101.335 kPa using the Parrinello-Rahman barostat [303]. We used the TIP3P water model [159] for all our liquid state simulations. More simulation details and example input files with full details can be found in the Supporting Information.

5.3.2 Calculation of the absolute free energy of molecular crystals

The absolute free energies of the solids were calculated from trajectories of simulation boxes with 64 ASA molecules, 100 OPLS methanol molecules, and 864 argon atoms with periodic boundary conditions. ASA's unit cell was obtained from Mercury CSD 3.8 [304] and the fcc argon crystal was obtained from the literature [301]. Simulation box sizes were chosen to be approximately between 2 nm and 3 nm to ensure that box sizes were large enough that atoms and their periodic copies were not within cutoff distance of one another. α-methanol's crystal was obtained from the Supporting information of Aragonès *et al.*[297] We used Amber14's ambertools [238, 305, 306, 307] and ParmEd [205] to generate the ASA's and argon's solid state input files. All atoms but one were subjected to harmonic constraints in the x, y, and z coordinates. A single atom was kept fixed in space to act as the reference point for the calculations, as explained in the Introduction.

Monte Carlo integration yielded A_{EM}, the free energy of the Einstein molecule, as it was previously done for α-methanol in the literature[297]. $\Delta A_{id \rightarrow IEM}$ and $\Delta A_{IEM \rightarrow solid}$ were estimated using Thermodynamic Integration (TI) [4] and the Multistate Bennett Acceptance Ratio (MBAR) [2]. We used force constants of $4000\,\mathrm{k_B T/\AA^2}$ to restrain atoms to their lattice positions in acetylsalicylic acid and argon simulations because it allowed us to use a reasonable time step of 1.0 fs in all simulations. α-methanol simulations used the same force constant that had been previously used by Aragonès *et al.* [297].

We used alchemical free energy calculations to obtain the difference in free energy between the reference Einstein molecule and the solid. This step was divided in two parts: (a) the force field parameters are alchemically turned on, and (b) the harmonic constraints are turned off.

Here, we deviate from earlier work which calculated the absolute free energy of a solid using EMM by introducing additional intermediate states to improve accuracy, along with using a superior free energy estimator.

For the calculation of $\Delta A_{id \to IEM}$, we found it was crucial to introduce intermediate states; we also switched to using the MBAR estimator. The original EMM calculation of the absolute free energy of a solid [48, 295, 296, 44, 104, 297] estimated $\Delta A_{id \to IEM}$ using exponential averaging (EXP) with just two states: the Einstein molecule (EM) and the interacting Einstein molecule (IEM) [48, 295, 296, 44, 104, 297, 42, 105]. As EXP is known to have convergence issues and biases [5, 84, 6, 123], we switched to the superior MBAR free energy estimator[2]. Additionally, when we did so, we found that overlap of states (as measured by the overlap matrix [267]) was insufficient so we created a series of intermediate states connecting both ends of the transformation.

For $\Delta A_{IEM \to solid}$., the original work used Thermodynamic Integration (TI) [4]. Here, we replaced TI with MBAR as our analysis method of choice. Generally, the literature shows that TI performs as well as more efficient methods like BAR and MBAR when the integrand is smooth [5, 84, 6], but it is sensitive to the choice and number of intermediate states [127]. MBAR is the most consistently well-performing free energy estimator [7] and exploits the overlap between states more thoroughly than its predecessor, the Bennett Acceptance Ratio (BAR) estimator [2]. Here, we chose to compare performance of MBAR and TI for calculation of $\Delta A_{IEM \to solid}$ for ASA and α-methanol; we also applied EXP as a comparison in the latter case only.

5.3.3 Chemical potential calculations

The chemical potential of a pure solid is its molar free energy:

$$\mu = \frac{A}{N} \tag{5.12}$$

where N is the number of molecules in the solid, and A its Helmholtz free energy.

The chemical potential of a substance i in water is defined as the derivative of the free energy of the system with respect to the composition:

$$\mu_i = \left(\frac{\partial G}{\partial N_i}\right)_{P,T,N_{H_2O}} \tag{5.13}$$

where G is the Gibbs free energy, and N_i is the number of molecules of i in solution; P, T, and N_{H_2O} are the pressure, absolute temperature, and number of water molecules in solution, and are kept constant in the calculation.

One important aspect to discuss is reason why we chose to calculate the Helmholtz free energy for the solid and Gibbs free energies for each solution. Solid state simulations with position restraints required running under constant temperature and constant volume conditions due to software limitations, therefore we were able to calculate A for the solids.. At constant pressure, both kinds of free energy are related by:

$$\Delta G = \Delta A + P\Delta V \tag{5.14}$$

Since solids are much less susceptible to volume changes than liquids, it is reasonable to consider that $P\Delta V$ is negligible and $\Delta G \approx \Delta A$. For instance, the difference in volume between the experimental ASA crystal structure and the simulation box after a constant pressure equilibration stage is 0.14 nm^3. The $P\Delta V$ term – i.e., the free energy difference

discounting possible structure relaxation effects – would be much smaller than the simulation error.

As we explain in more detail in the Results section, successful absolute free energy calculations for molecular solids require a pathway involving a large number of alchemical intermediate states. The calculation of the absolute free energies of α-methanol at 150 K and acetylsalicylic acid required 600 states. Since GROMACS reads each λ until its fourth decimal place and the states need to be spaced more closely together as as the harmonic restraints are turned off (See Supporting Information), we decided to split each free energy calculation in sets of 100 states.

Liquid state simulation boxes were generated using the SolvationToolkit [40], a Python package that uses packmol [204], OpenMolTools (v0.6.7) [308] and OpenEye Python Toolkits [202, 155, 156]. Excess chemical potentials were obtained with the same solvation free energy protocol used in previous studies [187]: Starting from a fully interacting system, we progressively decouple the interactions of a single solute molecule with the remaining of the system, which allows us to calculate the free energy difference between a solute molecule in vacuum and in solution, i.e, the solvation free energy.

We also used alchemical free energy calculations using a single Einstein molecule as a reference state to estimate the standard chemical potential of a substance, μ_i^0:

$$\mu_i^0 = \mu_i^{ideal} - (\mu_i^{FF\ off} + \mu_i^{restraining}) \tag{5.15}$$

where $\mu_i^{FF\ off}$ and $\mu_i^{restraining}$ respectively are the chemical potential associated with turning off the force field and chemical potential of restraining the atoms of the molecule to their

Table 20: Absolute free energy components for α-methanol at 150 K, in k_BT.

	Literature [297]	**Our replica**
A_{EM}	29.05	29.24
		-38.04 ± 0.07 (EXP)
$\Delta A_{id \to IEM}$	-41.27 ± 0.01	$-41.306\,56 \pm 0.000\,04$ (MBAR, 20 states)
		$-41.275\,719 \pm 0.000\,007$ (MBAR, 40 states)
		-18.421 ± 0.005 (TI, 18 states)
$\Delta A_{IEM \to solid}$	-17.33 ± 0.03	-18 ± 3 (MBAR, 18 states)
		-17.1712 ± 0.0006 (TI 600 states)
		-17.1692 ± 0.0004 (MBAR, 600 states)

lattice positions. μ_i^{ideal} is calculated using the Monte Carlo integration procedure that we used to calculate A^{EM} to a single molecule.

5.4 Results

5.4.1 Chemical potential of molecular solids

The first step to predict aqueous solubilities with the aid of absolute free energy calculations was the assessment of the methodologies we chose to use. Since our method is the same one used by Aragonès *et al.*[297] and wanted to be sure that we could reproduce previous results, we ran simulations for α-methanol at 150 K and estimated the free energies of solids using MBAR. Turning off the harmonic restraints was the challenging step. Our MBAR calculation of $\Delta A_{IEM \to solid}$ for α-methanol using 18 intermediate states yielded $(-18 \pm 3)\,\mathrm{k_B T}$ while our TI result was $(-18.421 \pm 0.005)\,\mathrm{k_B T}$ and the literature result was $(-17.33 \pm 0.03)\,\mathrm{k_B T}$ using 17 states [297]. The MBAR error was unusually high ($3\,\mathrm{k_B T}$) which is usually a signal of overlap problems or other serious concerns.

MBAR is a free energy estimation method that minimizes the free energy variance and considers the overlap between a given state and all the others in the transformation path

[6], which means that high uncertainties ($\pm 3k_BT$) suggest the presence of problems in the transformation's path. TI's uncertainty estimates are much lower, but we believe this is an artifact. Error analysis for TI simply does not work the same way and does not give insight into whether exploration of phase space is adequate, unlike MBAR. Specifically, uncertainty estimates from TI usually factor in only the uncertainty in the integrand at each sampled lambda value and could potentially also factor in the smoothness of the integrand (i.e. numerical integration error) but do nothing to factor in whether the integrand will in fact vary smoothly in between lambda points; usually no data is available on this. BAR and MBAR, in contrast, factor in information about how well the intermediate states overlap in phase space and reflect high uncertainties when phase space overlap is poor. In our experience, usually TI would suffer from similar problems if additional intermediate states were added, but uncertainties in TI typically do not reflect this, as is the case here. Thus, the high uncertainty of the MBAR value indicates a sampling/convergence problem which warrants further exploration.

To explore the high uncertainty of our MBAR free energy estimates, we examined the degree of overlap the intermediate states had with each other. Phase space overlap analysis [309, 194, 128] quantifies the probability that any given configuration of an intermediate state can be found in other states. A good rule of thumb for designing a set of free energy calculations spanning between two states is to ensure that the states along the path have significant overlap with their neighbors as shown in Fig. **23**. More overlap improves the quality of the MBAR free energy estimation: Fig. **23**(b) represents a set of restraining simulations where the free energy uncertainty can potentially be accurately estimated using BAR and MBAR; Fig. **23**(a) shows a case where it cannot. In our case we find that the α-methanol simulation using 18 intermediate states does not have adequate overlap (Fig. **24**)– specifically, the states $4 \leq \lambda_i \leq 17$ do not have overlapping configurations with other states, which explains the $3\,\mathrm{k_BT}$ uncertainty in our MBAR estimate.

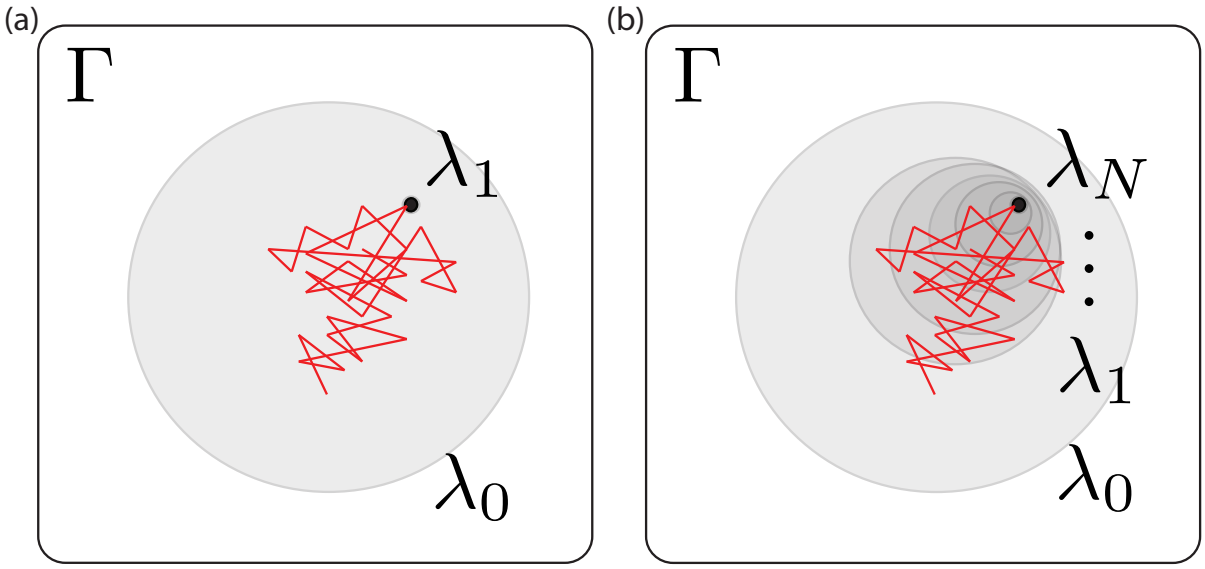

Figure 23: Phase space overlap between the states in a thermodynamic path for removing restraints with λ. Γ represents the phase space that contains all the configurations for all the states in the path. λ_0 and λ_1 (left) or λ_N (right) represent the end states along the path, each shaded region represents a state in phase space and the red lines represent the configurations visited by the simulation run in the λ_0 state. The restrained state is a subset of the unrestrained one. (a) and (b) represent simulations with different numbers of intermediate states along the path between a fully restrained state (λ_1 (a) or λ_N (b)) and an unrestrained state (λ_0). In (a), the simulation (red) only visits very few configurations consistent with the restrained state – i.e, there is poor phase space overlap – indicating a need for more intermediate states, otherwise any free energy estimates will be subject to very high uncertainties; in (b) there is still almost no overlap between the simulation and states consistent with λ_N, but there is overlap with the next shaded region, λ_1, indicating the potential for overlap and accurate free energy estimates. Thus simulations run in each shaded region are more likely to have a bigger phase space overlap with λ_N than simulations run in λ_0.

Since prior work had appeared to do this estimation successfully [297], we were uncertain why we were encountering such overlap problems, so we studied an even simpler system. We calculated $\Delta A_{IEM \rightarrow solid}$ of fcc argon at 4 K with 18 states as in our α-methanol free energy estimation. MBAR yielded an error estimate of infinity while TI estimated $\Delta A_{IEM \rightarrow solid}$ to be $(-1666.5 \pm 0.8)\,k_BT$ which, as we show below, is incorrect. This path resulted phase space overlap diagram without overlap between the states after state number 2 (Fig. **25**). Apparently as the harmonic potential that holds atoms in their lattice positions tends to

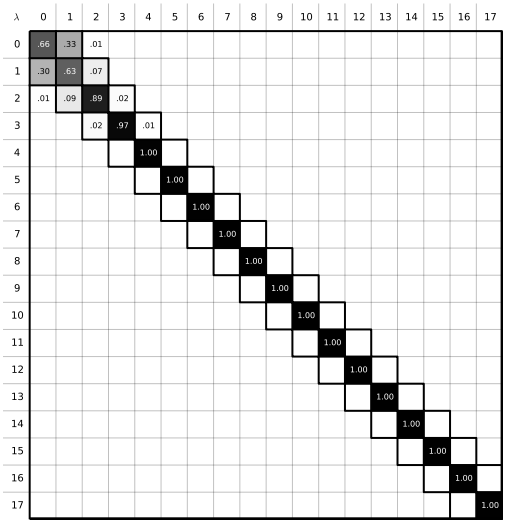

Figure 24: Phase space overlap between the states in the path between IEM and the α-methanol solid. The sum of all the elements in a row should yield 1.0, a probability of 100 %. A good free energy estimate is obtained when the states along the alchemical path contain configurations that can be found in other intermediate states. In these situations, the phase space overlap is non-zero, which results in non-zero off-diagonal elements. Here, however, the phase space overlap plot shows that there is no overlap between the states λ_i, $4 \leq i \leq 17$ indicating poor free energy estimates will result.

zero, atoms become rather mobile, dramatically decreasing phase space overlap and leading to poor free energy estimates.

To improve phase space overlap, we introduced more intermediate states along the path for removing the restraints (see Fig **23**). We chose to break down the simulation in smaller parts, adding a significant amount of states near the point where the harmonic restraints are approximately zero. The MBAR estimate of $\Delta A_{IEM \rightarrow solid}$ for fcc argon is $(-1016.0 \pm 0.2)\,\mathrm{k_B T}$ using 300 states. TI's corresponding value was $(-1017 \pm 1)\,\mathrm{k_B T}$, differing by far from the (incorrect) value of $(-1666.5 \pm 0.8)\,\mathrm{k_B T}$ obtained above with fewer states. Phase space overlap diagrams showed significant improvement in the configuration overlap between the states (Supporting Information). Thus, increasing the number of states was an effective strategy, and we used it in all subsequent calculations.

Even though our α-methanol results were similar to results previously published by other authors [297], we need to emphasize that reliable free energies resulted from simulations with

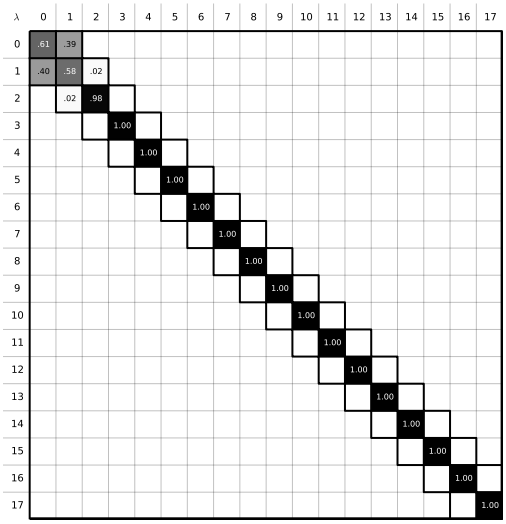

Figure 25: Phase space overlap between the states in the path between IEM and the fcc argon solid. A good free energy estimate is obtained when the states along the alchemical path contain configurations that can be found in other intermediate states. Here, however, the phase space overlap diagram shows that there is no overlap between the states λ_i, $3 \leq i \leq 17$, which explains the poor quality of the free energy result.

a large number of intermediate states, as can be seen in Table **20**. Despite its conceptual simplicity, calculating the components of the absolute free energy of a solid to a point where there is significant phase space overlap between the intermediate states is computationally demanding. A 900-atom OPLS α-methanol system required 40 states to calculate $\Delta A_{id \to IEM}$, and 600 states for $\Delta A_{IEM \to solid}$.

We chose these intermediate states in advance, and these ultimately led to free energy errors smaller than $0.1\,\mathrm{k_B T}$; the estimated TI and MBAR values differed by no more than $0.3\,\mathrm{k_B T}$. Our results for ASA using an optimal number of states can be seen in Table **21**. The MBAR chemical potential of ASA at 298.15 K equals to $(-220.67 \pm 0.03)\,\mathrm{k_B T}$.

The computational cost of calculating A^{ASA} was high; Each state required a separate simulation (of a 1344-atom ASA system), with 718 states in total. Simulations typically required 11 hours on a single CPU, so the calculation of a single absolute free energy of a molecular solid required approximately 7898 CPU-hours.

Table 21: Absolute free energy components for polymorph I of acetylsalicylic acid (ASA) at 298.15 K, in k_BT.

	Acetylsalicylic Acid
A_{EM}	48.047
$\Delta A_{id \to IEM}$	-167.316 ± 0.001 (TI, 118 states)
	-167.07 ± 0.03 (MBAR, 118 states)
$\Delta A_{IEM \to solid}$	-101.656 ± 0.002 (TI, 600 states)
	-101.644 ± 0.002 (MBAR, 600 states)

5.4.2 Chemical potential of solutions and the solubility of GAFF acetylsalicylic acid in TIP3P water.

Equation 5.11 states that the absolute chemical potential of a solution is determined by three quantities: μ_i^0, the standard chemical potential; μ_i^{ex}, the excess chemical potential of the component at a concentration of χ; and a volume-dependent ideal gas component of $k_BT \cdot \ln(\Lambda_i^3 \cdot N_{ASA}/\langle V \rangle_{solution})$. μ_{ASA}^0 only required information regarding the internal structure of the molecule [170], thus we estimated μ_{ASA}^0 by alchemically transforming a single solute molecule into a single Einstein molecule (Table **22**), whose absolute free energy we know how to calculate. We used the same number of states that we chose for the solid state simulations and we found that μ_{ASA}^0 is equal to $(-150.7 \pm 0.2)\,\mathrm{k_B T}$, as discussed in the last subsection of the Methods section.

Table 22: Standard chemical potential of acetylsalicylic acid (ASA) at 298.15 K, in k_BT.

	Acetylsalicylic Acid
μ_{ASA}^{ideal}	9.3
$\mu_{ASA}^{FF\,off}$	65.7409 ± 0.0009 (MBAR 118 states)
$\mu_{ASA}^{restraining}$	94.3 ± 0.2 (MBAR 600 states)

Concentrations, volumes and excess chemical potentials can be seen in Table **23**. We obtained the excess chemical potentials from solvation free energy calculations [53, 57, 187]. Volumes were obtained from the state in the alchemical path where the solute was fully coupled to the rest of the system.

Table 23: Simulation data for solutions of acetyl salicylic acid in water in different concentrations.

Molar fraction (%)	**Volume** (nm^3)	**# solute molecules**	**# solvent molecules**	μ^{ex} (k_BT)
2.000 e-03	3035.99 ± 0.05	2	99998	-16.80 ± 0.05
6.666 e-03	911.17 ± 0.02	2	30002	-15.88 ± 0.04
7.999 e-03	759.33 ± 0.01	2	25000	-15.51 ± 0.05
9.998 e-03	911.45 ± 0.03	3	30003	-15.65 ± 0.04
9.999 e-03	607.59 ± 0.02	2	20000	-15.47 ± 0.05
1.3330 e-02	911.72 ± 0.02	4	30004	-15.77 ± 0.04
1.3332 e-02	455.84 ± 0.02	2	15000	-15.61 ± 0.04
1.666 e-02	912.00 ± 0.03	5	30005	-15.96 ± 0.05
1.9992 e-02	912.27 ± 0.02	6	30006	-15.78 ± 0.04
1.9996 e-02	304.01 ± 0.01	2	10000	-15.62 ± 0.05
3.998 e-02	152.25 ± 0.01	2	5000	-15.41 ± 0.06
1.996 e-01	30.835 ± 0.007	2	1000	-16.37 ± 0.05
2.991 e-01	31.069 ± 0.003	3	1000	-16.40 ± 0.06
3.984 e-01	31.309 ± 0.007	4	1000	-16.62 ± 0.06
4.975 e-01	31.547 ± 0.003	5	1000	-17.1 ± 0.1

The experimental aqueous solubility of ASA is approximately 0.038 % in water at 298 K [300], but our model predicts that ASA is effectively insoluble in water (Figure **26**). While all-atom simulations can yield solubility estimates given adequate simulation time and a correct method, the computed solubility will be that dictated by the underlying energy model or force field, and will not necessarily match experiment. Here, we use GAFF, a general-purpose force field with known limitations [69, 306, 73, 187]; apparently, here, the right answer **for the force field** is not correct. Perhaps this is because of limitations in describing the solid state, as the force field is parameterized for liquid state simulations. Indeed, classical fixed charge force fields have shown severe limitations for polymorph prediction for these reasons [288, 291, 290, 274, 292] . Also, point partial atomic charges regularly used in molecular dynamics do not describe electrostatic interactions in a solid particularly well [310]. In the case of the ASA crystal, it is possible that its hydrogen bonds and π-stacking interactions add layers of complexity that are not properly described by GAFF.

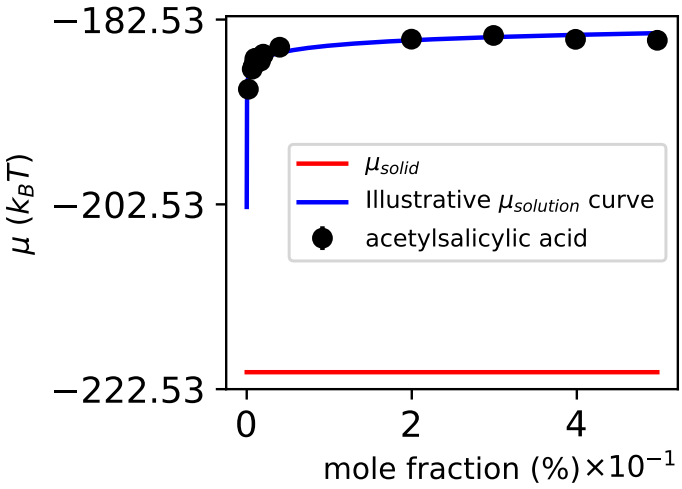

Figure 26: Chemical potentials of ASA, solid and solution in different concentrations, with respect to mole fraction.

5.5 Discussion and Conclusions

Despite its theoretical rigor, solubility prediction from absolute free energy calculations is a difficult task: it is computationally expensive and, at least in the present approach, requires many different steps and a great deal of care. Here, we attempted to develop and test a general approach to compute the solubility of molecular solids by adapting the Einstein Molecule Method (EMM) to tackle this problem, as discussed above.

To tune our methodology, we initially decided to reproduce the absolute free energy of solid α-methanol, one of methanol's polymorphs, at 150 K using EMM before doing the same calculations for our compound of choice, ASA. We verified that the free energy differences between the Einstein molecule and the interactive Einstein molecule ($\Delta A_{EM \to IEM}$) and between the latter state and the solid ($\Delta A_{IEM \to solid}$) were more reliably estimated with the multistate Bennett acceptance ratio (MBAR). The absolute free energy of the crystal (as computed for united-atom OPLS α-methanol) agreed with results found in the literature, which suggested that we were on the right path. We did, however, require a very large

number of intermediate alchemical states to obtain accurate free energy estimates, making these simulations fairly computationally demanding.

We then chose to calculate the solubility of ASA due to its pharmacological importance and due to its relative complexity in comparison to previous molecular solids whose absolute free energies have been computed via EMM previously [297]. As for α-methanol, this calculation required a large number of intermediate alchemical states and considerable computational cost – approximately 8000 CPU hours for a single absolute free energy calculation for the molecular solid, even with the crystal structure as input. Perhaps the number of intermediate states could be further optimized, but clearly a large number of intermediate simulations was required and thus considerable computational cost. Despite all of this, we still could not reproduce the experimental aqueous solubility of acetylsalicylic acid (ASA); experimentally it is modestly soluble, whereas our work would suggest it is essentially completely insoluble in water, likely due to force field limitations.

The solubility of naphthalene was recently estimated using a similar methodology, the Extended Einstein Crystal Method [284], but with additional approximations. Specifically, since naphthalene molecules interact very weakly with each other in the crystal lattice and with water molecules in solution, the differences between the internal partition function of a naphthalene molecule in the solid and in the solution were assumed to be negligible. This allowed the authors to drop some complexities in treatment of the solution-phase part of the calculation. However, that approach is only suitable for compounds that are only very weakly interacting in solution and in the crystal. ASA, in contrast, is a molecule that interacts strongly with other ASA molecules in its crystal lattice and with water molecules in solution via hydrogen bonds. For instance, an important crystalline feature that is not necessarily present in solution is the dimer structure, with two ASA molecules bound together via hydrogen bonds between the carboxylic acid groups. Differences between the internal partition functions of the molecule in the solid (q_{ASA}^{solid}) and in solution ($q_{ASA}^{solution}$) would prob-

ably not be negligible in this scenario, thus a more general approach is needed for treatment of such cases. Our work here provides one attempt in that direction.

Overall, the present approach seems to have significant limitations – most notably that the computational expense is considerable, and the resulting estimated solubility is quite inaccurate. Perhaps both of these may be surmountable; GPU-based free energy calculations can be dramatically faster potentially making an 8000 CPU-hour calculation be 80 GPU hours which would amount to overnight on 8 GPUs, and perhaps this could be optimized via changes to simulation time and number of intermediate states. And with better force fields, perhaps accuracy could be improved; the AMOEBA-based approach of Schnieders shows considerable promise [103]. Alteratively, other approaches may be of interest. Solubility has been predicted by simulations using pseudocritical paths (i.e., paths were molecular crystals are transformed in tractable Einstein crystal-like states between the ending states of the transformation [311, 312, 313, 314]) and a single experimental reference point [313]), and with the aid of a thermodynamic cycle formed by the molecular crystal, the molecule in vacuum, and the solvated molecule [103].

We believe the time has come for routine physical methods for estimation of solubility, even if improved force fields prove necessary before results have significant accuracy for application to biomolecular design problems.

5.6 Supporting Information

5.6.1 Simulation details

The following are GROMACS 4.6.7 simulation input parameters, as are the MDP files with full details which are deposited in the Supporting Information.

General information

- Friction coefficient = $\text{mass}_{particle}/\tau_t$, $\tau_t = 2.0\,ps$.
- Parrinello-Rahman barostat (when applicable): $\tau_p = 10\,ps$ and compressibility = $4.5 \cdot 10^{-5}\,bar^{-1}$.

Electrostatics (solid)

- PME cut-off: 1.0 nm.
- PME order: 4
- Fourier spacing = 0.10 nm
- we used the same parameters as the Aragonès et al in solid state simulations. Additional details can be found in the MDP files deposited with this paper.

Electrostatics (solution)

- PME cut-off: 1.2 nm.
- PME order: 6
- Fourier spacing = 0.10 nm
- we used the same parameters as the Aragonès et al. for solid state simulations. Additional details can be found in the MDP files deposited with this paper.

vdW interactions

- Cut-off: $1.0\,nm$

- Switch at $0.9\,nm$
- DispCorr = AllEnerPres
- additional details can be found in the MDP files deposited with this paper.

Solution simulation files were generated using the `SolvationToolkit` module found at `https://github.com/MobleyLab/SolvationToolkit`. Solvation Toolkit relies on `openmoltools`, `mdtraj`, `packmol`, `ParmEd`, and OpenEye tools. As noted in the main body of the text, AM1-BCC charges were assigned with OpenEye's quacpac python module; we used `openmoltools` to drive this process. Specific source code used for charging is available at `https://github.com/choderalab/openmoltools/blob/v0.6.7/openmoltools/openeye.py#L13`. The code generates molecular conformations prior to charging, as was recommended at `http://docs.eyesopen.com/toolkits/cookbook/python/modeling/am1-bcc.html`.

We generated the simulation files for the acetylsalicylic solid simulations using the `pdb` file containing the crystal structure, Antechamber and `ParmEd`. We used the Antechamer AM1-BCC procedure to generate the charges for the solid state simulations.

Free energy estimation was done using the `alchemical_analysis` script which can be found at `https://github.com/MobleyLab/alchemical-analysis`.

5.6.2 Supporting details

The `DA_ideal_to_IEM.csv` file containing the elements of the phase space overlap matrix of a $\Delta A_{EM \to IEM}$ estimated from an alchemical path of 118 states can be found at `https://f1000research.com/articles/7-686/`.

Chapter 6

Concluding Remarks

Computational resources now are more powerful and more accessible than they were in the early days of computational chemistry and can perhaps be used to further the research on the application of free energy calculations to more sophisticated problems. Amid free energy calculations, solvation and hydration free energies have some degree of prominence: despite their relative simplicity – one just decouples the interaction of a solute molecule with its solvent surroundings and calculates ΔG from the energy differences or from how the energy varies with respect to the state – solvation free energy calculations are applicable in force field parameterization, modeling implicit solvents, drug design, and method development. In this dissertation, I described and discussed the results of four different projects involving solvation free energy calculations.

Chapter 2 outlines a general protocol to run solvation free energy calculations. It suggests to decouple electrostatic and Lennard-Jones terms separately from each other, to use soft-core potentials to avoid numerical instabilities, and to use MBAR whenever possible. Besides the protocol for hydration free energy simulations, Ch. 2 contains calculated hydration enthalpies and entropies and a few experimental enthalpies obtained from ORCHYD [169].

Since enthalpies are more sensitive to force field parameters than free energies, a possible future development would be an investment in obtaining experimental values of different hydration enthalpies [153, 166, 167].

Chapter 3 discusses the calculation of infinite dilution activity coefficients (IDACs) and their potential use in force field testing and assisting with force field parameterization. I calculated a plethora of solvation free energies using Yank in OpenEye Software's Orion platform, and a series of densities using OpenMM. The results of the calculations were encouraging and hopefully will drive future research in expanding the number of experimental IDAC values and in perhaps using IDACs instead of hydration free energies in method development.

In chapter 4 I detailed a collaborative project that aimed at benchmarking relative alchemical free energy calculations and reporting simulation protocols that could be reproduced using different codes. Chapter 4 is a good example of how solvation free energies can be used to test methods and validate simulation protocols. We found that relative alchemical free energy calculations in AMBER [238], GROMACS [150], SOMD [241, 208], and CHARMM [239] reproduce results to within 0.2 $kcal \cdot mol^{-1}$. Our results were validated by comparison with $\Delta\Delta G^{hydration}$ obtained from absolute free energy calculations. It was not possible to define a universal protocol for all four codes, but we made available our input data and protocols to be used in follow-up studies.

Chapter 5 discusses the prediction of aqueous solubilities using solvation free energy calculations and the absolute chemical potential of acetylsalicylic acid. The chapter contains a thorough discussion of the effect of poor phase space overlap between intermediate states and the considerable expense of the method. A low number of intermediate states, as Aragonès et al. had done it [297] does not generate a reliable estimate for the free energy uncertainty. Since the method is physically exact, the chapter also discusses how we could more accurately predict the solubility of molecular solids with the aid of better force fields and computational resources.

There is much room for growth in the field of free energy calculations. Solvation free energy calculations of small molecules can be very useful in this effort due to the simplicity of their protocol and the ease to generate a set of uncorrelated samples. IDACs, on their turn, show a lot of promise as tools for assisting force field development and method testing. I hope the data and studies presented here will drive new work on the use of free energy calculations in complex systems with the assistance of solvation free energies.

Bibliography

[1] Charles H Bennett. Efficient estimation of free energy differences from Monte Carlo data. *Journal of Computational Physics*, 22(2):245–268, October 1976.

[2] Michael R. Shirts and John D. Chodera. Statistically optimal analysis of samples from multiple equilibrium states. *The Journal of Chemical Physics*, 129(12):124105, September 2008.

[3] Cristophe Chipot and Andrew Pohorille. *Free Energy Calculations Theory and Applications in Chemistry and Biology*. Springer, 2007.

[4] John G. Kirkwood. Statistical Mechanics of Fluid Mixtures. *The Journal of Chemical Physics*, 3(5):300–313, May 1935.

[5] F. Marty Ytreberg, Robert H. Swendsen, and Daniel M. Zuckerman. Comparison of free energy methods for molecular systems. *The Journal of Chemical Physics*, 125(18):184114, November 2006.

[6] Michael R. Shirts and Vijay S. Pande. Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration. *The Journal of Chemical Physics*, 122(14):144107, April 2005.

[7] Himanshu Paliwal and Michael R. Shirts. A Benchmark Test Set for Alchemical Free Energy Transformations and Its Use to Quantify Error in Common Free Energy Methods. *Journal of Chemical Theory and Computation*, 7(12):4115–4134, 2011.

[8] Robert W. Zwanzig. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *The Journal of Chemical Physics*, 22(8):1420–1426, August 1954.

[9] J. P. Valleau and D. N. Card. Monte Carlo Estimation of the Free Energy by Multistage Sampling. *The Journal of Chemical Physics*, 57(12):5457–5462, December 1972.

[10] Thomas C. Beutler, Alan E. Mark, René C. van Schaik, Paul R. Gerber, and Wilfred F. van Gunsteren. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chemical Physics Letters*, 222(6):529–539, June 1994.

[11] M. Zacharias, T. P. Straatsma, and J. A. McCammon. Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. *The Journal of Chemical Physics*, 100(12):9025–9031, June 1994.

[12] Michael R Shirts, David L Mobley, and Scott P Brown. Free-energy calculations in structure-based drug design. *Drug Design*, pages 61–86, 2010.

[13] Thomas Steinbrecher, Insuk Joung, and David a. Case. Soft-core potentials in thermodynamic integration: Comparing one-and two-step transformations. *Journal of Computational Chemistry*, 32(15):3253–3263, 2011.

[14] David L. Mobley. Let's get honest about sampling. *Journal of computer-aided molecular design*, 26(1):93–5, January 2012.

[15] William L. Jorgensen and C. Ravimohan. Monte Carlo simulation of differences in free energies of hydration. *J. Chem. Phys.*, 83(6):3050–3054, 1985.

[16] Nandou Lu, David A. Kofke, and Thomas B. Woolf. Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods. *Journal of Computational Chemistry*, 25(1):28–40, January 2004.

[17] G. M. Torrie and J. P. Valleau. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, 23(2):187–199, February 1977.

[18] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov. New approach to Monte Carlo calculation of the free energy: Method of expanded ensembles. *The Journal of Chemical Physics*, 96(3):1776–1783, February 1992.

[19] E. Marinari and G. Parisi. Simulated Tempering: A New Monte Carlo Scheme. *EPL (Europhysics Letters)*, 19(6):451, 1992.

[20] Ulrich H. E. Hansmann. Parallel tempering algorithm for conformational studies of biological molecules. *Chemical Physics Letters*, 281(1):140–150, December 1997.

[21] Yuji Sugita, Akio Kitao, and Yuko Okamoto. Multidimensional replica-exchange method for free-energy calculations. *The Journal of Chemical Physics*, 113(15):6042–6051, October 2000.

[22] Hiroaki Fukunishi, Osamu Watanabe, and Shoji Takada. On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *The Journal of Chemical Physics*, 116(20):9058–9067, May 2002.

[23] Christopher J. Woods, Jonathan W. Essex, and Michael A. King. The Development of Replica-Exchange-Based Free-Energy Methods. *The Journal of Physical Chemistry B*, 107(49):13703–13710, December 2003.

[24] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092, June 1953.

[25] I. R. McDonald and K. Singer. Machine Calculation of Thermodynamic Properties of a Simple Fluid at Supercritical Temperatures. *The Journal of Chemical Physics*, 47(11):4766–4772, December 1967.

[26] I. R. McDonald and K. Singer. Calculation of thermodynamic properties of liquid argon from Lennard-Jones parameters by a Monte Carlo method. *Discussions of the Faraday Society*, 43(0):40–49, January 1967.

[27] Jong K. Lee, J. A. Barker, and Farid F. Abraham. Theory and Monte Carlo simulation of physical clusters in the imperfect vapor. *The Journal of Chemical Physics*, 58(8):3166–3180, April 1973.

[28] David J. McGinty. Molecular dynamics studies of the properties of small clusters of argon atoms. *The Journal of Chemical Physics*, 58(11):4733–4742, June 1973.

[29] Susumu Okazaki, Koichiro Nakanishi, Hidekazu Touhara, and Yoshinori Adachi. Monte Carlo studies on the hydrophobic hydration in dilute aqueous solutions of nonpolar molecules. *The Journal of Chemical Physics*, 71(6):2421–2429, September 1979.

[30] Michael R. Mruzik, Farid F. Abraham, Donald E. Schreiber, and G. M. Pound. A Monte Carlo study of ion–water clusters. *J. Chem. Phys.*, 64(2):481–491, 1976.

[31] Johan P. M. Postma, Herman J. C. Berendsen, and Jan R. Haak. Thermodynamics of cavity formation in water. A molecular dynamics study. *Faraday Symp. Chem. Soc.*, 17(0):55–67, 1982.

[32] Charles L. Brooks III. Thermodynamics of ionic solvation: Monte Carlo simulations of aqueous chloride and bromide ions. *The Journal of Physical Chemistry*, 111(9):6680–6684, 1986.

[33] Hsiang-Ai Yu and Martin Karplus. A thermodynamic analysis of solvation. *The Journal of Chemical Physics*, 89(4):2366–2379, August 1988.

[34] Bhalachandra L. Tembre and J. Andrew Mc Cammon. Ligand-receptor interactions. *Computers & Chemistry*, 8(4):281–283, January 1984. His name is Tembe, not Tembre.

[35] M K Gilson, J A Given, B L Bush, and J A McCammon. The statistical-thermodynamic basis for computation of binding affinities: A critical review. *Biophysical journal*, 72(3):1047–69, March 1997.

[36] P. A. Bash, U. C. Singh, R. Langridge, and P. A. Kollman. Free energy calculations by computer simulation. *Science*, 236(4801):564–568, May 1987.

[37] Bin Chen and J. Ilja Siepmann. Partitioning of Alkane and Alcohol Solutes between Water and (Dry or Wet) 1-Octanol. *Journal of the American Chemical Society*, 122(27):6464–6467, 2000.

[38] Alexander P. Lyubartsev, Sven P. Jacobsson, G"\{o}ran Sundholm, and Aatto Laaksonen. Solubility of Organic Compounds in Water/Octanol Systems. A Expanded Ensemble Molecular Dynamics Simulation Study of log P Parameters. *The Journal of Physical Chemistry B*, 105(32):7775–7782, 2001.

[39] Julien Michel, Mario Orsi, and Jonathan W. Essex. Prediction of Partition Coefficients by Multiscale Hybrid Atomic-Level/Coarse-Grain Simulations. *The Journal of Physical Chemistry B*, 112(3):657–660, 2008.

[40] Caitlin C. Bannan, Gaetano Calabró, Daisy Y. Kyu, and David L. Mobley. Calculating Partition Coefficients of Small Molecules in Octanol/Water and Cyclohexane/Water. *Journal of Chemical Theory and Computation*, 12(8):4015–4024, 2016.

[41] Caitlin C. Bannan, Kalistyn H. Burley, Michael Chiu, Michael R. Shirts, Michael K. Gilson, and David L. Mobley. Blind Prediction of Cyclohexane-water Distribution Coefficients from the SAMPL5 Challenge. *J. Comput.-Aided Mol. Des.*, 30(11):927–944, September 2016.

[42] Mauro Ferrario, Giovanni Ciccotti, Eckhard Spohr, Thierry Cartailler, and Pierre Turq. Solubility of KF in water by molecular dynamics using the Kirkwood integration method. *The Journal of Chemical Physics*, 117(10):4947–4953, August 2002.

[43] Kelly E Anderson and J Ilja Siepmann. Molecular simulation approaches to solubility. In *Developments and Applications in Solubility*, pages 171–187. IUPAC, Oxford, 2007.

[44] E. Sanz and C. Vega. Solubility of KF and NaCl in water by molecular simulation. *The Journal of Chemical Physics*, 126(1):014507, January 2007.

[45] Antonio Llinàs, Robert C. Glen, and Jonathan M. Goodman. Solubility Challenge: Can You Predict Solubilities of 32 Molecules Using a Database of 100 Reliable Measurements? *Journal of Chemical Information and Modeling*, 48(7):1289–1303, July 2008.

[46] Nigel B. Wilding. Computer simulation of fluid phase transitions. *American Journal of Physics*, 69(11):1147–1155, October 2001.

[47] Nandou Lu, C. Daniel Barnes, and David a. Kofke. Free-energy calculations for fluid and solid phases by molecular simulation. *Fluid Phase Equilibria*, 194-197:219–226, March 2002.

[48] C Vega, E Sanz, J L F Abascal, and E G Noya. Determination of phase diagrams via computer simulation: Methodology and applications to water, electrolytes and proteins. *Journal of Physics: Condensed Matter*, 20(15):153101, April 2008.

[49] M. F. Gautreaux and Jesse Coates. Activity coefficients at infinite dilution. *AIChE Journal*, 1(4):496–500, December 1955.

[50] C. A. Eckert, B. A. Newman, G. L. Nicolaides, and T. C. Long. Measurement and application of limiting activity coefficients. *AIChE Journal*, 27(1):33–40, 1981.

[51] M. S. H. Bader and K. A. M. Gasem. Determination of Infinite Dilution Activity Coefficients for Organic-Aqueous Systems Using a Dilute Vapor-Liquid Equilibrium Method. *Chemical Engineering Communications*, 140(1):41–72, October 1995.

[52] J Peter Guthrie. A blind challenge for computational solvation free energies: Introduction and overview. *The Journal of Physical Chemistry B*, 113(14):4501–4507, April 2009.

[53] Pavel V. Klimovich and David L. Mobley. Predicting hydration free energies using all-atom molecular dynamics simulations and multiple starting conformations. *Journal of Computer-Aided Molecular Design*, 24(4):307–316, 2010.

[54] David L. Mobley, Ken A. Dill, and John D. Chodera. Treating Entropy and Conformational Changes in Implicit Solvent Simulations of Small Molecules. *The Journal of Physical Chemistry B*, 112(3):938–946, January 2008.

[55] Robert C. Harris and B. Montgomery Pettitt. Effects of geometry and chemistry on hydrophobic solvation. *Proceedings of the National Academy of Sciences of the United States of America*, 111(41):14681–14686, October 2014.

[56] Anthony Nicholls, David L. Mobley, J. Peter Guthrie, John D. Chodera, Christopher I. Bayly, Matthew D. Cooper, and Vijay S. Pande. Predicting small-molecule solvation free energies: An informal blind test for computational chemistry. *Journal of Medicinal Chemistry*, 51(4):769–779, February 2008.

[57] David L. Mobley and J. Peter Guthrie. FreeSolv: A database of experimental and calculated hydration free energies, with input files. *Journal of Computer-Aided Molecular Design*, 28(7):711–720, July 2014.

[58] Matthew T Geballe, A Geoffrey Skillman, Anthony Nicholls, J Peter Guthrie, and Peter J Taylor. The SAMPL2 blind prediction challenge: Introduction and overview. *Journal of Computer-Aided Molecular Design*, 24(4):259–279, April 2010.

[59] Matthew T Geballe and J Peter Guthrie. The SAMPL3 blind prediction challenge: Transfer energy overview. *Journal of Computer-Aided Molecular Design*, 26(5):489–496, May 2012.

[60] Emilio Gallicchio and Ronald M. Levy. Prediction of SAMPL3 host-guest affinities with the binding energy distribution analysis method (BEDAM). *Journal of Computer-Aided Molecular Design*, 26(5):505–516, February 2012.

[61] Emilio Gallicchio, Nanjie Deng, Peng He, Lauren Wickstrom, Alexander L. Perryman, Daniel N. Santiago, Stefano Forli, Arthur J. Olson, and Ronald M. Levy. Virtual screening of integrase inhibitors by large scale binding free energy calculations: The SAMPL4 challenge. *Journal of Computer-Aided Molecular Design*, 28(4):475–490, February 2014.

[62] Hari S. Muddana, Neil V. Sapra, Andrew T. Fenley, and Michael K. Gilson. The SAMPL4 hydration challenge: Evaluation of partial charge sets with explicit-water molecular dynamics simulations. *Journal of Computer-Aided Molecular Design*, 28(3):277–287, March 2014.

[63] Samuel Genheden. Predicting Partition Coefficients with a Simple All-Atom/Coarse-Grained Hybrid Model. *Journal of Chemical Theory and Computation*, 12(1):297–304, 2016.

[64] Nuno M. Garrido, António J. Queimada, Miguel Jorge, Eugénia A. Macedo, and Ioannis G. Economou. 1-Octanol/Water Partition Coefficients of n-Alkanes from Molecular Simulations of Absolute Solvation Free Energies. *Journal of Chemical Theory and Computation*, 5(9):2436–2446, 2009.

[65] Nuno M. Garrido, Ioannis G. Economou, António J. Queimada, Miguel Jorge, and Eugénia A. Macedo. Prediction of the n-hexane/water and 1-octanol/water partition coefficients for environmentally relevant compounds using molecular simulation. *AIChE Journal*, 58(6):1929–1938, June 2012.

[66] Li Yang, Alauddin Ahmed, and Stanley I. Sandler. Comparison of two simulation methods to compute solvation free energies and partition coefficients. *Journal of Computational Chemistry*, 34(4):284–293, February 2013.

[67] J M Wang, R M Wolf, J W Caldwell, P a Kollman, and D a Case. Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9):1157–1174, 2004.

[68] Michael R. Shirts, Jed W. Pitera, William C. Swope, and Vijay S. Pande. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *The Journal of Chemical Physics*, 119(11):5740–5761, September 2003.

[69] David L. Mobley, Elise Dumont, John D. Chodera, and Ken A. Dill. Comparison of charge models for fixed-charge force fields: Small-molecule hydration free energies in explicit solvent. *The Journal of Physical Chemistry. B*, 111(9):2242–2254, March 2007.

[70] Shuai Liu, Shannon Cao, Kevin Hoang, Kayla L. Young, Andrew S. Paluch, and David L. Mobley. Using MD Simulations To Calculate How Solvents Modulate Solubility. *Journal of Chemical Theory and Computation*, 12(4):1930–1941, April 2016.

[71] Jin Zhang, Badamkhatan Tuguldur, and David van der Spoel. Force Field Benchmark of Organic Liquids. 2. Gibbs Energy of Solvation. *Journal of Chemical Information and Modeling*, 55(6):1192–1201, June 2015.

[72] Chris Oostenbrink, Alessandra Villa, Alan E. Mark, and Wilfred F. Van Gunsteren. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*, 25(13):1656–1676, October 2004.

[73] Christopher J. Fennell, Karisa L. Wymer, and David L. Mobley. A fixed-charge model for alcohol polarization in the condensed phase, and its role in small molecule hydration. *Journal of Physical Chemistry B*, 118(24):6438–6446, 2014.

[74] Joakim P. M. Jämbeck and Alexander P. Lyubartsev. Update to the General Amber Force Field for Small Solutes with an Emphasis on Free Energies of Hydration. *The Journal of Physical Chemistry B*, 118(14):3793–3804, 2014.

[75] Stefano Bosisio, Antonia S. J. S. Mey, and Julien Michel. Blinded predictions of host-guest standard free energies of binding in the SAMPL5 challenge. *Journal of Computer-Aided Molecular Design*, pages 1–10, August 2016.

[76] David B. Harwood, Cornelis J. Peters, and J. Ilja Siepmann. A Monte Carlo simulation study of the liquid–liquid equilibria for binary dodecane/ethanol and ternary dodecane/ethanol/water mixtures. *Fluid Phase Equilibria*, 407:269–279, January 2016.

[77] Andreas Klamt, Frank Eckert, Jens Reinisch, and Karin Wichmann. Prediction of cyclohexane-water distribution coefficients with COSMO-RS on the SAMPL5 data set. *Journal of Computer-Aided Molecular Design*, pages 1–9, July 2016.

[78] Ariën S. Rustenburg, Justin Dancer, Baiwei Lin, Jianwen A. Feng, Daniel F. Ortwine, David L. Mobley, and John D. Chodera. Measuring experimental cyclohexane-water distribution coefficients for the SAMPL5 challenge. *Journal of Computer-Aided Molecular Design*, 30(11):945–958, July 2016.

[79] Anthony Nicholls and Barry Honig. A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson–Boltzmann equation. *Journal of Computational Chemistry*, 12(4):435–445, May 1991.

[80] J. Andrew Grant, Barry T. Pickup, and Anthony Nicholls. A smooth permittivity function for Poisson–Boltzmann solvation methods. *Journal of Computational Chemistry*, 22(6):608–640, April 2001.

[81] Jennifer L. Knight and Charles L. Brooks. Surveying implicit solvent models for estimating small molecule absolute hydration free energies. *Journal of Computational Chemistry*, 32(13):2909–2923, October 2011.

[82] Pei-Kun Yang. Modifying Poisson equation for near-solute dielectric polarization and solvation free energy. *Chemical Physics*, 472:229–240, June 2016.

[83] J. Peter Guthrie. SAMPL4, a blind challenge for computational solvation free energies: The compounds considered. *Journal of Computer-Aided Molecular Design*, 28(3):151–168, April 2014.

[84] Michael R. Shirts, David L. Mobley, and John D. Chodera. Alchemical Free Energy Calculations: Ready for Prime Time? In D. C. Spellmeyer and R. Wheeler, editors, *Annual Reports in Computational Chemistry*, volume 3, pages 41–59. Elsevier, January 2007.

[85] Michael R. Shirts. Best Practices in Free Energy Calculations for Drug Design. In *Methods in Molecular Biology*, pages 425–467. Springer, 2012.

[86] R. E. Skyner, J. L. McDonagh, C. R. Groom, T. van Mourik, and J. B. O. Mitchell. A review of methods for the calculation of solution free energies and the modelling of systems in solution. *Physical Chemistry Chemical Physics*, 17(9):6174–6191, February 2015.

[87] David L Mobley, Christopher I Bayly, Matthew D Cooper, and Ken A Dill. Predictions of hydration free energies from all-atom molecular dynamics simulations. *The Journal of Physical Chemistry B*, 113(14):4533–4537, April 2009.

[88] Christopher J. Cramer and Donald G. Truhlar. A Universal Approach to Solvation Modeling. *Accounts of Chemical Research*, 41(6):760–768, June 2008.

[89] A. Z. Panagiotopoulos, N. Quirke, M. Stapleton, and D. J. Tildesley. Phase equilibria by simulation in the Gibbs ensemble. *Molecular Physics*, 63(4):527–545, 1988.

[90] Marcus G. Martin and J. Ilja Siepmann. Calculating Gibbs free energies of transfer from Gibbs ensemble Monte Carlo simulations. *Theoretical Chemistry Accounts*, 99(5):347–350, 1998.

[91] Matthew J. McGrath, I.-F. Will Kuo, Brice F. Ngouana W, Julius N. Ghogomu, Christopher J. Mundy, Aleksandr V. Marenich, Christopher J. Cramer, Donald G. Truhlar, and J. Ilja Siepmann. Calculation of the Gibbs free energy of solvation and dissociation of HCl in water via Monte Carlo simulations and continuum solvation models. *Physical chemistry chemical physics: PCCP*, 15(32):13578–13585, August 2013.

[92] Yow-Lin Huang, Svetlana Miroshnichenko, Hans Hasse, and Jadran Vrabec. Henry's Law Constant from Molecular Simulation: A Systematic Study of 95 Systems. *International Journal of Thermophysics*, 30(6):1791, 2009.

[93] Ling Zhang and J. Ilja. Siepmann. Direct calculation of Henry's law constants from Gibbs ensemble Monte Carlo simulations: Nitrogen, oxygen, carbon dioxide and methane in ethanol. *Theoretical Chemistry Accounts*, 115(5):391–397, 2006.

[94] Joseph T. Slusher. Estimation of infinite dilution activity coefficients in aqueous mixtures via molecular simulation. *Fluid Phase Equilibria*, 153(1):45–61, November 1998.

[95] Eirik F. da Silva. Use of free energy simulations to predict infinite dilution activity coefficients. *Fluid Phase Equilibria*, 221(1–2):15–24, July 2004.

[96] Chinghang Tong, Simon L. Clegg, and John H. Seinfeld. Comparison of activity coefficient models for atmospheric aerosols containing mixtures of electrolytes, organics, and water. *Atmospheric Environment*, 42(21):5459–5482, July 2008.

[97] Renan P. Gerber and Rafael de P. Soares. Prediction of Infinite-Dilution Activity Coefficients Using UNIFAC and COSMO-SAC Variants. *Industrial & Engineering Chemistry Research*, 49(16):7488–7496, 2010.

[98] Ming-Lan Ge, Xiao-Mei Deng, Li-Hui Zhang, Jin-Yuan Chen, Jie-Ming Xiong, and Wen-Hao Li. Activity coefficients at infinite dilution of organic solutes in the ionic liquid 1-butyl-3-methylimidazolium methyl sulfate. *The Journal of Chemical Thermodynamics*, 77:7–13, October 2014.

[99] Ming-Lan Ge, Chun-Yang Lu, Xiao-Yan Liu, Xiang-Bo Li, Jin-Yuan Chen, and Jie-Ming Xiong. Activity coefficients at infinite dilution of alkanes, alkenes, alkyl benzenes in dimethylphosphate based ionic liquids using gas–liquid chromatography. *The Journal of Chemical Thermodynamics*, 91:279–285, December 2015.

[100] Ming-Lan Ge, Qin Zhang, Sai-Nan Li, Yin-Juan Li, Xiu-Zhen Zhang, and Zhao Mu. Thermodynamics and activity coefficients at infinite dilution for organic solutes in the ionic liquid 1-hexyl-2,3-dimethylimidazolium bis(trifluoromethylsulfonyl)imide. *The Journal of Chemical Thermodynamics*, 102:303–309, November 2016.

[101] Michael Krummen, Detlef Gruber, and Jürgen Gmehling. Measurement of Activity Coefficients at Infinite Dilution in Solvent Mixtures Using the Dilutor Technique. *Industrial & Engineering Chemistry Research*, 39(6):2114–2123, 2000.

[102] Rodrigo Martínez, María Teresa Sanz, Sagrario Beltrán, and Elena Corcuera. Activity Coefficients at Infinite Dilution of Volatile Compounds in Water: Effect of Temperature and Salt Concentration. *Journal of Chemical & Engineering Data*, 57(5):1480–1485, 2012.

[103] Michael J. Schnieders, Jonas Baltrusaitis, Yue Shi, Gaurav Chattree, Lianqing Zheng, Wei Yang, and Pengyu Ren. The Structure, Thermodynamics, and Solubility of Organic Crystals from Simulation with a Polarizable Force Field. *Journal of Chemical Theory and Computation*, 8(5):1721–1736, May 2012.

[104] J. L. Aragones, E. Sanz, and C. Vega. Solubility of NaCl in Water by Molecular Simulation Revisited. *J. Chem. Phys.*, 136(24):244508, June 2012.

[105] A. L. Benavides, J. L. Aragones, and C. Vega. Consensus on the solubility of NaCl in water from computer simulations using the chemical potential route. *The Journal of Chemical Physics*, 144(12):124504, March 2016.

[106] Jan Westergren, Lennart Lindfors, Tobias Höglund, Kai Lüder, Sture Nordholm, and Roland Kjellander. In silico prediction of drug solubility: 1. Free energy of hydration. *The Journal of Physical Chemistry. B*, 111(7):1872–1882, February 2007.

[107] Kai Lüder, Lennart Lindfors, Jan Westergren, Sture Nordholm, and Roland Kjellander. In silico prediction of drug solubility: 2. Free energy of solvation in pure melts. *The Journal of Physical Chemistry. B*, 111(7):1883–1892, February 2007.

[108] Kai Lüder, Lennart Lindfors, Jan Westergren, Sture Nordholm, and Roland Kjellander. In silico prediction of drug solubility. 3. Free energy of solvation in pure amorphous matter. *The Journal of Physical Chemistry. B*, 111(25):7303–7311, June 2007.

[109] Kai Lüder, Lennart Lindfors, Jan Westergren, Sture Nordholm, Rasmus Persson, and Mikaela Pedersen. In silico prediction of drug solubility: 4. Will simple potentials suffice? *Journal of Computational Chemistry*, 30(12):1859–1871, September 2009.

[110] Andrew S. Paluch, Sreeja Parameswaran, Shuai Liu, Anasuya Kolavennu, and David L. Mobley. Predicting the excess solubility of acetanilide, acetaminophen, phenacetin, benzocaine, and caffeine in binary water/ethanol mixtures via molecular simulation. *The Journal of Chemical Physics*, 142(4):044508, January 2015.

[111] Andrew L. Ferguson, Pablo G. Debenedetti, and Athanassios Z. Panagiotopoulos. Solubility and Molecular Conformations of n-Alkane Chains in Water. *J. Phys. Chem. B*, 113(18):6405–6414, 2009.

[112] Richard Olsen, Bjorn Kvamme, and Tatiana Kuznetsova. Free energy of solvation and Henry's law solubility constants for mono-, di- and tri-ethylene glycol in water and methane. *Fluid Phase Equilibria*, 418:152–159, 2016.

[113] Javad Noroozi, Cyrus Ghotbi, Jaber Jahanbin Sardroodi, Javad Karimi-Sabet, and Marc A. Robert. Solvation free energy and solubility of acetaminophen and ibuprofen in supercritical carbon dioxide: Impact of the solvent model. *The Journal of Supercritical Fluids*, 109:166–176, March 2016.

[114] Samuel Genheden and Jonathan W. Essex. All-atom/coarse-grained hybrid predictions of distribution coefficients in SAMPL5. *Journal of Computer-Aided Molecular Design*, July 2016.

[115] Kee-Choo Chung and Hwangseo Park. Extended solvent-contact model approach to blind SAMPL5 prediction challenge for the distribution coefficients of drug-like molecules. *Journal of Computer-Aided Molecular Design*, July 2016.

[116] Nicholas Bodor and Peter Buchwald. Mechanism of Drug Action: Basic Concepts. In *Retrometabolic Drug Design and Targeting*, pages 9–38. John Wiley & Sons, Inc., 2012.

[117] John D. Chodera, David L. Mobley, Michael R. Shirts, Richard W. Dixon, Kim Branson, and Vijay S. Pande. Alchemical free energy methods for drug discovery: Progress and challenges. *Current Opinion in Structural Biology*, 21(2):150–160, 2011.

[118] Yuqing Deng and Benoît Roux. Computations of Standard Binding Free Energies with Molecular Dynamics Simulations. *The journal of physical chemistry. B*, 113(8), February 2009.

[119] Julien Michel and Jonathan W. Essex. Prediction of protein-ligand binding affinity by free energy simulations: Assumptions, pitfalls and expectations. *Journal of Computer-Aided Molecular Design*, 24(8):639–658, August 2010.

[120] Ayesha Zafar and Jóhannes Reynisson. Hydration Free Energy as a Molecular Descriptor in Drug Design: A Feasibility Study. *Molecular Informatics*, 35(5):207–214, May 2016.

[121] Nandou Lu, Jayant K. Singh, and David A. Kofke. Appropriate methods to combine forward and reverse free-energy perturbation averages. *The Journal of Chemical Physics*, 118(7):2977–2984, February 2003.

[122] Di Wu and David A. Kofke. Phase-space overlap measures. II. Design and implementation of staging methods for free-energy calculations. *Journal of Chemical Physics*, 123(8), 2005.

[123] Di Wu and David A. Kofke. Asymmetric bias in free-energy perturbation measurements using two Hamiltonian-based models. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, 70(6 Pt 2):066702, December 2004.

[124] Di Wu and David A. Kofke. Phase-space overlap measures. I. Fail-safe bias detection in free energies calculated by molecular simulation. *Journal of Chemical Physics*, 123(5), 2005.

[125] Alan M. Ferrenberg and Robert H. Swendsen. Optimized Monte Carlo data analysis. *Physical Review Letters*, 63(12):1195–1198, September 1989.

[126] Shankar Kumar, John M. Rosenberg, Djamal Bouzida, Robert H. Swendsen, and Peter A. Kollman. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry*, 13(8):1011–1021, 1992.

[127] Anita de Ruiter, Stefan Boresch, and Chris Oostenbrink. Comparison of thermodynamic integration and Bennett acceptance ratio for calculating relative protein-ligand binding free energies. *Journal of Computational Chemistry*, 34(12):1024–1034, May 2013.

[128] Pavel V. Klimovich, Michael R. Shirts, and David L. Mobley. Guidelines for the analysis of free energy calculations. *Journal of Computer-Aided Molecular Design*, 29(5):397–411, May 2015.

[129] Junmei Wang, Tingjun Hou, and Xiaojie Xu. Recent Advances in Free Energy Calculations with a Combination of Molecular Mechanics and Continuum Models. *Current Computer - Aided Drug Design*, 2(3):287–306, 2006.

[130] Thomas Steinbrecher, David L. Mobley, and David A. Case. Nonlinear scaling schemes for Lennard-Jones interactions in free energy calculations. *The Journal of Chemical Physics*, 127(21):214108, December 2007.

[131] Levi N. Naden and Michael R. Shirts. Linear Basis Function Approach to Efficient Alchemical Free Energy Calculations. 2. Inserting and Deleting Particles with Coulombic Interactions. *Journal of Chemical Theory and Computation*, 11(6):2536–2549, June 2015.

[132] Tri T. Pham and Michael R. Shirts. Identifying low variance pathways for free energy calculations of molecular transformations in solution phase. *The Journal of Chemical Physics*, 135(3):034114, July 2011.

[133] Vytautas Gapsys, Daniel Seeliger, and Bert L. de Groot. New Soft-Core Potential Function for Molecular Dynamics Based Alchemical Free Energy Calculations. *Journal of Chemical Theory and Computation*, 8(7):2373–2382, July 2012.

[134] Tri T. Pham and Michael R. Shirts. Optimal pairwise and non-pairwise alchemical pathways for free energy calculations of molecular transformation in solution phase. *The Journal of Chemical Physics*, 136(12):124120, March 2012.

[135] Levi N. Naden, Tri T. Pham, and Michael R. Shirts. Linear basis function approach to efficient alchemical free energy calculations. 1. Removal of uncharged atomic sites. *Journal of Chemical Theory and Computation*, 10(3):1128–1149, 2014.

[136] Clara Christ and Thomas Fox. Accuracy assessment and automation of free energy calculations for drug design. *Journal of chemical information and modeling*, 2013.

[137] Niels Hansen and Wilfred F van Gunsteren. Practical Aspects of Free-Energy Calculations: A Review. *J. Chem. Theory Comput.*, 10(7):2632–2647, July 2014.

[138] Bao Wang and G. W. Wei. Parameter optimization in differential geometry based solvation models. *The Journal of Chemical Physics*, 143(13):134119, October 2015.

[139] Ian J. Nessler, Jacob M. Litman, and Michael J. Schnieders. Toward polarizable AMOEBA thermodynamics at fixed charge efficiency using a dual force field approach: Application to organic crystals. *Physical Chemistry Chemical Physics*, 18(44):30313–30322, 2016.

[140] Katarzyna B. Koziara, Martin Stroet, Alpeshkumar K. Malde, and Alan E. Mark. Testing and validation of the Automated Topology Builder (ATB) version 2.0: Prediction of hydration free enthalpies. *Journal of Computer-Aided Molecular Design*, 28(3):221–233, 2014.

[141] Stefan Canzar, Mohammed El-Kebir, René Pool, Khaled Elbassioni, Alpeshkumar K. Malde, Alan E. Mark, Daan P. Geerke, Leen Stougie, and Gunnar W. Klau. Charge Group Partitioning in Biomolecular Simulation. *Journal of Computational Biology*, 20(3):188–198, March 2013.

[142] Alpeshkumar K. Malde, Le Zuo, Matthew Breeze, Martin Stroet, David Poger, Pramod C. Nair, Chris Oostenbrink, and Alan E. Mark. An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *Journal of Chemical Theory and Computation*, 7(12):4026–4037, 2011.

[143] David L. Mobley, Christopher I. Bayly, Matthew D. Cooper, Michael R. Shirts, and Ken A. Dill. Small Molecule Hydration Free Energies in Explicit Solvent: An Extensive Test of Fixed-Charge Atomistic Simulations. *Journal of Chemical Theory and Computation*, 5(2):350–358, 2009.

[144] Sreeja Parameswaran and David L Mobley. Box size effects are negligible for solvation free energies of neutral solutes. *Journal of Computer-Aided Molecular Design*, 28(8):825–829, August 2014.

[145] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91(1):43–56, September 1995.

[146] Erik Lindahl, Berk Hess, and David van der Spoel. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Model.*, 7(8):306–317, 2001.

[147] David Van Der Spoel, Erik Lindahl, Berk Hess, Gerrit Groenhof, Alan E. Mark, and Herman J. C. Berendsen. GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, 26(16):1701–1718, December 2005.

[148] Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. GROMACS 4: Algorithms for Highly Efficient, Load-balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.*, 4(3):435–447, March 2008.

[149] Sander Pronk, Szilárd Páll, Roland Schulz, Per Larsson, Pär Bjelkmar, Rossen Apostolov, Michael R. Shirts, Jeremy C. Smith, Peter M. Kasson, David van der Spoel, Berk Hess, and Erik Lindahl. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29(7):845–854, April 2013.

[150] Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C Smith, Berk Hess, and Erik Lindahl. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2:19–25, September 2015.

[151] Michael R. Shirts, David L. Mobley, John D. Chodera, and Vijay S. Pande. Accurate and Efficient Corrections for Missing Dispersion Interactions in Molecular Simulations. *The Journal of Physical Chemistry B*, 111(45):13052–13063, November 2007.

[152] M. Parrinello and A. Rahman. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Physical Review Letters*, 45(14):1196–1199, October 1980.

[153] Niel M. Henriksen, Andrew T. Fenley, and Michael K. Gilson. Computational Calorimetry: High-Precision Calculation of Host–Guest Binding Thermodynamics. *Journal of Chemical Theory and Computation*, 11(9):4377–4394, September 2015.

[154] Inc. OpenEye Scientific Software. OEChem, 2010.

[155] Inc. OpeneEye Scientific Software. QUACPAC 1.7.0.2. OpenEye Scientific Software.

[156] Paul C D Hawkins and Anthony Nicholls. Conformer generation with OMEGA: Learning from the data set and the analysis of failures. *Journal of Chemical Information and Modeling*, 52(11):2919–2936, November 2012.

[157] Araz Jakalian, Bruce L. Bush, David B. Jack, and Christopher I. Bayly. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *Journal of Computational Chemistry*, 21(2):132–146, January 2000.

[158] Araz Jakalian, David B. Jack, and Christopher I. Bayly. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *Journal of Computational Chemistry*, 23(16):1623–1641, December 2002.

[159] William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, July 1983.

[160] Emilio Gallicchio and Ronald M Levy. AGBNP: An analytic implicit solvent model suitable for molecular dynamics simulations and high-resolution modeling. *Journal of computational chemistry*, 25(4):479–499, 2004.

[161] Emilio Gallicchio, Kristina Paris, and Ronald M. Levy. The AGBNP2 Implicit Solvation Model. *Journal of Chemical Theory and Computation*, 5(9):2544–2564, September 2009.

[162] Stefan Boresch, Georgios Archontis, and Martin Karplus. Free energy simulations: The meaning of the individual contributions from a component analysis. *Proteins: Structure, Function, and Bioinformatics*, 20(1):25–33, 1994.

[163] M. Zacharias and T. P. Straatsma. Path Dependence of Free Energy Components in Thermodynamic Integration. *Molecular Simulation*, 14(6):417–423, June 1995.

[164] Ken A. Dill. Additivity Principles in Biochemistry. *Journal of Biological Chemistry*, 272(2):701–704, 1997.

[165] R M Levy and E Gallicchio. Computer simulations with explicit solvent: Recent progress in the thermodynamic decomposition of free energies and in modeling electrostatic effects. *Annual review of physical chemistry*, 49:531–567, 1998.

[166] Berk Hess and Nico F. A. van der Vegt. Hydration Thermodynamic Properties of Amino Acid Analogues: A Systematic Comparison of Biomolecular Force Fields and Water Models. *The Journal of Physical Chemistry B*, 110(35):17616–17626, September 2006.

[167] A. T. Fenley, H. S. Muddana, and M. K. Gilson. Entropy-enthalpy transduction caused by conformational shifts can obscure the forces driving protein-ligand binding. *Proceedings of the National Academy of Sciences*, 109(49):20006–20011, December 2012.

[168] Hans W. Horn, William C. Swope, Jed W. Pitera, Jeffry D. Madura, Thomas Dick, Greg L. Hura, and Teresa Head-Gordon. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *Journal of Chemical Physics*, 120(20):9665–9678, 2004.

[169] N. V. Plyasunova, A. V. Plyasunov, and E. L. Shock. Database of Thermodynamic Properties for Aqueous Organic Compounds. *International Journal of Thermophysics*, 25(2):351–360, 2004.

[170] Arieh Ben-Naim. *Molecular Theory of Solutions*. Oxford University Press, Oxford, 2006.

[171] Charles A. Eckert and Steven R. Sherman. Measurement and prediction of limiting activity coefficients. *Fluid Phase Equilibria*, 116(1):333–342, March 1996.

[172] Ariel A. Chialvo. Determination of excess Gibbs free energy by the single-charging-integral approach. Infinite dilution activity coefficients and related quantities. *The Journal of Physical Chemistry*, 95(17):6683–6687, August 1991.

[173] John M. Prausnitz, Ruediger N. Lichtenthaler, and Edmundo Gomes de Azevedo. *Molecular Thermodynamics of Fluid-Phase Equilibria*. Prentice Hall, 2nd edition, 1986.

[174] Yoshimori Miyano, Arisa Kimura, Mikiko Kuroda, Asami Matsushita, Ai Yamasaki, Yusuke Yamaguchi, Akihiro Yoshizawa, and Yoko Tateishi. Henry's Law Constants and Infinite Dilution Activity Coefficients of Propane, Propene, Butane, 2-Methylpropane, 1-Butene, 2-Methylpropene, trans-2-Butene, cis-2-Butene, 1,3-Butadiene, Dimethyl Ether, Chloroethane, and 1,1-Difluoroethane in Benzene, Toluene, o-Xylene, m-Xylene, p-Xylene, and Styrene. *Journal of Chemical & Engineering Data*, 52(1):291–297, January 2007.

[175] Kazuo Kojima, Suojiang Zhang, and Toshihiko Hiaki. Measuring methods of infinite dilution activity coefficients and a database for systems including water. *Fluid Phase Equilibria*, 131(1):145–179, May 1997.

[176] J. David Raal. Characterization of differential ebulliometers for measuring activity coefficients. *AIChE Journal*, 46(1):210–220, January 2000.

[177] Michael Diedenhofen, Frank Eckert, and Andreas Klamt. Prediction of Infinite Dilution Activity Coefficients of Organic Compounds in Ionic Liquids Using COSMO-RS. *Journal of Chemical & Engineering Data*, 48(3):475–479, March 2003.

[178] Eirik F. da Silva. Erratum to "Use of free energy simulations to predict infinite dilution activity coefficients" [Fluid Phase Equilb. 221 (2004) 15]. *Fluid Phase Equilibria*, 231(2):252–253, April 2005.

[179] Christian S. Schacht, Lawien Zubeir, Theo W. de Loos, and Joachim Gross. Application of Infinite Dilution Activity Coefficients for Determining Binary Equation of State Parameters. *Industrial & Engineering Chemistry Research*, 49(16):7646–7653, August 2010.

[180] Pratik Dhakal, Sydnee N. Roese, Erin M. Stalcup, and Andrew S. Paluch. Application of MOSCED To Predict Limiting Activity Coefficients, Hydration Free Energies,

Henry's Constants, Octanol/Water Partition Coefficients, and Isobaric Azeotropic Vapor–Liquid Equilibrium. *Journal of Chemical & Engineering Data*, 63(2):352–364, February 2018.

[181] Fernando M. Lisboa and Josefredo R. Pliego. Infinite dilution activity coefficient from SMD calculations: Accuracy and performance for predicting liquid-liquid equilibria. *Journal of Molecular Modeling*, 24(3):56, February 2018.

[182] Hesam Ahmadian Behrooz and R. Bozorgmehry Boozarjomehry. Prediction of limiting activity coefficients for binary vapor-liquid equilibrium using neural networks. *Fluid Phase Equilibria*, 433:174–183, February 2017.

[183] Kourosh Tabar Heydar, Halime Gharehmoshk Gharavi, Mina Nazifi, Mojtaba Mirzaei, and Ali Sharifi. Using binary mixtures of dicationic ionic liquids for determination of activity coefficients at infinite dilution by gas–liquid chromatography. *Fluid Phase Equilibria*, 353:93–100, September 2013.

[184] Yingjie Xu, Wu Qian, Qiongqiong Gao, and Haoran Li. Prediction of vapor–liquid equilibria of alcohol+hydrocarbon systems by 1H NMR spectroscopy. *Chemical Engineering Science*, 74:211–218, May 2012.

[185] Ernesto Estrada, Gerardo A. Díaz, and Eduardo J. Delgado. Predicting infinite dilution activity coefficients of organic compounds in water by quantum-connectivity descriptors. *Journal of Computer-Aided Molecular Design*, 20(9):539–548, September 2006.

[186] Ekaterina L. Ratkova, David S. Palmer, and Maxim V. Fedorov. Solvation Thermodynamics of Organic Molecules by the Molecular Integral Equation Theory: Approaching Chemical Accuracy. *Chemical Reviews*, 115(13):6312–6356, July 2015.

[187] Guilherme Duarte Ramos Matos, Daisy Y. Kyu, Hannes H. Loeffler, John D. Chodera, Michael R. Shirts, and David L. Mobley. Approaches for Calculating Solvation Free Energies and Enthalpies Demonstrated with an Update of the FreeSolv Database. *Journal of Chemical & Engineering Data*, 62(5):1559–1569, May 2017.

[188] Martin Karplus and J. Andrew McCammon. Molecular dynamics simulations of biomolecules. *Nature Structural & Molecular Biology*, 9(9):646–652, September 2002.

[189] Dian Jiao, Pavel A. Golubkov, Thomas A. Darden, and Pengyu Ren. Calculation of protein–ligand binding free energy by using a polarizable potential. *Proceedings of the National Academy of Sciences*, 105(17):6290–6295, April 2008.

[190] Sushil K. Mishra, Gaetano Calabró, Hannes H. Loeffler, Julien Michel, and Jaroslav Koča. Evaluation of Selected Classical Force Fields for Alchemical Binding Free Energy Calculations of Protein-Carbohydrate Complexes. *J. Chem. Theory Comput.*, 11(7):3333–3345, July 2015.

[191] Lingle Wang, Yujie Wu, Yuqing Deng, Byungchan Kim, Levi Pierce, Goran Krilov, Dmitry Lupyan, Shaughnessy Robinson, Markus K. Dahlgren, Jeremy Greenwood, Donna L. Romero, Craig Masse, Jennifer L. Knight, Thomas Steinbrecher, Thijs Beuming, Wolfgang Damm, Ed Harder, Woody Sherman, Mark Brewer, Ron Wester, Mark Murcko, Leah Frye, Ramy Farid, Teng Lin, David L. Mobley, William L. Jorgensen, Bruce J. Berne, Richard A. Friesner, and Robert Abel. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *Journal of the American Chemical Society*, 137(7):2695–2703, February 2015.

[192] David L. Mobley and Michael K. Gilson. Predicting binding free energies: Frontiers and benchmarks. *Annual review of biophysics*, 46:531–558, May 2017.

[193] Clara D Christ, Alan E Mark, and Wilfred F van Gunsteren. Feature Article Basic Ingredients of Free Energy Calculations: A Review. *Journal of Computational Chemistry*, 31(8):1569–1582, June 2010.

[194] Andrew Pohorille, Christopher Jarzynski, and Christophe Chipot. Good Practices in Free-Energy Calculations. *The Journal of Physical Chemistry B*, 114(32):10235–10253, August 2010.

[195] Noor Asidah Mohamed, Richard T. Bradshaw, and Jonathan W. Essex. Evaluation of solvation free energies for small molecules with the AMOEBA polarizable force field. *Journal of Computational Chemistry*, pages n/a–n/a, October 2016.

[196] Jian Yin, Niel M. Henriksen, David R. Slochower, Michael R. Shirts, Michael W. Chiu, David L. Mobley, and Michael K. Gilson. Overview of the SAMPL5 host–guest challenge: Are we doing better? *Journal of Computer-Aided Molecular Design*, 31(1):1–19, 2017.

[197] Jian Yin, Niel M. Henriksen, David R. Slochower, and Michael K. Gilson. The SAMPL5 Host-Guest Challenge: Computing Binding Free Energies and Enthalpies from Explicit Solvent Simulations by the Attach-Pull-Release (APR) Method. *Journal of Computer-Aided Molecular Design*, 31(1):133–145, 2017.

[198] Michael Frenkel, Robert D. Chirico, Vladimir V. Diky, Qian Dong, Svetlana Frenkel, Paul R. Franchois, Dale L. Embry, Thomas L. Teague, Kenneth N. Marsh, and Randolph C. Wilhoit. ThermoML: An XML-Based Approach for Storage and Exchange of Experimental and Critically Evaluated Thermophysical and Thermochemical Property Data. 1. Experimental Data. *Journal of Chemical & Engineering Data*, 48(1):2–13, January 2003.

[199] Michael D. Frenkel, Robert D. Chirico, Vladimir Diky, Qian Dong, Kenneth N. Marsh, John H. Dymond, William A. Wakeham, Stephen E. Stein, Erich Koenigsberger, and Anthony R. Goodwin. XML-based IUPAC Standard for Experimental, Predicted, and Critically Evaluated Thermodymic property data storage and Capture (ThermoML): IUPAC Recommendations 2005. *Pure and Applied Chemistry*, 78(3), May 2006.

[200] Michael Frenkel, Robert D. Chirico, Vladimir Diky, Paul L. Brown, John H. Dymond, Robert N. Goldberg, Anthony R. H. Goodwin, Heiko Heerklotz, Erich Königsberger, John E. Ladbury, Kenneth N. Marsh, David P. Remeta, Stephen E. Stein, William A. Wakeham, and Peter A. Williams. Extension of ThermoML: The IUPAC standard for thermodynamic data communications (IUPAC Recommendations 2011). *Pure and Applied Chemistry*, 83(10):1937–1969, September 2011.

[201] Kyle A. Beauchamp, Julie M. Behr, Ariën S. Rustenburg, Christopher I. Bayly, Kenneth Kroenlein, and John D. Chodera. Towards Automated Benchmarking of Atomistic Forcefields: Neat Liquid Densities and Static Dielectric Constants from the ThermoML Data Archive. *The journal of physical chemistry. B*, 119(40):12912–12920, October 2015.

[202] Inc. OpeneEye Scientific Software. OEChem Toolkit, 2010. (accessed June 16, 2015).

[203] Andrea Rizzi, Patrick B. Grinaway, Daniel L. Parton, Michael R. Shirts, Kai Wang, Peter M. Eastman, Mark S. Friedrichs, Vijay S. Pande, Kim Branson, David L. Mobley, and John D. Chodera. YANK: A GPU-accelerated platform for alchemical free energy calculations.

[204] L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry*, 30(13):2157–2164, October 2009.

[205] Jason Swails, Carlos Hernandez, David L Mobley, Hai Nguyen, Lee-Ping Wang, and Pawel Janowski. Parmed, (accessed October 9, 2015).

[206] Peter Eastman and Vijay S. Pande. Efficient nonbonded interactions for molecular dynamics on a graphics processing unit. *Journal of Computational Chemistry*, 31(6):1268–1272, 2009.

[207] P. Eastman and V. Pande. OpenMM: A Hardware-Independent Framework for Molecular Simulations. *Computing in Science Engineering*, 12(4):34–39, July 2010.

[208] Peter Eastman, Mark S. Friedrichs, John D. Chodera, Randall J. Radmer, Christopher M. Bruns, Joy P. Ku, Kyle A. Beauchamp, Thomas J. Lane, Lee-Ping Wang, Diwakar Shukla, Tony Tye, Mike Houston, Timo Stich, Christoph Klein, Michael R. Shirts, and Vijay S. Pande. OpenMM 4: A Reusable, Extensible, Hardware Independent Library for High Performance Molecular Simulation. *Journal of Chemical Theory and Computation*, 9(1):461–469, January 2013.

[209] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh Ewald: An N log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12):10089–10092, June 1993.

[210] Ulrich Essmann, Lalith Perera, Max L. Berkowitz, Tom Darden, Hsing Lee, and Lee G. Pedersen. A smooth particle mesh Ewald method. *The Journal of Chemical Physics*, 103(19):8577–8593, November 1995.

[211] Jennifer L Knight, Joseph D Yesselman, and Charles L Brooks III. Assessing the quality of absolute hydration free energies among CHARMM-compatible ligand parameterization schemes. *Journal of Computational Chemistry*, 34(11):893–903, January 2013.

[212] Mark S. Friedrichs, Peter Eastman, Vishal Vaidyanathan, Mike Houston, Scott Legrand, Adam L. Beberg, Daniel L. Ensign, Christopher M. Bruns, and Vijay S. Pande. Accelerating molecular dynamic simulation on graphics processing units. *Journal of Computational Chemistry*, 30(6):864–872, 2008.

[213] Emilio Gallicchio and Ronald M. Levy. Recent theoretical and computational advances for modeling protein–ligand binding affinities. In Christo Christov, editor, *Computational Chemistry Methods in Structural Biology*, volume 85 of *Advances in Protein Chemistry and Structural Biology*, pages 27–80. Academic Press, 2011.

[214] Xiya Lu, Dong Fang, Shingo Ito, Yuko Okamoto, Victor Ovchinnikov, and Qiang Cui. QM/MM free energy simulations: Recent progress and challenges. *Mol. Sim.*, 42(13):1056–1078, 2016.

[215] J. M. Rickman and R. LeSar. Free-Energy Calculations in Materials Research. *Ann. Rev. Mat. Res.*, 32(1):195–217, 2002.

[216] Shunzhou Wan, Bernhard Knapp, David W. Wright, Charlotte M. Deane, and Peter V. Coveney. Rapid, Precise, and Reproducible Prediction of Peptide–MHC Binding Affinities from Molecular Dynamics That Correlate Well with Experiment. *J. Chem. Theory Comput.*, 11(7):3346–3356, 2015.

[217] D. L. Beveridge and F. M. Dicapua. Free Energy Via Molecular Simulation: Applications to Chemical and Biomolecular Systems. *Annu. Rev. Biophys. Biophys. Chem.*, 18(1):431–492, 1989.

[218] T. P. Straatsma and J. A. McCammon. Computational Alchemy. *Ann. Rev. Phys. Chem.*, 43(1):407–435, 1992.

[219] Peter. Kollman. Free energy calculations: Applications to chemical and biochemical phenomena. *Chemical Reviews*, 93(7):2395–2417, November 1993.

[220] David R. Squire and William G. Hoover. Monte Carlo Simulation of Vacancies in Rare–Gas Crystals. *J. Chem. Phys.*, 50(2):701–706, 1969.

[221] Stefan Boresch, Franz Tettinger, Martin Leitgeb, and Martin Karplus. Absolute Binding Free Energies: A Quantitative Approach for Their Calculation. *J. Chem. Phys. B*, 107(35):9535–9551, 2003.

[222] Vytautas Gapsys, Servaas Michielssens, Daniel Seeliger, and Bert L. de Groot. Pmx: Automated protein structure and topology generation for alchemical perturbations. *J. Comput. Chem.*, 36(5):348–354, 2015.

[223] Robert B. Sandberg, Martina Banchelli, Carlo Guardiani, Stefano Menichetti, Gabriella Caminati, and Piero Procacci. Efficient Nonequilibrium Method for Binding Free Energy Calculations in Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation*, 11(2):423–435, February 2015.

[224] William L Jorgensen, J Kathleen Buckner, Stephane Boudon, and Julian Tirado-Rives. Efficient computation of absolute free energies of binding by computer simulations. Application to the methane dimer in water. *J. Chem. Phys.*, 89(6):3742–3746, 1988.

[225] Stefan Boresch and Martin Karplus. The Role of Bonded Terms in Free Energy Simulations. 2. Calculation of Their Influence on Free Energy Differences of Solvation. *J. Phys. Chem. A*, 103(1):119–136, 1999.

[226] Christopher J. Woods, Maturos Malaisree, Supot Hannongbua, and Adrian J. Mulholland. A water–swap reaction coordinate for the calculation of absolute protein–ligand binding free energies. *J. Chem. Phys.*, 134(5):054114, 2011.

[227] Christopher J. Woods, Maturos Malaisree, Julien Michel, Ben Long, Simon McIntosh-Smith, and Adrian J. Mulholland. Rapid decomposition and visualisation of protein–ligand binding free energies by residue and by water. *Faraday Discuss.*, 169(0):477–499, 2014.

[228] Paulius Mikulskis, Samuel Genheden, and Ulf Ryde. A Large-Scale Test of Free-Energy Simulation Estimates of Protein–Ligand Binding Affinities. *Journal of Chemical Information and Modeling*, 54(10):2794–2806, 2014.

[229] Lingle Wang, Yuqing Deng, Yujie Wu, Byungchan Kim, David N. LeBard, Dan Wandschneider, Mike Beachy, Richard A. Friesner, and Robert Abel. Accurate Modeling of Scaffold Hopping Transformations in Drug Discovery. *J. Chem. Theory Comput.*, 13(1):42–54, 2017.

[230] David A. Pearlman. A Comparison of Alternative Approaches to Free Energy Calculations. *J. Phys. Chem.*, 98(5):1487–1493, 1994.

[231] Julien Michel, Marcel L. Verdonk, and Jonathan W. Essex. Protein–Ligand Complexes: Computation of the Relative Free Energy of Different Scaffolds and Binding Modes. *J. Chem. Theory Comput.*, 3(5):1645–1655, 2007.

[232] Gabriel J Rocklin, David L Mobley, Ken a Dill, and Philippe H Hünenberger. Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: An accurate correction scheme for electrostatic finite-size effects. *The Journal of chemical physics*, 139(18):184103, November 2013.

[233] Paul H. Axelsen and Daohui Li. Improved convergence in dual-topology free energy calculations through use of harmonic restraints. *J. Comput. Chem.*, 19(11):1278–1283, 1998.

[234] Shunzhou Wan, Agastya P. Bhati, Stefan J. Zasada, Ian Wall, Darren Green, Paul Bamborough, and Peter V. Coveney. Rapid and Reliable Binding Affinity Prediction of Bromodomain Inhibitors: A Computational Study. *J. Chem. Theory Comput.*, 13(2):784–795, 2017.

[235] S. Shobana, Benoît Roux, and Olaf S. Andersen. Free Energy Simulations: Thermodynamic Reversibility and Variability. *J. Phys. Chem. B*, 104(21):5179–5190, 2000.

[236] Stefan Boresch. The Role of Bonded Energy Terms in Free Energy Simulations - Insights from Analytical Results. *Molecular Simulation*, 28(1-2):13–37, January 2002.

[237] Shuai Liu, Lingle Wang, and David L. Mobley. Is Ring Breaking Feasible in Relative Binding Free Energy Calculations? *Journal of Chemical Information and Modeling*, 55(4):727–735, April 2015.

[238] David A Case, Thomas E Cheatham III, Tom Darden, Holger Gohlke, Ray Luo, Kenneth M Merz Jr., Alexey Onufriev, Carlos Simmerling, Bing Wang, and Robert J Woods. The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, 26(16):1668–1688, December 2005.

[239] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. CHARMM: The biomolecular simulation program. *J. Comput. Chem.*, 30(10):1545–1614, 2009.

[240] Walter R. P. Scott, Philippe H. Hünenberger, Ilario G. Tironi, Alan E. Mark, Salomon R. Billeter, Jens Fennen, Andrew E. Torda, Thomas Huber, Peter Krüger, and Wilfred F. van Gunsteren. The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A*, 103(19):3596–3607, 1999.

[241] Christopher Woods, Antonia SJ Mey, Gaetano Calabro, and Julien Michel. *Sire Molecular Simulations Framework*. 2016.

[242] Michael R. Shirts, Christoph Klein, Jason M. Swails, Jian Yin, Michael K. Gilson, David L. Mobley, David A. Case, and Ellen D. Zhong. Lessons learned from comparing molecular dynamics engines on the SAMPL5 dataset. *Journal of Computer-Aided Molecular Design*, 31(1):147–161, January 2017.

[243] Michael Schappals, Andreas Mecklenfeld, Leif Kröger, Vitalie Botan, Andreas Köster, Simon Stephan, Edder J. Garćia, Gabor Rutkai, Gabriele Raabe, Peter Klein, Kai Leonhard, Colin W. Glass, Johannes Lenhard, Jadran Vrabec, and Hans Hasse. Round Robin Study: Molecular Simulation of Thermodynamic Properties from Models with Internal Degrees of Freedom. *J. Chem. Theory Comput.*, 13(9):4270–4280, 2017.

[244] Sander Pronk, Per Larsson, Iman Pouya, Gregory R. Bowman, Imran S. Haque, Kyle Beauchamp, Berk Hess, Vijay S. Pande, Peter M. Kasson, and Erik Lindahl. Copernicus: A New Paradigm for Parallel Adaptive Molecular Dynamics. In *Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '11, pages 60:1–60:10, Seattle, Washington, 2011. ACM.

[245] S. Kashif Sadiq, David Wright, Simon J. Watson, Stefan J. Zasada, Ileana Stoica, and Peter V. Coveney. Automated Molecular Simulation Based Binding Affinity Calculator for Ligand-Bound HIV-1 Proteases. *J. Chem. Inf. Model*, 48(9):1909–1919, 2008.

[246] Magnus Lundborg and Erik Lindahl. Automatic GROMACS Topology Generation and Comparisons of Force Fields for Solvation Free Energy Calculations. *J. Phys. Chem. B*, 119(3):810–823, 2015.

[247] Hannes H. Loeffler, Julien Michel, and Christopher Woods. FESetup: Automating Setup for Alchemical Free Energy Simulations. *Journal of Chemical Information and Modeling*, 55(12):2485–2490, December 2015.

[248] Vivekanandan Balasubramanian, Iain Bethune, Ardita Shkurti, Elena Breitmoser, Eugen Hruska, Cecilia Clementi, Charles A. Laughton, and Shantenu Jha. *ExTASY: Scalable and Flexible Coupling of MD Simulations and Advanced Sampling Techniques*. 2016.

[249] T Simonson. Free-energy of particle insertion - An exact analysis of the origin singularity for simple liquids. *Mol. Phys.*, 80(2):{441–447}, October 1993.

[250] James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kalé, and Klaus Schulten. Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, 26(16):1781–1802, 2005.

[251] Stefan Boresch and Martin Karplus. The Role of Bonded Terms in Free Energy Simulations: 1. Theoretical Analysis. *J. Phys. Chem. A*, 103(1):103–118, 1999.

[252] Hannes H. Loeffler, Christoph A. Sotriffer, Rudolf H. Winger, Klaus R. Liedl, and Bernd M. Rode. Calculation of sequence–dependent free energies of hydration of dipeptides formed by alanine and glycine. *J. Comput. Chem.*, 22(8):846–860, 2001.

[253] Devleena Shivakumar, Edward Harder, Wolfgang Damm, Richard A. Friesner, and Woody Sherman. Improving the Prediction of Absolute Solvation Free Energies Using the Next Generation OPLS Force Field. *Journal of Chemical Theory and Computation*, 8(8):2553–2558, August 2012.

[254] Y. Deng and B. Roux. Hydration of Amino Acid Side Chains: Non-Polar and Electrostatic Contributions Calculated from Staged Molecular Dynamics Free Energy Simulations with Explicit Water Molecules. *J. Phys. Chem.*, 108:16567–16576, 2004.

[255] Jed W. Pitera and Wilfred F. van Gunsteren. A Comparison of Non-Bonded Scaling Approaches for Free Energy Calculations. *Molecular Simulation*, 28(1-2):45–65, January 2002.

[256] Jamshed Anwar and David M. Heyes. Robust and accurate method for free-energy calculation of charged molecular systems. *The Journal of Chemical Physics*, 122(22):224117, June 2005.

[257] Hans C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics*, 72(4):2384, 1980.

[258] J.A. Barker and R.O. Watts. Monte Carlo studies of the dielectric properties of water-like models. *Mol. Phys.*, 26(3):789–792, 1973.

[259] Stefano Bosisio, Antonia S. J. S. Mey, and Julien Michel. Blinded predictions of distribution coefficients in the SAMPL5 challenge. *J. Comput. Aided Mol. Des.*, 30(11):1101–1114, 2016.

[260] John D. Chodera, William C. Swope, Jed W. Pitera, Chaok Seok, and Ken A. Dill. Use of the Weighted Histogram Analysis Method for the Analysis of Simulated and Parallel Tempering Simulations. *Journal of Chemical Theory and Computation*, 3(1):26–41, 2007.

[261] Lee-Ping Wang, Todd J. Martinez, and Vijay S. Pande. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *J. Phys. Chem. Lett.*, 5(11):1885–1891, 2014.

[262] Tai-Sung Lee, Yuan Hu, Brad Sherborne, Zhuyan Guo, and Darrin M. York. Toward Fast and Accurate Binding Affinity Prediction with pmemdGTI: An Efficient Implementation of GPU-Accelerated Thermodynamic Integration. *J. Chem. Theory Comput.*, 13(7):3077–3084, 2017.

[263] Joseph W. Kaus, Levi T. Pierce, Ross C. Walker, and J. Andrew McCammon. Improving the Efficiency of Free Energy Calculations in the Amber Molecular Dynamics Package. *J. Chem. Theory Comput.*, 9(9):4131–4139, 2013.

[264] Piero Procacci and Chiara Cardelli. Fast switching alchemical transformations in molecular dynamics simulations. *Journal of Chemical Theory and Computation*, 10(7):2813–2823, 2014.

[265] Shuai Liu, Yujie Wu, Teng Lin, Robert Abel, Jonathan P. Redmann, Christopher M. Summa, Vivian R. Jaber, Nathan M. Lim, and David L. Mobley. Lead optimization mapper: Automating free energy calculations for lead optimization. *Journal of Computer-Aided Molecular Design*, 27(9):755–770, September 2013.

[266] Zoe Cournia, Bryce Allen, and Woody Sherman. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *Journal of Chemical Information and Modeling*, 57(12):2911–2937, December 2017.

[267] Pavel V. Klimovich and David L. Mobley. A Python tool to set up relative free energy calculations in GROMACS. *Journal of Computer-Aided Molecular Design*, 29(11):1007–1014, November 2015.

[268] Vijayaraj Ramadoss, Francois Dehez, and Christophe Chipot. AlaScan: A Graphical User Interface for Alanine Scanning Free-Energy Calculations. *J. Chem. Inf. Model*, 56(6):1122–1126, 2016.

[269] Agastya P. Bhati, Shunzhou Wan, David W. Wright, and Peter V. Coveney. Rapid, Accurate, Precise, and Reliable Relative Free Energy Prediction Using Ensemble Based Thermodynamic Integration. *J. Chem. Theory Comput.*, 13(1):210–222, 2017.

[270] Madhu Pudipeddi and Abu T.M. Serajuddin. Trends in solubility of polymorphs. *Journal of Pharmaceutical Sciences*, 94(5):929–939, May 2005.

[271] J. Bauer, S. Spanton, R. Henry, J. Quick, W. Dziki, W. Porter, and J. Morris. Ritonavir: An extraordinary example of conformational polymorphism. *Pharmaceutical Research*, 18(6):859–866, June 2001.

[272] W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Dzyabchenko, P. Erk, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, J. P. M. Lommerse, W. T. M. Mooij, S. L. Price, H. Scheraga, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer, and D. E. Williams. Crystal structure prediction of small organic molecules: A second blind test. *Acta Crystallographica Section B: Structural Science*, 58(4):647–661, August 2002.

[273] Jaakko Aaltonen, Morten Allesø, Sabiruddin Mirza, Vishal Koradia, Keith C. Gordon, and Jukka Rantanen. Solid form screening – A review. *European Journal of Pharmaceutics and Biopharmaceutics*, 71(1):23–37, January 2009.

[274] D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. Bartashevich, S. X. M. Boerrigter, D. E. Braun, A. J. Cruz-Cabeza, G. M. Day, R. G. Della Valle, G. R. Desiraju, B. P. van Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, D. W. M. Hofmann, F. Hofmann, K. V. J. Jose, P. G. Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, R. J. Needs, M. A. Neumann, D. Nikylov, A. M. Orendt, R. Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. Venuti, and I. K. Zhitkov. Towards crystal structure prediction of complex organic compounds – a report on the fifth blind test. *Acta Crystallographica Section B: Structural Science*, 67(6):535–551, December 2011.

[275] Doris E. Braun, Jennifer A. McMahon, Lien H. Koztecki, Sarah L. Price, and Susan M. Reutzel-Edens. Contrasting Polymorphism of Related Small Molecule Drugs Correlated and Guided by the Computed Crystal Energy Landscape. *Crystal Growth & Design*, 14(4):2056–2072, April 2014.

[276] Aurora J. Cruz-Cabeza and Joel Bernstein. Conformational Polymorphism. *Chemical Reviews*, 114(4):2170–2191, February 2014.

[277] M. A. Neumann, J. van de Streek, F. P. A. Fabbiani, P. Hidber, and O. Grassmann. Combined crystal structure prediction and high-pressure crystallization in rational pharmaceutical polymorph screening. *Nature Communications*, 6:7793, July 2015.

[278] Gregory J. O. Beran. Modeling Polymorphic Molecular Crystals with Electronic Structure Theory. *Chemical Reviews*, 116(9):5567–5613, May 2016.

[279] Anton J. Hopfinger, Emilio Xavier Esposito, A. Llinàs, R. C. Glen, and J. M. Goodman. Findings of the Challenge To Predict Aqueous Solubility. *Journal of Chemical Information and Modeling*, 49(1):1–5, January 2009.

[280] Maryam Salahinejad, Tu C. Le, and David A. Winkler. Aqueous Solubility Prediction: Do Crystal Lattice Interactions Help? *Molecular Pharmaceutics*, 10(7):2757–2766, July 2013.

[281] David S. Palmer and John B. O. Mitchell. Is Experimental Data Quality the Limiting Factor in Predicting the Aqueous Solubility of Druglike Molecules? *Molecular Pharmaceutics*, 11(8):2962–2972, August 2014.

[282] Alessandro Lusci, Gianluca Pollastri, and Pierre Baldi. Deep Architectures and Deep Learning in Chemoinformatics: The Prediction of Aqueous Solubility for Drug-Like Molecules. *Journal of Chemical Information and Modeling*, 53(7):1563–1575, July 2013.

[283] David S. Palmer, Antonio Llinàs, Iñaki Morao, Graeme M. Day, Jonathan M. Goodman, Robert C. Glen, and John B. O. Mitchell. Predicting Intrinsic Aqueous Solubility by a Thermodynamic Cycle. *Molecular Pharmaceutics*, 5(2):266–279, April 2008.

[284] Lunna Li, Tim Totton, and Daan Frenkel. Computational methodology for solubility prediction: Application to the sparingly soluble solutes. *The Journal of Chemical Physics*, 146(21):214110, June 2017.

[285] Devleena Shivakumar, Joshua Williams, Yujie Wu, Wolfgang Damm, John Shelley, and Woody Sherman. Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *Journal of Chemical Theory and Computation*, 6(5):1509–1519, May 2010.

[286] Eliot Boulanger, Lei Huang, Chetan Rupakheti, Alexander D. MacKerell, and Benoît Roux. Optimized Lennard-Jones Parameters for Druglike Small Molecules. *Journal of Chemical Theory and Computation*, April 2018.

[287] Sarah L. Price. The computational prediction of pharmaceutical crystal structures and polymorphism. *Advanced Drug Delivery Reviews*, 56(3):301 – 319, 2004.

[288] G. M. Day, W. D. S. Motherwell, H. L. Ammon, S. X. M. Boerrigter, R. G. Della Valle, E. Venuti, A. Dzyabchenko, J. D. Dunitz, B. Schweizer, B. P. van Eijck, P. Erk, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, F. J. J. Leusen, C. Liang, C. C. Pantelides, P. G. Karamertzanis, S. L. Price, T. C. Lewis, H. Nowell, A. Torrisi, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, and P. Verwer. A third blind test

of crystal structure prediction. *Acta Crystallographica Section B: Structural Science*, 61(5):511–527, October 2005.

[289] Scott M. Woodley and Richard Catlow. Crystal structure prediction from first principles. *Nature Materials*, 7(12):937, December 2008.

[290] G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf, and B. Schweizer. Significant progress in predicting the crystal structures of small organic molecules – a report on the fourth blind test. *Acta Crystallographica Section B: Structural Science*, 65(2):107–125, April 2009.

[291] Sarah (Sally) L. Price. Computed Crystal Energy Landscapes for Understanding and Predicting Organic Crystal Structures and Polymorphism. *Accounts of Chemical Research*, 42(1):117–126, January 2009.

[292] A. M. Reilly, R. I. Cooper, C. S. Adjiman, S. Bhattacharya, A. D. Boese, J. G. Brandenburg, P. J. Bygrave, R. Bylsma, J. E. Campbell, R. Car, D. H. Case, R. Chadha, J. C. Cole, K. Cosburn, H. M. Cuppen, F. Curtis, G. M. Day, R. A. DiStasio Jr, A. Dzyabchenko, B. P. van Eijck, D. M. Elking, J. A. van den Ende, J. C. Facelli, M. B. Ferraro, L. Fusti-Molnar, C.-A. Gatsiou, T. S. Gee, R. de Gelder, L. M. Ghiringhelli, H. Goto, S. Grimme, R. Guo, D. W. M. Hofmann, J. Hoja, R. K. Hylton, L. Iuzzolino, W. Jankiewicz, D. T. de Jong, J. Kendrick, N. J. J. de Klerk, H.-Y. Ko, L. N. Kuleshova, X. Li, S. Lohani, F. J. J. Leusen, A. M. Lund, J. Lv, Y. Ma, N. Marom, A. E. Masunov, P. McCabe, D. P. McMahon, H. Meekes, M. P. Metz, A. J. Misquitta, S. Mohamed, B. Monserrat, R. J. Needs, M. A. Neumann, J. Nyman, S. Obata, H. Oberhofer, A. R. Oganov, A. M. Orendt, G. I. Pagola, C. C. Pantelides, C. J. Pickard, R. Podeszwa, L. S. Price, S. L. Price, A. Pulido, M. G. Read, K. Reuter, E. Schneider, C. Schober, G. P. Shields, P. Singh, I. J. Sugden, K. Szalewicz, C. R. Taylor, A. Tkatchenko, M. E. Tuckerman, F. Vacarro, M. Vasileiadis, A. Vazquez-Mayagoitia, L. Vogt, Y. Wang, R. E. Watson, G. A. de Wijs, J. Yang, Q. Zhu, and C. R. Groom. Report on the sixth blind test of organic crystal structure prediction methods. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, 72(4):439–459, August 2016.

[293] Christophe Chipot. Frontiers in free-energy calculations of biological systems. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4(1):71–89, January 2014.

[294] D. Frenkel and A. J. C. Ladd. New Monte Carlo method to compute the free energy of arbitrary solids. Application to the fcc and hcp phases of hard spheres. *The Journal of Chemical Physics*, 81(7):3188–3193, 1984.

[295] E. G. Noya, M. M. Conde, and C. Vega. Computing the free energy of molecular solids by the Einstein molecule approach: Ices XIII and XIV, hard-dumbbells and a patchy model of proteins. *The Journal of chemical physics*, 129(10):104704, September 2008.

[296] Carlos Vega and Eva G. Noya. Revisiting the Frenkel-Ladd method to compute the free energy of solids: The Einstein molecule approach. *The Journal of Chemical Physics*, 127(15):154113, October 2007.

[297] J. L. Aragones, E. G. Noya, C. Valeriani, and C. Vega. Free energy calculations for molecular solids using GROMACS. *The Journal of Chemical Physics*, 139(3):034104, July 2013.

[298] Jooyeon Park, Ian Nessler, Brian McClain, Dainius Macikenas, Jonas Baltrusaitis, and Michael J. Schnieders. Absolute Organic Crystal Thermodynamics: Growth of the Asymmetric Unit into a Crystal via Alchemy. *Journal of Chemical Theory and Computation*, 10(7):2781–2791, July 2014.

[299] Y. Kim, K. Machida, T. Taga, and K. Osaki. Structure redetermination and packing analysis of aspirin crystal. *Chemical & Pharmaceutical Bulletin*, 33(7):2641–2647, July 1985.

[300] Samuel H. Yalkowsky, Yan He, and Parijat Jain. *Handbook of Aqueous Solubility Data*. CRC Press, second edition edition, 2010.

[301] D. G. Henshaw. Atomic Distribution in Liquid and Solid Neon and Solid Argon by Neutron Diffraction. *Physical Review*, 111(6):1470–1475, September 1958.

[302] Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation*, 4(3):435–447, March 2008.

[303] M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12):7182, 1981.

[304] C. F. Macrae, P. R. Edgington, P. McCabe, E. Pidcock, G. P. Shields, R. Taylor, M. Towler, and J. van de Streek. Mercury: Visualization and analysis of crystal structures. *Journal of Applied Crystallography*, 39(3):453–457, June 2006.

[305] Romelia Salomon-Ferrer, David A Case, and Ross C Walker. An overview of the Amber biomolecular simulation package. *WIREs Computational Molecular Science*, 3(2):198–210, September 2012.

[306] Thomas E. Cheatham and David A. Case. Twenty-five years of nucleic acid simulations. *Biopolymers*, 99(12):969–977, December 2013.

[307] D. A. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. Cheatham, T. A. Darden, R. E. Duke, H. Gohlke, A. W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, F. Paesani, D. R. Roe, A. Roitberg,

C. Sagui, R. Salomon-Ferrer, G. Seabra, C. L. Simmerling, W. Smith, J. Swails, Walker, J. Wang, R. M. Wolf, X. Wu, and P. A. Kollman. Amber 14. University of California, San Francisco, 2014.

[308] Kyle A Beauchamp, Ariën S Rustenburg, Andrea Rizzi, Julie M Behr, Guilherme Matos, Lee-Ping Wang, Robert T McGibbon, David L Mobley, and John D Chodera. OpenMolTools.

[309] Daan Frenkel and Berend Smit. *Understanding Molecular Simulation: From Algorithms to Applications*, volume 1. Academic Press Inc. (London) Ltd., 1996.

[310] Christian Kramer, Alexander Spinn, and Klaus R. Liedl. Charge Anisotropy: Where Atomic Multipoles Matter Most. *Journal of Chemical Theory and Computation*, 10(10):4488–4496, October 2014.

[311] D. M. Eike, J. F. Brennecke, and E. J. Maginn. Toward a robust and general molecular simulation method for computing solid-liquid coexistence. *The Journal of chemical physics*, 122(1):14115, January 2005.

[312] D. M. Eike and E. J. Maginn. Atomistic simulation of solid-liquid coexistence for molecular systems: Application to triazole and benzene. *The Journal of chemical physics*, 124(16):164503, April 2006.

[313] A. S. Paluch, S. Jayaraman, J. K. Shah, and E. J. Maginn. A method for computing the solubility limit of solids: Application to sodium chloride in water and alcohols. *The Journal of chemical physics*, 133(12):124504, September 2010.

[314] Eric C. Dybeck, Natalie P. Schieber, and Michael R. Shirts. Effects of a More Accurate Polarizable Hamiltonian on Polymorph Free Energies Computed Efficiently by Reweighting Point-Charge Potentials. *Journal of Chemical Theory and Computation*, 12(8):3491–3505, August 2016.