JCTC Journal of Chemical Theory and Computation

Perspective

pubs.acs.org/JCTC

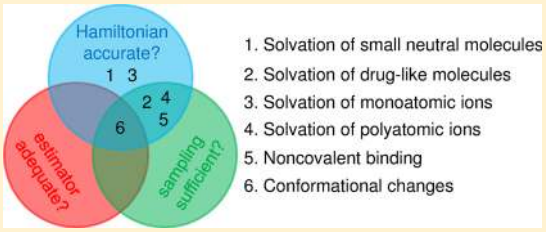
Practical Aspects of Free-Energy Calculations: A Review

Niels Hansen[†,‡,*] and Wilfred F. van Gunsteren[‡]

[†]Institute of Thermodynamics and Thermal Process Engineering, University of Stuttgart, D-70569 Stuttgart, Germany

[‡]Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH, CH-8093 Zürich, Switzerland

ABSTRACT: Free-energy calculations in the framework of classical molecular dynamics simulations are nowadays used in a wide range of research areas including solvation thermodynamics, molecular recognition, and protein folding. The basic components of a free-energy calculation, that is, a suitable model Hamiltonian, a sampling protocol, and an estimator for the free energy, are independent of the specific application. However, the attention that one has to pay to these components depends considerably on the specific application. Here, we review six different areas of application and discuss the relative importance of the three main components to provide the reader with an organigram and to make nonexperts aware of the many pitfalls present in free energy calculations.



1. INTRODUCTION

A wide range of fundamental chemical quantities such as binding or equilibrium constants, solubilities, partition coefficients, and adsorption coefficients are related to the difference in free energy between particular (non)physical states of a system.[1] By means of statistical mechanics, free-energy differences may be expressed in terms of averages over ensembles of atomic configurations for the molecular system of interest. Such an ensemble can be generated by Monte Carlo (MC) or molecular-dynamics (MD) techniques. Much of the statistical-mechanical framework for calculating free energy differences has been developed some time ago.[2−6] The free energy F of a system in the canonical ensemble (i.e., at constant number of particles, volume, and temperature) is given as

$$F = -k_B T \ln\{[h^{3N}N!]^{-1} \int\int e^{-H(\vec{p}^N,\vec{r}^N)/k_B T} d\vec{p}^N d\vec{r}^N\} \quad (1)$$

where k_B is Boltzmann's constant, T the temperature, h Planck's constant, N the number of particles or atoms in the classically treated molecular system, and $\vec{r}^N = (\vec{r}_1,\vec{r}_2,...\vec{r}_N)$ and $\vec{p}^N = (\vec{p}_1,\vec{p}_2,...\vec{p}_N)$ the Cartesian coordinates and conjugate momenta of all N atoms, respectively. The factor $N!$ is only present for indistinguishable particles. The Hamiltonian $H(\vec{p}^N, \vec{r}^N)$ consists of the kinetic energy $K(\vec{p}^N)$ and the potential energy $V(\vec{r}^N)$ of the system. The latter describes the interactions between the various atoms and is also called a molecular force field.[7] Since the integral in eq 1 is $6N$-dimensional and the integrand is positive definite, the absolute free energy can only be calculated in particular cases, that is, for very simple model systems for which an analytical expression for the partition function can be obtained. In an isothermal−isobaric ensemble, the corresponding free enthalpy[8] or Gibbs energy[9] is given by

$$G = -k_B T \ln\{[Vh^{3N}N!]^{-1} \int\int\int e^{-[H(\vec{p}^N,\vec{r}^N)+pV]/k_B T} d\vec{p}^N d\vec{r}^N dV\} \quad (2)$$

In condensed phase systems such as biomolecules in aqueous solutions the pressure−volume work term pV is usually negligible such that we use the term *free energy* throughout this review.

For practical applications, it generally suffices to calculate relative free energies,

$$\Delta F_{BA} = F_B - F_A \quad (3)$$

where A and B denote two different systems, $H_A(\vec{p}^N, \vec{r}^N)$ and $H_B(\vec{p}^N, \vec{r}^N)$, or two different states that are connected by a coupling parameter λ, $H(\vec{p}^N, \vec{r}^N; \lambda)$ leading to

$$F_\lambda(\lambda') = -k_B T \ln\{[h^{3N}N!]^{-1} \int\int e^{-H(\vec{p}^N,\vec{r}^N;\lambda')/k_B T} d\vec{p}^N d\vec{r}^N\} \quad (4)$$

or two different states along a phase space (reaction) coordinate $R(\vec{r}^N)$, leading to

$$F_R(R') = -k_B T \ln\{[h^{3N}N!]^{-1} \int\int \delta(R' - R(\vec{r}^N)) \times e^{-H(\vec{p}^N,\vec{r}^N)/k_B T} d\vec{p}^N d\vec{r}^N\} \quad (5)$$

How to apply this framework in computer simulations in an effective fashion remains, however, a very active area of research, giving rise to over 3500 papers using the most popular free energy methods that were published in the first decade of

Special Issue: Free Energy Calculations: Three Decades of Adventure in Chemistry and Biophysics

Received: February 24, 2014
Published: May 20, 2014

   dx.doi.org/10.1021/ct500161f | J. Chem. Theory Comput. 2014, 10, 2632−2647

this century with the publication rate increasing ∼17% per year.[10] This rapid development comes at the price that it is increasingly difficult for researchers to find their way through the maze of available computational techniques. For reviews on the methodology to calculate free energy via molecular simulation, we refer to the literature.[11−20] Reviews that are more application oriented can be found in refs 21−23. No single method for free-energy calculation can be considered as clearly superior to the others, and the proper choice depends very much on the system under consideration. Christ et al.[18] identified three main challenges that have to be met in free energy estimation from molecular simulation, which are (i) the choice of a suitable model Hamiltonian, (ii) the choice of a sampling protocol that allows generating a representative ensemble of configurations, and (iii) the choice of an estimator for the free energy difference. The aim of the latter review was to enable the reader to classify the vast array of methods by identifying which choices have to be made for these three basic components. Here, we attempt to enable the reader to be aware of the various peculiarities and pitfalls that are characteristic for the various (bio)molecular systems for which free-energy calculations are performed. We do this by discussing equilibrium free-energy computation from the systems point of view.

In classical molecular simulation three categories of free-energy differences can be distinguished, which are *conformational*, *alchemical*, and *thermodynamic* free energy differences. Conformational free-energy differences refer to differences between two distinct conformational states (e.g., a left and a right-handed helix) of the same system. Alchemical free-energy differences refer to differences between two states differing in their Hamiltonian. Thermodynamic free-energy differences refer to differences between two thermodynamic state points (e.g., two different temperatures). In this review, we focus on the first two types of calculations, their differences can be stated as follows. The free-energy difference between two conformational states is a logarithmic measure of the relative partition functions corresponding to a common Hamiltonian but integrated over different regions of configurational space of the system. The free-energy difference between two alchemical states is a logarithmic measure of the relative partition functions corresponding to the entire space in terms of the degrees of freedom of the system, but performed considering two different Hamiltonians (that are both functions of the same number of degrees of freedom).

Due to the explosion in the number of publications since the early 1990s no review of the state of the art can hope to be comprehensive. However, there is broad consensus about the necessity of further methodological developments and the definition of best practices in all areas of applications in (bio)molecular modeling, of which the main ones are (i) solvation of small neutral molecules, (ii) solvation of larger (drug-like) heteromolecules, (iii) solvation of (single) monatomic ions, (iv) solvation of polyatomic ions, (v) noncovalent binding, and (vi) conformational changes. In the following these application areas are reviewed separately with an eye to identify current obstacles and challenges.

2. SOLVATION OF SMALL NEUTRAL MOLECULES

The free energy of solvation corresponds to the free energy of transferring a compound from a well-defined state (gas) to another (solution), allowing a direct comparison with experiment. The calculation of the free energy of solvation usually involves an alchemical transformation and was one of the first practical applications of the free energy perturbation and integration methodology.[21,22,24] Today, solvation free energies remain of primary importance for developing and testing force fields, for testing new methodology, for gaining fundamental insights into the solvation process, and for specific applications such as predicting how molecular compounds will partition between different environments.

Condensed-phase or biomolecular force fields (e.g., CHARMM,[25−29] AMBER,[30−33] OPLS,[34−38] and GROMOS[39−49]) mainly aim at the description of (bio)molecules in solution. They focus on the description of torsional-angle properties, nonbonded interactions, and solvation effects. However, the parameter sets used to describe most compounds have historically been based primarily on structural properties and fitting to results from quantum-mechanical calculations.[50,51] This is despite the fact that many properties of interest, especially in biomolecular systems, such as protein or peptide folding, depend on how compounds or their constituent moieties partition between different environments. The primary reason why thermodynamic properties have until recently not been more generally incorporated into force-field parametrization was cost, and the difficulty in obtaining converged results.

However, solvation free energies have been used extensively for the verification of the OPLS force field,[52−54] especially to rationalize the choice of partial atomic charges,[55−57] which are empirical parameters. For the most recent versions of the OPLS all-atom (AA) force field the partial charge assignment was based on hydration free energies calculated in explicit solvent for a set of 239 small molecules spanning diverse chemical functional groups commonly found in drugs and drug-like molecules.[58,59]

When reparametrizing the GROMOS force field for the aliphatic CH_n united atoms in 1998, Daura et al.[40] fitted the repulsive van der Waals parameter of the water oxygen for interactions with nonpolar atoms to reproduce the experimental free energies of hydration of five of the alkanes studied: methane, ethane, propane, butane, and isobutane. Subsequent versions of the GROMOS force field have specifically been developed to reproduce the free enthalpy of hydration and apolar solvation in explicit solvent. The 53A5 force field[44] was developed for an accurate description of the thermodynamic properties of pure liquids. Since it did not seem possible at that time to reconcile both pure liquid and hydration properties with a sufficient accuracy, the 53A6 force field[44] aimed primarily at reproducing solvation properties of (neutral) amino acid side chain analogues in aqueous and nonaqueous solvents. The $53A6_{OXY}$ parameter set[60] strikes a balance between 53A5 and 53A6 versions, by providing a single set to reproduce both pure-liquid properties and solvation free energies of small oxygen-containing molecules. The 54A7 force field[48] is based on 53A6 with a number of force field adjustments, among which the Lennard-Jones interaction parameters of the sodium and chloride ions were changed based on calculated absolute single ion solvation free energies (see also below).[61] In subsequent work, also, other functional groups such as occurring in sulfones[62] or in amines, amides, thiols, sulfides, and aromatics were (or are currently) (re)parametrized aiming at simultaneously improving pure-liquid and solvation properties. Recently, force-field parameters describing 110 post-translationally modified (PTM) amino acids and protein termini, compatible with the GROMOS force

fields 45A3 and 54A7, were derived. Validation against the hydration free energy of side-chain analogues showed that the newly generated parameters compatible with the GROMOS 54A7 parameter set reproduce experimental data almost equally well as the original parameters reproduce hydration free energies for side-chain analogues of naturally occurring amino acids.[63] A web server for automated introduction of PTMs to protein 3D structures is available.[64] Parameters for post-translationally modified amino acids have also been reported for the AMBER ff03 force field. However, the assessment was based on structural properties of proteins rather than on hydration free energies of side-chain analogues.[65]

Recently, Mobley et al.[66,67] have reported hydration free energies for 504 small molecules parametrized using the AMBER Antechamber program[68] to assign parameters for a general AMBER force field (GAFF),[69] which is compatible with the all-atom AMBER biomolecular force field. They identified systematic errors in force-field parameters for particular functional groups such as alkynes, which could be fixed by using OPLS Lennard-Jones parameters for triple bonded carbons. Subsequent work focused on ethane, biphenyl, and dioxin and their chlorinated derivatives.[70] Except for the ethane derivatives, the agreement with experimental data was less good compared to their earlier set of small molecules. The introduction of virtual charge sites led to some improvement but the remaining error might not be reducable using fixed charge models. Hydration free energies were calculated with both implicit[66] and explicit[67] solvent. It was concluded that the explicit solvent simulations reliably outperform today's implicit solvent models. Based on their extensive set of data, Mobley et al.[67] also proposed several remedies to improve implicit solvent models. The better performance of explicit solvent simulations over implicit ones was also noted by Shivakumar et al.[71] using the GAFF[69] force field and the all-atom CHARMm-MSI[72] one.

More recently, Knight et al.[73] calculated the hydration free energies of a set of 457 small neutral molecules using the CHARMM general force field CGenFF,[74,75] which is compatible with the all-atom CHARMM biomolecular force field,[76] using three different implicit solvent models. The set of molecules was based on the one used by Mobley et al.[67] and parametrized automatically using the CHARMM-compatible ligand parametrization tools MATCH,[77] ParamChem,[74,75] and SwissParam,[78] as well as by converting the GAFF parameters from Mobley et al.[67] to CHARMM format. GAFF uses the restrained electrostatic potential (RESP) fitting approach[79] to generate charges for the entire molecule concurrently while the MATCH, ParamChem, and SwissParam tools use a fragment-based approach, where charge distributions of a molecule are built-up from charges assigned to the component fragments of the molecule. Of the 12 combinations of solvent model and parametrization scheme the Antechamber parameters (GAFF with semiempirical Merck-Frosst AM1-BCC partial atomic charges[80,81]) yielded the most accurate estimates.

The latter study and others[82−84] show that existing methods based on implicit solvation models are often reasonably accurate in calculating solvation free energies of a large number of compounds. However, explicit solvent simulations should still be regarded as the "gold standard" due to the aim of using these force fields for heterogeneous systems in solution.[67] Recently, an extension of the automatic parametrization schemes for GAFF and CGenFF has been reported that uses QM results as primary data.[85] This general automated atomic model parametrization (GAAMP) approach includes also the optimization of dihedral parameters, that often have limited transferability, in addition to electrostatic parameters. The parametrization can start from GAFF or CGenFF as the initial models.

It should be kept in mind that force-field parametrizations are usually carried out for one particular water model. While Shivakumar et al.[58] showed that there is little difference in overall performance of hydration free-energy predictions between the SPC, TIP3P, and TIP4P water models for a subset of 13 molecules out of the 239 molecules used in their study, Shirts and Pande showed that a modified version of the TIP3P model better reproduces hydration free energies of 15 amino acid side-chain analogues.[86] The influence of the solvent model on calculated hydration free energies of nucleobases and chloroform-to-water partition coefficients was recently studied by Wolf and Groenhof[87] using various combinations of force fields in conjunction with their native and non-native water models. It was found that the large differences in solvation free energies between different force fields are actually due to the nucleobase parameters and not the solvent models and that the difference in hydration free energy due to the use of a different water model is larger in case of aromatic amino acid analogues than in case of nucleobases. Hess and van der Vegt[88] also came to the conclusion that the choice of water model strongly influences the accuracy of calculated free energies of hydration of amino acid side-chain analogues.

Predictions of hydration free energies for compounds with multiple functional groups showed that the training sets currently used still lack sufficient coverage of chemical space.[89,90] For the class of nitroaromatic compounds, Ahmed and Sandler[91] tested 16 force-field/(charge+water) models, out of which only 6 performed approximately equally well in predicting measured hydration free energies.

To assess the state of the art of a force field other properties than the solvation free energy should also be considered. Using a total of 146 molecular liquids, Caleman et al.[92] compared the ability of the OPLS-AA and GAFF force fields to reproduce key properties of neat liquids such as density, enthalpy of vaporization, surface tension, heat capacity at constant volume and pressure, isothermal compressibility, volumetric expansion coefficient, and static dielectric constant. The overall performance of the OPLS-AA force field was found to be somewhat better, but both force fields have issues with reproduction of the surface tension and the dielectric constant. By including experimental dielectric response data, in addition to liquid density and heat of vaporization, into the parameter optimization, Fennell et al.[93] arrived at new GAFF parameters for hydroxyl groups that also lead to improvements in the calculation of hydration free energies for a test set of 41 small molecule alcohols.

The more diverse the test sets become, and the more diverse the chemical environments are in which small molecules should be simulated, the more likely it becomes that fixed charge models are reaching their limits. Mobley et al.[67,90] reported an RMS deviation from experimental numbers of 5.2 kJ mol^{-1} for their test set of 504 molecules, which is comparable to the RMSD of 6.5 kJ mol^{-1} Caleman et al.[92] found for the heat of vaporization. For larger molecules with multiple functional groups, the error in the hydration free energy increases to up to 10 kJ mol^{-1}.[89] To be compatible with statistical QSPR methods[94,95] to predict solubility, which have negligible computational cost, the RMS error has to decrease to about 4 kJ mol^{-1}, also for complex molecules, to justify the significant

computational overhead of explicit solvent simulations. Additional degrees of freedom can be included in the parametrization process by modifying the form of the potential energy function for the short-range interactions, the combination rules for unlike interaction partners, and by considering explicit polarization. While the Lennard-Jones 12−6 potential energy function is the most widely used one, it is well-known that increasing the repulsion exponent may improve the description of vapor−liquid equilibrium data.[96,97] Repulsive exponents smaller than 12 such as 9 or 8 have also been suggested.[98,99] The former is used for example in the condensed-phase force field COMPASS.[100] Regarding the combination rules, the standard arithmetic or geometric mean rules often perform poorly[101,102] and are mainly used due to the lack of systematic and transferable procedures for the derivation of improved combination rules. For including explicit polarization in empirical force fields, various strategies are currently developed for many of the condensed phase force fields[103] together with automated parametrization schemes that aim at reproducing hydration free energies.[85] However, from disagreement between an experimental and a calculated hydration free energy alone, it cannot be concluded whether explicit polarization is needed. This usually requires the consideration of additional properties such as the dielectric permittivity. Missing polarizability has, for example, been demonstrated to be the reason for the underestimation of the hydration free energy and the dielectric permittivity of *N*-methylacetamide by fixed charge models.[104] We note, however, that the calculation of solvation free energies for small molecules is only one aspect of the complex process of (bio)molecular force-field development.

For protein simulations, force-field improvements can only be evaluated on the basis of comparing simulation results to experimental data for a diverse set of protein structures. Recently, Nerenberg et al.[105] presented a parametrization strategy for a fixed-charge protein force field based on the AMBER ff99SB parameters[106] combined with the TIP4P-Ew water model,[107] which is a reparameterized version of the standard TIP4P water model for use with Ewald summation techniques. They calibrated the solute−solvent van der Waals interaction parameters of a set of 47 small molecules representing all of the chemical functionalities of standard protein side chains and backbone groups. To be consistent with the charge model used in AMBER ff99SB, partial charges were obtained by fitting calculated electrostatic potentials at the HF/6-31G* level using the RESP method.[79] The RMSE in solvation free energies could be reduced from 7.3 to 2.5 kJ mol^{-1}. With the combination of original AMBER ff99SB parameters and TIP4P-Ew water, nearly every chemical moiety was undersolvated. The new solute−solvent interaction parameters were evaluated based on simulations of dipeptide solutions, of short disordered peptides and of ubiquitin. For the latter case, the favorable enhancement of solute−water interactions resulted in partial unfolding although the hydrophic core remained intact. By reintroducing a 12−10 hydrogen bonding potential energy term[108,109] (instead of 12−6), this problem could be remedied.

The latter study made use of reweighting for evaluating the influence of modified van der Waals parameters based on trajectories generated with the unmodified ones. This approach is very efficient for parameter studies if the perturbed ensemble has sufficient overlap with the unperturbed one.[110,111] Through the use of Jacobian factors, it can also be used in case of geometric changes.[112]

We conclude that the use of solvation free energies in force-field calibration, testing, and comparison has become a standard tool. However, it is again noted that the force-field parameters are empirical. They have been derived using a specific set of conditions to reproduce a specific set of properties. As with any empirical force field, the choice of temperature, treatment of long-range nonbonded interactions, pressure coupling scheme, and so on, are implicitly incorporated into the parametrization but are in fact properties of the underlying algorithms used to integrate Newton's equations of motion. This should be kept in mind when using force fields in conjunction with MD codes that were not used for the original parametrization or when changing the implicitly included parameters. The density of liquid octanol, for example, was reported to be significantly overestimated while, at the same time, the heat of vaporization was reported to be underestimated with the GROMOS 53A6 force field, using the GROMACS code presumably with nonstandard settings[113] compared to the recommended settings in GROMOS.[60] This problem might become more subtle for polarizable force fields, which include additional choices, such as how they treat intramolecular polarization.[114]

The current momentum in the area of solvation free energy calculations will soon lead to a further reduction of the discrepancy between experiment and simulation, and it remains to be seen what the residual error will be that defines the limit of classical force fields.

In terms of the three choices, we conclude that the choice of the Hamiltonian is the one that has the strongest influence on the accuracy of solvation free energies of small neutral molecules. For the other two choices, the current literature demonstrates that a variety of techniques is available (see below), which can, in many cases, be used interchangeably to calculate a solvation free energy with a precision higher than the acurracy of the force field. However, also, relatively small molecules require special attention if they show slow torsional-angle transitions such as in carboxylic acids,[115] ibuprofen,[116] or dimethoxyethane.[117]

Because many methods can often be used interchangeably to calculate solvation free energies, such simulations are frequently used in method development, testing, and comparison. Prompted by difficulties with complex intramolecular potential energy surfaces, expanded ensemble methods,[118,119] a hybrid Monte Carlo−molecular dynamics approach, has been found to be significantly more efficient than TI[120] or BAR[4] due to an enhanced conformational sampling.[121−123] A combination of λ-dynamics and metadynamics[124] has been suggested recently to enhance sampling on virtual variables.

Alternatives to the widely used nonlinear soft-core scaling proposed by Beutler et al.[125] are being developed that flatten the potential energy only in a region that is energetically inaccessible under normal conditions.[126] Naden and Shirts[127] proposed a formalism based on splitting the potential energy function into a configurational and an alchemical part. This approach leads to a lower variance and can be efficiently implemented. Also, methods based on sampling along a reaction coordinate, such as the adaptive biasing force method,[128] have been used to calculate the solvation free energy directly by transferring a solute from the gas phase into the solvent across the gas/solvent interface.[129]

Hydration free energies provide a probe of the underlying physics. Due to the asymmetry of the charge sites in water, its

response to polar solutes depends on the internal charge distribution of the latter. Mobley et al.[130,131] studied the hydration of various polar solutes (both fictious ones with simple geometry and real ones) in different explicit water models. By inverting the charge distribution in the artificial solutes, large differences in the hydration free energy of up to 40 kJ mol^{-1} were found, largely driven by the structure of water in the first hydration shell. Charge hydration asymmetry also occurs in ionic solvation and the insights gained from explicit solvent simulations have helped to improve implicit solvent models.[132−134] Also, the treatment of nonpolar contributions to the solvation free energy in implicit solvent models could be improved based on insights gained from explicit solvent simulations.[135] By studying the solubility of alkanes up to *n*-eicosane in water Ferguson et al.[136] detected no sharp peak in the dependence of the solubility upon carbon numbers larger than 12 as suggested by different experimental sources, but rather a nearly exponential decrease. Moreover, remarkable similarities in the conformational ensemble in the gas and solvated phases were found suggesting that the effect of the solvent interaction is the appearance of a free energy barrier of order k_BT separating the compact and extended free energy basins for sufficiently long chains and a destabilization of the most extended chain conformations.

Finally, solvation free energy calculations are used to predict physicochemical properties of compounds for which experimental data are scarce and structure−property relationships are uncertain such as for nitroaromatic compounds.[137,138]

3. SOLVATION OF LARGER (DRUG-LIKE) HETEROMOLECULES

In many practical applications, topologies and force-field parameters of larger heteromolecules such as substrates, inhibitors, cofactors, or drug molecules[139−142] are needed. These parameters are not standardly available and often have no close analogues within the desired biomolecular force field. Therefore, they have to be specifically assigned manually or by automated procedures, which are available for all of the main families of force fields such as for Amber/GAFF,[68,85,143] CHARMM/CGenFF,[74,75,77,85,144] OPLS,[145] and GROMOS,[146,147] although with different levels of sophistication. Because the parametrization is an underdetermined problem, it has to be ensured that the parameters are consistent with the macromolecular force field applied to the other components of the system. For thermodynamically calibrated force fields, this validation should include the calculation of the solvation free energy in polar and nonpolar environments. Unfortunately, experimental values for solvation free energies are only known for less than one percent of the millions of organic compounds prepared to date.[148,149] It is of importance to develop efficient robust and reliable calculation procedures for such molecules that are often characterized by multiple conformational substates[150−154] that have to be sampled sufficiently in both the liquid and the gas phase such that the calculated free energy depends only on the force-field parameters. This makes the use of enhanced sampling techniques mandatory in many cases. Khavrutskii and Wallqvist[155] combined thermodynamic integration with Hamiltonian replica exchange and showed that converged results are obtained for molecules with internal rotational barriers of up to 60 kJ mol^{-1} using only a few nanoseconds of simulation time.

Paluch et al.[156] proposed a method to predict the solubility limit of low solubility solids based on a single experimental reference solubility for each solute and a single free energy simulation of the solute−solvent system. The latter was carried out using an expanded ensemble calculation along with a combined Wang−Landau/Bennett's acceptance ratio method. A particularly important aspect was that the proposed method requires fewer experimental data points than the Abraham general solvation model[157] used for comparison.

In the absence of any experimental data, solvation free energies are often predicted by group contribution methods, such as to predict the hydration free energies of amino acids or proteins. However, explicit-solvent MD-based hydration free energies for 15 N-acetyl-methylamide amino acids with neutral side chains differ considerably from those based on additive group contribution methods.[158]

Another example of missing experimental hydration free energies is that of barbiturates, which are of great pharmaceutical interest. Using parameters from the OPLS-AA force field, Garrido et al.[159] tested different simulation setups for an efficient calculation of the hydration free energy including alternative ways to account for electrostatic interactions, such as the reaction-field method and the particle-mesh Ewald method. If the former is used in an automated setup, the problem of assigning atoms to charge groups may arise. Recently Canzar et al.[160] showed how this problem can be solved efficiently.

A recent blind test including 63 complex drug-like molecules showed that solvation free energies are, at present, predicted with an RMS error of 10−14 kJ mol^{-1}.[148]

In terms of the three choices, we conclude that the choice of Hamiltonian for the solvation of drug-like molecules is as crucial as for the solvation of small neutral molecules but its accurracy is more difficult to assess due to scarse experimental data. Moreover, the sampling protocol and free-energy estimator determine the efficiency of the calculation much more than is the case for small molecules. The choice of an appropriate technique is more case-dependent than for small molecules.

This renders the question of convergence to be more crucial for drug-like molecules compared to the case of small molecules discussed above. Approaches such as "double-wide sampling" or forward and reverse transformations to check for hysteresis in an alchemical free energy difference are only of limited value if the forward and reverse directions show very different convergence behavior such as in case of insertion and deletion of particles.[161] If both equilibrium ensembles have been sampled, the two trajectories can be efficiently combined in the context of the Bennett acceptance ratio (BAR) method.[4]

Apart from sampling errors that originate from suboptimal sampling of important regions of phase space, the bias due to finite sample size[162,163] is more crucial for larger molecules with multiple possibly slow degrees of freedom. As pointed out by Lyman and Zuckerman,[164] the simulation length to achieve statistically independent configurations may be much longer than expected. On the way to install measures for quality assurance of simulation results,[165] several best practices and validation tests have been proposed that help standardize the setup and evaluation of free energy calculations.[166−168]

4. SOLVATION OF (SINGLE) MONOATOMIC IONS

Within the realm of classical thermodynamics, only sums of thermodynamic quantities associated with neutral sets of ions can be probed. These sums may be partitioned into single-ion contributions, but only within an unknown offset constant

weighted by the integer ion charge. Single-ion thermodynamics becomes accessible when spectroscopic techniques such as photoionization or laser photodetachment are combined with statistical mechanics.[169] However, the resulting parameters, also referred to as *real* single-ion solvation parameters, still account for the mixture of two physically very different effects: bulk solvation and liquid−surface properties. The estimation of the surface term requires extrathermodynamic considerations, which may be either experimental or theoretical, and show considerable spread (e.g., as much as 100 kJ mol^{-1} in terms of solvation free energies for monovalent ions), making the estimation of *intrinsic* absolute solvation parameters (i.e., those which originate exclusively from the ion−solvent interactions) uncertain. For a comprehensive discussion of these issues, we refer to the book of Hünenberger and Reif.[169] In the context of atomistic (explicit-solvent) simulations, the single-ion solvation free energy is usually measured as the work of coupling the solute (ion) to the solvent (water) and thus reflects naturally the intrinsic absolute solvation free energy of an ion I^z (denoted by $\Delta_s G^{\boxminus}[I_g^z]$ in ref 169), that is, ideal solvation in a surface-free fluid at a fixed reference concentration (the same number of moles per liter in the vapor and solution) under which condition the entropy of translation in the vapor equals the entropy of liberation from a fixed point in solution.

Since the earliest calculations of ionic solvation free energies in the 1980s,[170−173] it has been realized that the raw results can be dramatically sensitive to the boundary conditions and treatment of electrostatic interactions used during these simulations, with typical variations on the order of 100 kJ mol^{-1} or larger in terms of solvation free energies for monovalent ions in water. As a result, the parametrization of ion−solvent interaction parameters suffers from problems related to the ambiguity of the experimental training set as well as from the need to correct the raw methodology-dependent simulation results, so that methodology-independent values are obtained. The main factors that lead to discrepancies between simulation (using fixed-charge water models) and experiment in the context of single-ion solvation are

(1) approximate force-field representation (functional form, e.g. absence of explicit electronic polarizability in the model),
(2) approximate force-field parameters (water and ion−water parameters),
(3) finite sampling errors,
(4) approximate electrostatics errors,
(5) finite size errors,
(6) improper summation errors,
(7) additional approximations involved in the evaluation of properties other than the solvation free energy (see below),
(8) inaccuracy or ambiguity of the experimental data.

Points 1 and 2 are related to the choice of the Hamiltonian and are thus not different from other free-energy calculations. Also, point 3 has to be considered carefully in every free-energy calculation. For points 4−6, a scheme has been developed that allows to correct *raw* solvation free energies *ex post*, so that methodology-independent values are obtained.[61,169,174−178] The corrected results are then exclusively characteristic of the underlying molecular model, as determined by the representation of the solvent, of the ion, and of the ion−solvent van der Waals interactions, and no longer depend on arbitrary simulation parameters such as the system size or the electrostatic cutoff distance. These values correspond to the idealized situation of an infinite bulk phase exempt of surface polarization, in which electrostatic interactions are exactly Coulombic. Using this correction scheme, a reparametrization of the ion−solvent Lennard-Jones interaction coefficients for Na^+ and Cl^- (among other ions) with the SPC water model[179] (as well as the SPC/E model[180]) against experimental hydration free energies was conducted by Reif and Hünenberger.[61] Three different parameter sets (L, M, and H) were calibrated, corresponding to different assumed values for the absolute intrinsic hydration free energy $\Delta G^{\ominus}_{hyd}[H^+]$ of the proton at P° = 1 bar and T^- = 298.15 K, which is an experimentally elusive quantity.[169] Recently, Dahlgren et al.[181] extended the correction scheme to derivative thermodynamic hydration and aqueous partial molar properties (point 7 in the above list) and showed that approximate internal consistency and qualitative agreement with experimental results can only be achieved when an appropriate correction scheme is applied, along with careful consideration of standard-state issues. As for the free energy itself, the correction terms for derivative thermodynamic hydration and aqueous partial molar properties are substantial. Directions for future improvements, with the ultimate goal of reaching a consistent and quantitative description of single-ion hydration thermodynamics in atomistic simulations, are also provided in the latter work. Apart from point 8, the correction scheme permits a thermodynamically consistent calibration of ion−solvent interaction parameters. Not correcting for points 4−6 necessarily leads to some kind of system-size dependence in the calculated quantities, the degree of which depends on the actual electrostatic scheme used. This might be acceptable for simulations focusing on structural properties such as radial distribution functions, which show remarkable insensitivity to the detailed treatment of electrostatic forces. Simulations attempting to calculate quantities related to the energy of the system, however, lump all the overlooked errors into the non-Coulomb interaction parameters and are therefore not transferable to other system sizes and electrostatic schemes. When reviewing the most recent literature related to ionic force-field parametrization or the use of atomistic simulations to probe single-ion solvation properties,[182−191] however, it is striking that correction terms are seldom taken into account. One of the reasons for not including such corrections might be that in practical simulations ions are propagated with approximate electrostatic interactions within systems of finite sizes which differs from the ideal situation of Coulombic electrostatic interactions in a macroscopic nonperiodic system underlying the interaction parameters calibrated against methodology independent hydration free energies. This might strongly affect the configurational sampling and might lead to a significant undersolvation especially for small box sizes,[190] because the correction terms representing methodology-independent ion hydration free energies are predominantly negative. It would therefore be desirable to design effective electrostatic interaction schemes which correct the approximate electrostatics and finite-size effects at the level of the forces, so as to achieve a solvent polarization around ionic groups that is exempt of artifacts. Attemps to include the correction terms into the equation of motion by means of restraints are in progress.[192]

We note that calibrating ion-pairs instead of single ions leads to an approximate cancellation of improper summation errors because these are linear in the ionic charge, but it has the

disadvantage that the underlying single-ion solvation free energies can only be determined up to an additive constant leading to a range of possible interaction parameters that all lead to the correct ion pair properties in dilute solutions. When using simulations of finite concentrations to determine these additive constants and thus the appropriate interaction parameters, one faces the problem of calibrating simultaneously ion−ion and ion−solvent interactions, even for low concentration.[193,194] Note that the problem of an unknown additive constant also appears in case of single-ion solvation where all values are anchored on the elusive absolute intrinsic hydration free energy of the proton. Also, in that case, a range of interaction parameters consistent with what is known about pair properties in dilute solutions is obtained.[61] However, the advantage of calibrating against absolute intrinsic single-ion properties is the possibility to sequentially calibrate ion−solvent (infinite dilution regime) and ion−ion (finite concentration regime) interaction parameters.

In terms of the three choices, we conclude that neither the sampling protocol nor the free energy estimator constitute an obstacle in calculating the solvation free energies of ions. The choice of the Hamiltonian is classically reduced to the choice of the ion−solvent interaction parameters and the choice of the water model, possibly complicated by the need to use nonstandard mixing rules for ion−ion interactions.[61,183] Whether the water model is chosen to be polarizable depends on the purpose of the ion force field, that is, whether it should be applied in combination with a biomolecular force field that is calibrated for a specific water model. The consideration of the correction scheme represents a fourth choice in the simulation of ions in solution and allows to obtain solvation free energies which are independent of the method used to treat the electrostatic interactions. However, it is currently only available as an *ex post* scheme. As a result, ions parametrized in the context of this scheme might have the tendency to be undersolvated in practical applications such as using them as counterions in an MD simulation of a biomolecule. We note that while it is commonly taken for granted to choose the ionic partial charges equal to their formal (integer) charges in nonpolarizable force fields, an alternative suggestion has been made to use reduced charges of 0.7 e as a first-order approximation for electronic polarizability.[195−198]

5. SOLVATION OF POLYATOMIC IONS

In force-field calibration for monatomic ions, the only degree of freedom is usually the Lennard-Jones repulsion parameter, while the dispersion parameter can be derived on the basis of approximate formulas relating it to the ionic polarizability[61,169,199] and the ionic charge is either set to the formal charge or scaled by a factor of approximately 0.7 to implicitly account for the effect of electronic polarization.[195−198] For a given value of the standard absolute intrinsic hydration free energy of the proton, $\Delta G^{\ominus}_{hyd}[H^+]$, and a given water model, there is a unique relationship between the repulsion parameter and the experimentally accessible conventional hydration free energy. In contrast, the parametrization of polyatomic ions is an underdetermined problem; it allows for multiple solutions of similar quality with respect to the reproduction of a single experimental value. This is due to the multiplicity of atomic partial charges and Lennard-Jones repulsion parameters, as well as to the loss of a direct connection between dispersion parameters and atomic polarizabilities, the latter being ill-defined when considering (united) atoms within molecules.

Another difference to free energy calculations of monatomic ions is that the electrostatic interaction correction scheme[61,176,177] requires a numerical solution of the Poisson equation to obtain a continuum-electrostatics estimate for the charging free energy of the ionic solute in a macroscopic nonperiodic system with Coulombic electrostatic interactions and based on the experimental solvent permittivity, and in a periodic system with a specific electrostatic interaction scheme and based on the model solvent permittivity. These two calculations allow to account for the combined effect of approximate electrostatics errors, finite size errors, and deviations between experimental and model solvent permittivity. In the case of rigid ions, these calculations can be performed on the basis of a single structure. The improper summation errors can be accounted for in a way similar to the case of monatomic ions.

The latter scheme was proposed by Reif et al.[49] and used to recalibrate the nonbonded interaction parameters for the charged amino acid side chains in the GROMOS force field, based on ionic side-chain analogues. As for the monatomic ions, the parametrization was based on $\Delta G^{\ominus}_{hyd}[H^+] = -1100$ kJ mol^{-1}. The resulting GROMOS 54A8 force field is the first of its kind to contain nonbonded parameters for charged amino acid side chains that are derived in such a rigorously thermodynamic fashion. Subsequently, the force field was tested on structural properties of electrolyte solutions, lipid bilayers, and proteins.[200]

Note that the recently revised AMBER parameters[201] for phosphate ions are anchored to a different value of $\Delta G^{\ominus}_{hyd}[H^+]$ ($= -1052$ kJ mol^{-1}) compared to the GROMOS force field illustrating an additional difficulty when comparing force-field parameters for ions.

In terms of the three choices, we note that all the difficulties present in the context of monatomic ions also apply to polyatomic ions. Depending on the size and flexibility of the latter, sampling problems similar to the ones discussed in the context of neutral flexible molecules might appear.

6. NONCOVALENT BINDING

Molecular recognition forms the basis of virtually all biological processes. Understanding the interactions between proteins and their ligands is key to rationalize the molecular aspect of enzymatic processes and the mechanisms by which cellular systems integrate and respond to regulatory signals. From a medicinal perspective, there is great interest in the development of computer models capable of predicting accurately the strength of protein−ligand association[202] making the accurate computation of free energies of binding a key challenge for computer-aided drug design.[202−207] The main advantage of atomistic simulations over faster, empirical scoring functions is a more realistic inclusion of all thermodynamically relevant phenomena such as protein or/and ligand flexibility[208−211] and the possibility of the explicit inclusion of the solvent, which is usually necessary to account for the entropic contribution to the free energy. However, despite their potential, the effectiveness of atomistic simulations as predictive tools for protein−ligand binding remains uncertain.[10,212]

In contrast to the preceding sections, binding free energy calculations are usually affected by a mixture of inaccuracies in all three choices, that is, by errors in the force field, by insufficient sampling, and by the propagation of the former two effects by the free energy estimator. For sizable systems such as protein−ligand complexes, convergence may be hard to assess

because of slow degrees of freedom or rare events such as side-chain flipping at the binding site, both introducing considerable noise into the calculated free energy. It is therefore important to ask whether the time scale characteristic of the slowest degree of freedom is crucial for the free-energy change being estimated.[161]

The theoretical framework for calculating binding free energies in the realm of statistical mechanics is well established.[203] However, depending on the desired accuracy and available computational recources, different approaches can be used to calculate binding affinities. The state of the art for calculating binding free energies has been discussed in several recent reviews and perspective articles,[23,201,206,207,213−218] such that we limit our discussion to the three choices and some general challenges, which are independent of the methodology chosen.

Different choices for the three basic ingredients give rise to a multitude of methods each with a different trade-off between accuracy and computational efficiency. Three main classes of methods can be envisioned to calculate binding free energies, which are (i) end-point methods, (ii) methods based on the calculation of the free energy along a reaction coordinate, the potential of mean force (PMF), and (iii) methods based on the calculation of the free energy along a thermodynamic pathway employing alchemical transformations.

End-point and PMF methods are usually used to calculate the free energy change associated with the process that brings a ligand from an unbound state to bind to a receptor. This free energy change is often referred to as the *absolute* free energy of binding, while it is in fact a free energy difference. Here, we prefer the term *binding free energy* for this quantity and distinguish it from a *relative binding free energy*, which denotes the difference in binding free energies between two compounds. The calculation of relative binding free energies is a common application of alchemical free energy methods. However, these methods can also be used for the calculation of binding free energies, for example, in the context of the double-decoupling approach.[203] Comparison to experimental data requires the consideration of standard state corrections[203,219] leading to the *standard binding free energy*.

End-point methods such as the molecular mechanics/Poisson−Boltzmann surface area (MM/PBSA)[220,221] and molecular mechanics/generalized Born surface area(MM/GBSA)[220,222] only consider the two end states of interest, such as the protein free from and bound to a ligand, and calculate their absolute free energies. The binding free energy is then obtained by subtraction. Each of the free energies is decomposed into the mean enthalpic energy of the solute, the mean solute entropy, the polar solvation free energy, and the nonpolar solvation free energy. The entropic contributions are restricted to the conformational entropy and can be estimated from normal-mode analysis,[223−226] quasi-harmonic analysis,[225,227] and the restrain and release approach[225,228,229] or are simply neglected.[213,230−232] Inaccuracies inherent in these methods arise from a sensitivity toward the choice of the dielectric constant for the PB and GB calculations,[226] the use of implicit solvent,[233−235] inaccuracies in calculating entropic contributions,[229,236] and also from subtracting two large numbers (the approximate absolute free energies before and after binding), typically orders of magnitude larger than the binding free energy.[237] Recently, Silver et al.[238] proposed a novel end-point method based on uniform, rotameric enumeration of ligand torsional degrees of freedom to map out and explicitly integrate over the potential energy landscape. The method is structured around the use of the dead-end elimination and A* algorithms, which sort configurations by their energies and explicitly computes their contribution to the Boltzmann distribution. Linear interaction energy (LIE) is another end-point method that is based on the assumption that the free energy of binding shows a linear dependence on the polar (with parameter β) and nonpolar (α) changes in ligand-surroundings energies from MD averages.[239] Because LIE has several parameters, care should be taken not to overfit the data.[229] PDLD/S-LRA/β combines the semimicroscopic protein dipoles Langevin dipoles method, the linear response approximation (LRA), and the nonelectrostatic part of LIE. This approach seems to be more effective than MM-PBSA, LRA, or LIE.[229] An alternative end-point method, not based on molecular dynamics or Monte Carlo simulations, is the "mining minima" method,[240] which estimates the partition function through a harmonic approximation to the Hessian matrix. The exploration of minima local to the starting configuration, by transformations along low-frequency eigenvectors, allows for the inclusion of multiple possibly relevant states in the partition function estimate.

Methods based on the calculation of the free energy along a reaction coordinate are based on the definition of a pathway that connects the two states of interest and are reviewed, for example, in refs 241−243. The major challenges are to determine which degrees of freedom are important and how they participate in the reaction coordinate. An accurate reaction coordinate should convey the reaction mechanism, provide kinetically meaningful free energy surfaces, and facilitate calculations of the rate constant. For any reaction, the exact reaction coordinate is the committor probability,[244−246] the fraction of trajectories initiated from an atomic configuration $\vec{r}^N$ that commit to the product basin (B).[247−249] Unfortunately, the committor probability $p_B(\vec{r}^N)$ is not easy to compute and usually approximated in terms of collective variables that are functions of the configuration that compress many atomistic details into variables supposed to be physically important. The key challenge is to learn which collective variables are important and how they are involved in the reaction coordinate. For ion-pair dissociation, Mullen et al.[250] showed that dynamic recrossing of the dividing surface at the $p_B(\vec{r}^N) = 1/2$ isosurface is an inescapable consequence of dimensionality reduction to a single coordinate. An erroneous reduction of energy barriers by one-dimensional potentials of mean force was recently reported by Kopelevich[251] in the context of transport of a hydrophobic nanoparticle into a lipid bilayer. However, if calculation of the free energy of binding is the main purpose, one-dimensional PMFs combined with restraints, orthogonal to the direction of binding, can be an efficient approach.[252] It has also to be kept in mind that molecular recognition can be significantly more complex than a two-state process.[253] Recently, de Ruiter and Oostenbrink introduced the distance field (DF) as a reaction coordinate for the calculation of reversible protein−ligand binding.[254] DF is a grid-based method in which the shortest distance between the binding site and a ligand is determined avoiding routes that pass through the protein.

Methods based on the calculation of the free energy along a thermodynamic pathway employing alchemical transformations are widely used in the calculation of relative binding free energies. Apart from the established methods such as TI and staged FEP many other approaches including EDS,[255] BAR,[4] MBAR,[256] and λ-dynamics[17] have been tested on relevant

protein−ligand systems. In any case, errors due to insufficient sampling are more significant than the differences in the free-energy estimator.[257] Binding free energies can also be calculated by alchemical transformation, usually in the context of the double-decoupling method,[203,212,258] in which two separate calculations are carried out for decoupling the ligand from the solution and receptor environments. The latter process is often carried out using restraints to lock the noninteracting ligand into the binding pocket to enhance convergence of the free energy. The free-energy contribution associated with these restraints has to be taken into account in the thermodynamic cycle, for example by analytical approaches.[203,227,259] From a computational perspective the determination of the binding site volume in the presence of restraining potential energy terms in the Hamiltonian is crucial for obtaining reliable free-energy estimates.[206] Another alchemical method to calculate (absolute) binding free energies is the binding energy distribution analysis method (BEDAM), which is, in practice, only applicable using an implicit solvent.[260] It involves simulation of the ligand restricted to the protein binding site, but without interactions between the protein and the ligand. From this, the probability distribution for the binding energy is determined.

In the course of molecular design projects, end-point methods are most useful in the early, exploratory stages, while methods involving conformational sampling are most useful in later stages, when the goal is to optimize promising lead compounds.[203] Such a hierarchy of methods has been used recently to shed light on the molecular recognition of the coreceptor CXCR4 by the HIV-1 glycoprotein gp120.[261] Another approach is to combine the strength of different methods, such as LIE and one-step perturbation, in order to improve efficiency while maintaining accuracy.[262] When screening drug candidates prior to more elaborate free-energy calculations, it should be kept in mind, however, that MM-PBSA, in general, cannot be expected to reliably resolve differences within 12 kJ mol^{-1}.[213] The availability of automated workflows for the setup and analysis of binding free energy calculations[263−266] is expected to facilitate a more realistic evaluation of the different methods and how these could be combined in an efficient way to guide molecular design projects. However, to be acceptable for inclusion into work flows for lead-optimization, binding free energies have to be converged reliably with 4.2 kJ mol^{-1} variance error.[213]

Current challenges, which are of general nature and not bound to a particular method, are as follows:

(1) Multiple binding modes. Often multiple binding modes are of importance either due to different possible ligand orientations within the binding site,[209] or due to different conformations a ligand may adopt[267] or due to a combination of both effects.[268] These multiple orientations have to be sampled with the correct relative populations to avoid any bias. In the context of relative free-energy calculations, the two ligands to be compared may have distinct orientations giving rise to a slowly converging reorientation step in the calculation. Such cases can be treated more efficiently using the recently suggested "separated topologies" method.[269] This approach can also be used to determine the relative free energies of multiple orientations of the same ligand. However, also, the characterization of the ligand's unbound state can be of major relevance.[270] If a ligand has different conformations in the solvent but only one conformer binds to the protein, the free energy of focusing the different conformers to the one that binds to the protein has to be accounted for in binding free-energy calculations.[271,272] In the same line, the protein may also adopt several metastable states that contribute to the binding free energy. Considering only one conformational state neglects the free energy associated with confining the protein to that particular configuration.[273]

(2) Binding-site hydration. Upon binding, the ligand may replace several water molecules in the binding site. Depending on the free-energy change, these water molecules experience when leaving the site, this replacement might contribute (un)favorably to the binding free energy.[274] Several studies on model cavity−ligand systems[275−279] demonstrated that the role of water-mediated interactions and ligand dehydration can be far more relevant than the direct cavity−ligand interaction, owing to electrostatic screening and to entropic terms arising from solvent reorganization.[280,281] For proteins with solvent-exposed binding sites, these effects are difficult to capture with methods relying on implicit solvent models.[282,283] For a binding site deeply buried, the exchange of water molecules with the bulk region may be very slow, which may lead to convergence problems.[258] When estimating the effect of ligand modifications on binding free energies, it is crucial to take into account that water molecules are maintained for one variant but may be displaced for others.[284,285] To study which water molecules can favorably be replaced by a ligand, the free-energy difference between a water molecule and an apolar probe was calculated for a selection of water sites in the binding pockets of two proteins.[286] Such an analysis may give valuable insights for potency optimization in drug design.

(3) Definition of the bound state. The standard binding free energy depends on the definition of the bound state. If the binding is strong and specific, this does usually not pose a problem as long as the specific choice covers all important conformations of the complex.[203,260,287,288] For weak and less localized binding the dependence on the binding site volume would be noticeable. As pointed out by Mihailescu and Gilson,[288] the theoretical expression for the binding constant depends on the experimental technique used.

(4) Standard state correction. To convert calculated binding free energies into standard binding free energies, a correction term needs to be added that can be expressed in terms of the system volume or ligand concentration in the unbound state.[203,219] However, alternative approaches to estimate the required corrections in practice may differ significantly.[289]

(5) Finite size effects. Another underappreciated aspect are finite-size effects in binding free energy calculations. At present, finite-size effects on charging free energies are best understood in the context of the solvation of monatomic ions.[61,169,175−177,181,290] In this case, a numerical correction scheme[176] and a corresponding approximate analytical version[177] are available. The numerical version of this scheme has recently been extended to the case of small polyatomic ions[49] and to insertion of such ions into a simple model receptor, namely a functionalized C60 buckyball in water.[291]

However, considering the most general case of a complex polyatomic charged ligand inserted into a charged protein in solution, none of the schemes available at present[227,254,291−296] are sufficiently general, practical, and accurate. Recently, a new method for removing finite size effects has been proposed based on a continuum-electrostatics analysis. It requires performing Poisson−Boltzmann calculations on the protein−ligand system.[178] The approach introduces the concept of the *residual integrated potential* to account for the finite-size effect related to the solvent-excluded volume of the protein and the ligand, an effect that is absent in monatomic ion solvation.

(6) Force-field inaccuracy. Due to the hundreds of parameters involved in empirical force fields the propagation of errors in these parameters on calculated binding free energies is a complex problem. Recently, Rocklin et al.[297] investigated the sensitivity of binding free-energy calculations to the nonbonded energy parameters in force fields—atomic radii, dispersion well-depths, and partial charges—by performing tens of thousands of small parameter perturbations. They estimated that random, uncorrelated errors in force-field nonbonded parameters must be smaller than 0.02 e per charge, 0.06 Å per radius, and 0.04 kJ mol^{-1} per well depth in order to obtain 68% confidence that a computed binding affinity for a moderately sized lead compound will fall within 4.2 kJ mol^{-1} of the true affinity, if these are the only sources of error considered. Fixed charge models of ligands, parametrized against hydration free energies, might easily have larger uncertainty in the partial charges, especially in nonpolar binding sites.

7. CONFORMATIONAL CHANGES

The importance of knowing the change in free energy associated with a change in molecular conformation was already mentioned in the context of hydration or binding free energy calculations. It relies on the definition of conformational states as well as on the ability to define a reduced set of (spatial) coordinates $R(\vec{r}^N)$ on which the free energy is projected. Such a hypersurface is commonly called a reaction coordinate and, in configurational space, is a function of the positions of atoms in the system. Note that the term reaction *coordinate* is eventually associated with the minimum-free-energy pathway connecting the reference state to the target state but is commonly employed to characterize the order parameter along which the variation of the free energy is determined. The free energy as a function of the reaction coordinate $R'(\vec{r}^N)$, or the potential of mean force, is given by eq 5, where the term in curly brackets is the probability of finding the system lying on the reaction coordinate. Difficulties related to the representation of the reaction coordinate have been sketched in the previous section. A comprehensive discussion of methods to obtain reaction coordinates[298−300] is beyond the scope of the present review. Another difficulty is related to the definition of conformational states. For small systems such as carbohydrates, states can be relatively well-defined due to the rigid nature of the glycosidic linkage. Carbohydrates are therefore often used for testing new methodology[301,302] or to calibrate force fields.[303] Small peptides in solution show significantly more flexibility. Dipeptides are often used to test new methodology[304] or to investigate the pathway dependence of the free energy.[305] For larger (oligo)peptides the free-energy difference between different helical forms may depend on the definition of states. Relative free energies of 4, 0, and 12 kJ mol^{-1} for the π, α, and 3_{10} helical forms of a deca-alanine peptide in water were calculated by BS-LEUS using the sum of the two dihedral angles, $\phi_{n+1} + \psi_n$, encompassing the successive peptide bonds to define the helical states,[306] and values of 5, 0, and 47 kJ mol^{-1} were obtained by EDS using an RMSD criterion to define the helical states.[307] As is the definition of optimal reaction coordinates, the structure classification of biopolymeric structures is also an active field of research.[308,309] For proteins the characterization of the unfolded state represents a major obstacle in the calculation of folding free energies.[310]

A further difficulty in conformational free-energy calculations is the need to use enhanced sampling and biasing techniques[19,241,301,311−318] along with appropriate reweighting to the original Hamiltonian[319,320] for all but a few simple systems for which simple counting of configurations from a long MD trajectory works.[321]

In terms of the three choices, we conclude that conformational free energies are among the most challenging systems due to the many choices involved in setting up these calculations. As in the case of the choice of reaction coordinate, the free-energy difference will depend on the choice of definition of the conformational states. Next, if such conformational states are of high energy, the accuracy of the Hamiltonian is difficult to assess because conformations high in free energy are difficult to sample. Recently, EDS has been applied to solve this problem.[322] So, state definition and securing sufficient sampling are the main challenges when calculating free-energy differences between different conformational states.

8. CONCLUSION AND OUTLOOK

Although the two fundamental problems of inaccuracies in the Hamiltonian and of insufficient sampling are still prevalent, the calculation of free energy differences has seen some consolidation through the definition of best practices[166,167,323] and through a healthy skepticism toward the performance of computational models.[324] One of the great challenges remains the accurate calculation of the entropic contribution in molecular processes, which may or may not be a substantial portion of the free energy.[325] Some developments in the field that may lead to further consolidation are as follows:

(1) Evaluation and standards. Validation sets such as those provided at alchemistry.org,[326] in the binding database,[327] or through various blind tests[148,328] will help to generate a commonly accepted set of benchmark data useful for method and force-field development.

(2) Sensitivity analysis. Because free energies may be sensitive to the force-field parameters,[297] the use of efficient perturbation approaches may help to provide an estimate for the uncertainty due to an inaccurate Hamiltonian. The use of different free-energy methods for the same problem may give information regarding convergence and the sensitivity toward the free-energy estimator.

(3) Comparison of methods. Because different free-energy calculation methods use different information from the Hamiltonian, a comparison of methods on a pure theoretical basis may not always be sufficient to provide

practical recommendations. Comparison of methods in practical settings is therefore as important.[257,329−332]

(4) Interavailability of code. The transfer of a new method developed, implemented and tested for one particular MD software package to another one is far from being trivial because today's molecular dynamics packages are very complex pieces of software developed over decades often by a diverse group of contributors with different backgrounds and experience.[51] However, such transfer is essential for a wider acceptance of a particular method and also for a better comparison to other methods not implemented in the original software. Plug-ins with interfaces to different MD codes may help to disseminate new methods to a wider community of users.[333]

(5) Critical use of experimental data. Experimental measurements are invariably contaminated with errors, which may affect the maximally possible correlation between simulation and experiment that can be achieved.[324,334,335] Often modelers try too hard to reproduce experimental data as precisely as possible, ignoring the fact that these data are also subject to uncertainty.[334,336−338]

AUTHOR INFORMATION

Corresponding Author

*Phone: +49 711 685-66112. Fax: +49 711 685-66140. Email: hansen@itt.uni-stuttgart.de.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Philippe Hünenberger, Chris Oostenbrink, Alan Mark, and David Mobley for insightful discussions. This work was financially supported by the National Center of Competence in Research (NCCR) in Structural Biology, by grant No. 200020-137827 from the Swiss National Science Foundation, and by grant No. 228076 of the European Research Council. N.H. thanks the German Research Foundation (DFG) for financial support within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart.

REFERENCES

(1) *Free Energy Calculations: Theory and Applications in Chemistry and Biology*; Chipot, C., Pohorille, A., Eds.; Springer: Berlin, 2007.

(2) Landau, L. D.; Lifshitz, E. M. *Statistical Physics*; The Clarendon Press: Oxford, 1938.

(3) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420−1426.

(4) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245−268.

(5) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.

(6) Tuckerman, M. E. *Statistical Mechanics: Theory and Molecular Simulation*; Oxford University Press: Oxford, 2010.

(7) Hünenberger, P. H.; van Gunsteren, W. F. In *Computer Simulation of Biomolecular Systems*; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; Kluwer Academic: Dordrecht, The Netherlands, 1997; Vol. 3; pp 3−82.

(8) IUPAP. *Physica (Amsterdam)* **1978**, *93A*, 1−60.

(9) IUPAC, *Quantities, Units and Symbols in Physical Chemistry*; Blackwell Scientific Publications: Oxford, 1988.

(10) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150−160.

(11) Mezei, M.; Beveridge, D. L. *Ann. N.Y. Acad. Sci.* **1986**, *482*, 1−23.

(12) van Gunsteren, W. F. In *Computer Simulation of Biomolecular Systems, Theoretical and Experimental Applications*; van Gunsteren, W. F., Weiner, P. K., Eds.; Escom Science Publishers: Leiden, The Netherlands, 1989; pp 27−59.

(13) Frenkel, D. In *Computer simulations in Materials Science*; Meyer, M., Pontikis, V., Eds.; NATO ASI Series E205; Kluwer Academic Publishers: Dordrecht, 1991; pp 85−117.

(14) Straatsma, T. P.; McCammon, J. A. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407−435.

(15) Mark, A. E. In *Encyclopedia of Computational Chemistry*; von Ragué Schleyer, P., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; John Wiley & Sons: Chichester, 1998; Vol. 2; pp 1070−1083.

(16) Chipot, C.; Pearlman, D. A. *Mol. Sim.* **2002**, *28*, 1−12.

(17) Knight, J. L.; Brooks, C. L., III *J. Comput. Chem.* **2009**, *30*, 1692−1700.

(18) Christ, C. D.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2010**, *31*, 1569−1582.

(19) Kästner, J. *WIREs Comput. Mol. Sci.* **2011**, *1*, 932−942.

(20) Singh, S.; Chopra, M.; de Pablo, J. J. *Annu. Rev. Chem. Biomol. Eng.* **2012**, *3*, 369−394.

(21) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431−492.

(22) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395−2417.

(23) Chipot, C. *WIREs Comput. Mol. Sci.* **2014**, *4*, 71−89.

(24) Straatsma, T. P.; Berendsen, H. J. C.; Postma, J. P. M. *J. Chem. Phys.* **1986**, *85*, 6720−6727.

(25) MacKerell, A. D.; Wiórkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.

(26) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(27) Feller, S. E.; MacKerell, A. D. *J. Phys. Chem. B* **2000**, *104*, 7510−7515.

(28) Hatcher, E.; Guvench, O.; MacKerell, A. D. *J. Phys. Chem. B* **2009**, *113*, 12466−12476.

(29) Brooks, B. R.; et al. *J. Comput. Chem.* **2009**, *30*, 1545−1614.

(30) Weiner, P. K.; Kollman, P. A. *J. Comput. Chem.* **1981**, *2*, 287−303.

(31) Pearlman, D. A.; Case, D. A.; Caldwell, J. D.; Ross, W. S.; Cheatham, T. E., III; DeBolt, S.; Fergusson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1−41.

(32) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(33) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outerino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2008**, *29*, 622−655.

(34) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810−1819.

(35) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.

(36) Damm, W.; Frontera, A.; Rivado-Rives, J. *J. Comput. Chem.* **1997**, *18*, 1955−1970.

(37) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.

(38) Price, M. L. P.; Ostrovsky, D.; Jorgensen, W. L. *J. Comput. Chem.* **2001**, *22*, 1340−1352.

(39) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Vdf Hochschulverlag AG an der ETH Zürich: Zürich, Groningen, 1996.

(40) Daura, X.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **1998**, *19*, 535−547.

(41) van Gunsteren, W. F.; Daura, X.; Mark, A. E. In *Encyclopedia of Computational Chemistry*; Schleyer, P., Ed.; John Wiley & Sons: Chichester, Vol. 2; pp 1211−1216.

(42) Schuler, L. D.; Daura, X.; van Gunsteren, W. F. *J. Comput. Chem.* **2001**, *22*, 1205−1218.

(43) Chandrasekhar, I.; Kastenholz, M. A.; Lins, R. D.; Oostenbrink, C.; Schuler, L. D.; Tieleman, D.; van Gunsteren, W. F. *Eur. Biophys. J.* **2003**, *32*, 67−77.

(44) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656−1676.
(45) Soares, T. A.; Hünenberger, P. H.; Kastenholz, M. A.; Kräutler, V.; Lenz, T.; Lins, R. D.; Oostenbrink, C.; van Gunsteren, W. F. *J. Comput. Chem.* **2005**, *26*, 725−737.
(46) Lins, R. D.; Hünenberger, P. *J. Comput. Chem.* **2005**, *26*, 1400−1412.
(47) Poger, D.; van Gunsteren, W. F.; Mark, A. E. *J. Comput. Chem.* **2010**, *31*, 1117−1125.
(48) Schmid, N.; Eichenberger, A.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. *Eur. Biophys. J.* **2011**, *40*, 843−856.
(49) Reif, M. M.; Hünenberger, P. H.; Oostenbrink, C. *J. Chem. Theory Comput.* **2012**, *8*, 3705−3723.
(50) Ponder, J. W.; Case, D. A. *Adv. Prot. Chem.* **2003**, *66*, 27−85.
(51) van Gunsteren, W. F.; Dolenc, J. *Mol. Sim.* **2012**, *38*, 1271−2181.
(52) Jorgensen, W. L.; Laird, E. R.; Nguyen, T. B.; Tirado-Rives, J. *J. Comput. Chem.* **1993**, *14*, 206−215.
(53) Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. L. *J. Phys. Chem.* **1994**, *98*, 13077−13082.
(54) Jorgensen, W. L.; Tirado-Rives, J. *Perspect. Drug Discovery Des.* **1995**, *3*, 123−138.
(55) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.
(56) Chipot, C. *J. Comput. Chem.* **2003**, *24*, 409−415.
(57) Udier-Blagović, M.; De Tirado, P. M.; Pearlman, S. A.; Jorgensen, W. L. *J. Comput. Chem.* **2004**, *25*, 1322−1332.
(58) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. *J. Chem. Theory Comput.* **2010**, *6*, 1509−1519.
(59) Shivakumar, D.; Harder, E.; Damm, W.; Friesner, R. A.; Sherman, W. *J. Chem. Theory Comput.* **2012**, *8*, 2553−2558.
(60) Horta, B. A. C.; Fuchs, P. F. J.; van Gunsteren, W. F.; Hünenberger, P. H. *J. Chem. Theory Comput.* **2011**, *7*, 1016−1031.
(61) Reif, M. M.; Hünenberger, P. H. *J. Chem. Phys.* **2011**, *134*, 144104.
(62) Hansen, N.; Kraus, P.; Saßmannshausen, H.; Timmerscheidt, T.; van Gunsteren, W. F. *Mol. Phys.* **2011**, *109*, 2593−2605.
(63) Petrov, D.; Margreitter, C.; Grandits, M.; Oostenbrink, C.; Zagrovic, B. *PLoS Comput. Biol.* **2013**, *9*, e1003154.
(64) Margreitter, C.; Petrov, D.; Zagrovic, B. *Nucleic Acids Res.* **2013**, *41*, W422−W426.
(65) Khoury, G. A.; Thompson, J. P.; Smadbeck, J.; Kieslich, C. A.; Floudas, C. A. *J. Chem. Theory Comput.* **2013**, *9*, 5653−5674.
(66) Mobley, D. L.; Dill, K. A.; Chodera, J. D. *J. Phys. Chem. B* **2008**, *112*, 938−946.
(67) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. *J. Chem. Theory Comput.* **2009**, *5*, 350−358.
(68) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25*, 247−260.
(69) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(70) Mobley, D. L.; Liu, S.; Cerutti, D. S.; Swope, W. C.; Rice, J. E. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 551−562.
(71) Shivakumar, D.; Deng, Y.; Roux, B. *J. Chem. Theory Comput.* **2009**, *5*, 919−930.
(72) Momany, F. A.; Rone, R. *J. Comput. Chem.* **1992**, *13*, 888−900.
(73) Knight, J. L.; Yesselman, J. D.; Brooks, C. L., III *J. Comput. Chem.* **2013**, *34*, 893−903.
(74) Vanommeslaeghe, K.; MacKerell, A. D. *J. Chem. Inf. Model.* **2012**, *52*, 3144−3154.
(75) Vanommeslaeghe, K.; Prabhu Raman, E.; MacKerell, A. D. *J. Chem. Inf. Model.* **2012**, *52*, 3155−3168.
(76) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell, A. D. *J. Comput. Chem.* **2010**, *31*, 671−690.
(77) Yesselman, J. D.; Price, D. J.; Knight, J. L.; Brooks, C. L. *J. Comput. Chem.* **2012**, *33*, 189−202.
(78) Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. *J. Comput. Chem.* **2011**, *32*, 2359−2368.
(79) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(80) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132−146.
(81) Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623−1641.
(82) Cramer, C. J.; Truhlar, D. G. In *Free Energy Calculations in Rational Drug Design*; Reddy, M. R., Erion, M. D., Eds.; Kluwer: New York, 2001; pp 63−95.
(83) Cramer, C. J.; Truhlar, D. G. In *Trends and Perspectives in Modern Computational Science*; Maroulis, G., Simons, T., Eds.; Brill/VSP: Leiden, 2006; pp 112−140.
(84) Knight, J. L.; Brooks, C. L., III *J. Comput. Chem.* **2011**, *32*, 2909−2922.
(85) Huang, L.; Roux, B. *J. Chem. Theory Comput.* **2013**, *9*, 3543−3556.
(86) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 134508.
(87) Wolf, M. G.; Groenhof, G. *J. Comput. Chem.* **2012**, *33*, 2225−2232.
(88) Hess, B.; van der Vegt, N. F. A. *J. Phys. Chem. B* **2006**, *110*, 17616−17626.
(89) Nicholls, A.; Mobley, D. L.; Guthrie, J. P.; Chodera, J. D.; Bayly, C. I.; Cooper, M. D.; Pande, V. S. *J. Med. Chem.* **2008**, *51*, 769−779.
(90) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Dill, K. A. *J. Phys. Chem. B* **2009**, *113*, 4533−4537.
(91) Ahmed, A.; Sandler, S. I. *J. Chem. Theory Comput.* **2013**, *9*, 2774−2785.
(92) Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D. *J. Chem. Theory Comput.* **2012**, *8*, 61−74.
(93) Fennell, C. F.; Wymer, K. L.; Mobley, D. L. *J. Phys. Chem. B* **2014**, DOI: 10.1021/jp411529h.
(94) Delaney, J. *Drug Discovery Today* **2005**, *10*, 289−295.
(95) Johnson, S. R.; Chen, X.-Q.; Murphy, D.; Gudmundsson, O. *Mol. Pharmaceutics* **2007**, *4*, 513−523.
(96) Okumura, H.; Yonezawa, F. *J. Chem. Phys.* **2000**, *113*, 9162−9168.
(97) Potoff, J. J.; Bernard-Brunel, D. A. *J. Phys. Chem. B* **2009**, *113*, 14725−14731.
(98) Warshel, A. *J. Chem. Phys.* **1970**, *53*, 582−594.
(99) White, D. N. J. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 517−521.
(100) Sun, H. *J. Phys. Chem. B* **1998**, *102*, 7338−7364.
(101) Halgren, T. A. *J. Am. Chem. Soc.* **1992**, *114*, 7827−7843.
(102) Jämbeck, J. P. M.; Lyubartsev, A. P. *J. Phys. Chem. B* **2014**, *118*, 3793−3804.
(103) Cisneros, C. A.; Karttunen, M.; Ren, P.; Sagui, C. *Chem. Rev.* **2014**, *114*, 779−814.
(104) Harder, E.; Anisimov, V. M.; Whitfield, T.; MacKerell, A. D.; Roux, B. *J. Phys. Chem. B* **2008**, *112*, 3509−3521.
(105) Nerenberg, P. S.; Jo, B.; So, C.; Tripathy, A.; Head-Gordon, T. *J. Phys. Chem. B* **2012**, *116*, 4524−4534.
(106) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struc., Funct., Bioinf.* **2006**, *65*, 712−725.
(107) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665−9678.
(108) McGuire, R. F.; Momany, F. A.; Scheraga, H. A. *J. Phys. Chem.* **1972**, *76*, 375−393.
(109) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361−2381.
(110) Bachmann, S. J.; Dolenc, J.; van Gunsteren, W. F. *Mol. Phys.* **2013**, *111*, 2334−2344.
(111) Paliwal, H.; Shirts, M. R. *J. Chem. Theory Comput.* **2013**, *9*, 4700−4717.
(112) Paliwal, H.; Shirts, M. R. *J. Chem. Phys.* **2013**, *138*, 154108.
(113) Garrido, N. M.; Queimada, A. J.; Jorge, M.; Macedo, E. A.; Economou, I. G. *J. Chem. Theory Comput.* **2009**, *5*, 2436−2446.

(114) Yu, H.; van Gunsteren, W. F. *Comput. Phys. Commun.* **2005**, *172*, 69−85.
(115) Klimovich, P. V.; Mobley, D. L. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 307−316.
(116) Paluch, A. S.; Shah, J. K.; Maginn, E. J. *J. Chem. Theory Comput.* **2011**, *7*, 1394−1403.
(117) Fuchs, P. F. J.; Hansen, H. S.; Hünenberger, P. H.; Horta, B. A. C. *J. Chem. Theory Comput.* **2012**, *8*, 3943−3963.
(118) Escobedo, F. A.; de Pablo, J. J. *J. Chem. Phys.* **1995**, *103*, 2703−2710.
(119) Abreu, C. R. A.; Escobedo, F. A. *J. Chem. Phys.* **2006**, *124*, 054116.
(120) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300−313.
(121) Åberg, K. M.; Lyubartsev, A. P.; Jacobsson, S. P.; Laaksonen, A. *J. Chem. Phys.* **2004**, *120*, 3770−3776.
(122) Paluch, A. S.; Mobley, D. L.; Maginn, E. J. *J. Chem. Theory Comput.* **2011**, *7*, 2910−2918.
(123) Yang, L.; Ahmed, A.; Sandler, S. I. *J. Comput. Chem.* **2013**, *34*, 284−293.
(124) Wu, P.; Hu, X.; Yang, W. *J. Phys. Chem. Lett.* **2011**, *2*, 2099−2103.
(125) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529−539.
(126) Buelens, F. P.; Grubmüller, H. *J. Comput. Chem.* **2012**, *33*, 25−33.
(127) Naden, L. N.; Shirts, M. R. *J. Chem. Theory Comput.* **2014**, *10*, 1128−1149.
(128) Darve, E.; Pohorille, A. *J. Chem. Phys.* **2001**, *115*, 9169−9183.
(129) Bhatnagar, N.; Kamath, G.; Chelst, I.; Potoff, J. J. *J. Chem. Phys.* **2012**, *137*, 014502.
(130) Mobley, D. L.; Barber, A. E., II; Fennel, C. J.; Dill, K. A. *J. Phys. Chem. B* **2008**, *112*, 2405−2414.
(131) Mobley, D. L.; Baker, J. R.; Barber, A. E., II; Fennel, C. J.; Dill, K. A. *J. Phys. Chem. B* **2011**, *115*, 15145−15145.
(132) Purisima, E. O.; Sulea, T. *J. Phys. Chem. B* **2009**, *113*, 8206−8209.
(133) Mukhopadhyay, A.; Fenley, A. T.; Tolokh, I. S.; Onufriev, A. V. *J. Phys. Chem. B* **2012**, *116*, 9776−9783.
(134) Bardhan, J. P.; Jungwirth, P.; Makowski, L. *J. Chem. Phys.* **2012**, *137*, 124101.
(135) Fennell, C. J.; Kehoe, C.; Dill, K. A. *J. Am. Chem. Soc.* **2010**, *132*, 234−240.
(136) Ferguson, A. L.; Debenedetti, P. G.; Panagiotopoulos, A. Z. *J. Phys. Chem. B* **2009**, *113*, 6405−6414.
(137) Bhatnagar, N.; Kamath, G.; Potoff, J. J. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6467−6474.
(138) Ahmed, A.; Sandler, S. I. *J. Chem. Theory Comput.* **2013**, *9*, 2389−2397.
(139) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. *Nat. Rev. Drug Discovery* **2008**, *7*, 608−624.
(140) Kiss, L. E.; Ferreira, H. S.; Torrão, L.; Bonifácio, M. J.; Palma, P. N.; Soares-da-Silva, P.; Learmonth, D. A. *J. Med. Chem.* **2010**, *53*, 3396−3411.
(141) Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C. *Biochemistry* **2012**, *51*, 4990−5003.
(142) Mallinson, J.; Collins, I. *Future Med. Chem.* **2012**, *4*, 1409−1438.
(143) YASARA AutoSMILES Server. www.yasara.org/autosmilesserver.htm (accessed Jan. 14, 2014).
(144) Miller, B. T.; Singh, R. P.; Klauda, J. B.; Hodoscek, M.; Brooks, B. R.; Woodcock, H. L. *J. Chem. Inf. Model.* **2008**, *48*, 1920−1929.
(145) Ribeiro, A.; Horta, B. A. C.; de Alencastro, R. B. *J. Braz. Chem. Soc.* **2008**, *19*, 1433−1435.
(146) Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. *J. Chem. Theory Comput.* **2011**, *7*, 4026−4037.
(147) Koziara, K. B.; Stroet, M.; Malde, A. K.; Mark, A. E. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 221−233.
(148) Guthrie, J. P. *J. Phys. Chem. B* **2009**, *113*, 4501−4507.
(149) Ruddigkeit, L.; van Deursen, R.; Blum, L. C.; Reymond, J.-L. *J. Chem. Inf. Model.* **2012**, *52*, 2864−2875.
(150) Leitgeb, M.; Schröder, C.; Boresch, S. *J. Chem. Phys.* **2005**, *122*, 084109.
(151) Palma, P. N.; Bonifácio, M. J.; Loureiro, A. I.; Soares-da-Silva, P. *J. Comput. Chem.* **2012**, *33*, 970−986.
(152) Garate, J. A.; Oostenbrink, C. *J. Comput. Chem.* **2013**, *34*, 1398−1408.
(153) Arrar, M.; de Oliveira, C. A. F.; Fajer, M.; Sinko, W.; McCammon, J. A. *J. Chem. Theory Comput.* **2013**, *9*, 18−23.
(154) Hritz, J.; Oostenbrink, C. *J. Phys. Chem. B* **2009**, *113*, 12711−12720.
(155) Khavrutskii, I. V.; Wallqvist, A. *J. Chem. Theory Comput.* **2010**, *6*, 3427−3441.
(156) Paluch, A. S.; Cryan, D. D.; Maginn, E. J. *J. Chem. Eng. Data* **2011**, *56*, 1587−1595.
(157) Acree, W. E.; Abraham, M. H. *Can. J. Chem.* **2001**, *79*, 1466−1476.
(158) König, G.; Bruckner, S.; Boresch, S. *Biophys. J.* **2013**, *104*, 453−462.
(159) Garrido, N. M.; Jorge, M.; Queimada, A. J.; Economou, I. G.; Macedo, E. A. *Fluid Phase Equilib.* **2010**, *289*, 148−155.
(160) Canzar, S.; El-Kebir, M.; Pool, R.; Elbassioni, K.; Malde, A. K.; Mark, A. E.; Geerke, D. P.; Stougie, L.; Klau, G. W. *J. Comput. Biol.* **2013**, *20*, 188−198.
(161) Chipot, C. In *New Algorithms for Macromolecular Simulation*; Leimkuhler, B., Chipot, C., Elber, R., Laaksonen, A., Mark, A. E., Schlick, T., Schütte, C., Skeel, R., Eds.; Springer: Berlin, 2006; pp 185−211.
(162) Zuckerman, D. M.; Woolf, T. B. *J. Stat. Phys.* **2004**, *114*, 1303−1323.
(163) Ytreberg, F. M.; Zuckerman, D. M. *J. Comput. Chem.* **2004**, *25*, 1749−1759.
(164) Lyman, E.; Zuckerman, D. M. *J. Phys. Chem. B* **2007**, *111*, 12876−12882.
(165) Murdock, S. E.; Tai, K.; Ng, M. H.; Johnston, S.; Wu, B.; Fangohr, H.; Laughton, C. A.; Essex, J. W.; Sansom, M. S. P. *J. Chem. Theory Comput.* **2006**, *2*, 1477−1481.
(166) Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235−10253.
(167) Shirts, M. R.; Mobley, D. L. In *Biomolecular Simulations: Methods and Protocols*; Monticelli, L., Salonen, E., Eds.; Springer: New York, 2013; pp 271−311.
(168) Shirts, M. R. *J. Chem. Theory Comput.* **2013**, *9*, 909−926.
(169) Hünenberger, P. H.; Reif, M. M. *Single-ion solvation: Experimental and Theoretical Approaches to Elusive Thermodynamic Quantities*, 1st ed.; Royal Society of Chemistry: London, 2011.
(170) Warshel, A. *J. Phys. Chem.* **1982**, *86*, 2218−2224.
(171) Lybrand, T. P.; Ghosh, I.; McCammon, J. A. *J. Am. Chem. Soc.* **1985**, *107*, 7793−7794.
(172) Warshel, A.; Sussman, F.; King, G. *Biochemistry* **1986**, *25*, 8368−8372.
(173) Straatsma, T. P.; Berendsen, H. J. C. *J. Chem. Phys.* **1988**, *89*, 5876−5886.
(174) Hummer, G.; Pratt, L. R.; Garcia, A. E. *J. Chem. Phys.* **1997**, *107*, 9275−9277.
(175) Kastenholz, M. A.; Hünenberger, P. H. *J. Chem. Phys.* **2006**, *124*, 124106.
(176) Kastenholz, M. A.; Hünenberger, P. H. *J. Chem. Phys.* **2006**, *124*, 224501.
(177) Reif, M. M.; Hünenberger, P. H. *J. Chem. Phys.* **2011**, *134*, 144103.
(178) Rocklin, G. J.; Mobley, D. L.; Dill, K. A.; Hünenberger, P. H. *J. Chem. Phys.* **2013**, *139*, 184103.
(179) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullmann, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981; pp 331−342.
(180) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269−6271.

(181) Dahlgren, B.; Reif, M. M.; Hünenberger, P. H.; Hansen, N. *J. Chem. Theory Comput.* **2012**, *8*, 3542−3564.
(182) Deublein, S.; Vrabec, J.; Hasse, H. *J. Chem. Phys.* **2012**, *136*, 084501.
(183) Fyta, M.; Netz, R. R. *J. Chem. Phys.* **2012**, *136*, 124103.
(184) Trumm, M.; Martnez, Y. O. G.; Réal, F.; Masella, M.; Vallet, V.; Schimmelpfennig, B. *J. Chem. Phys.* **2012**, *136*, 044509.
(185) Mamatkulov, S.; Fyta, M.; Netz, R. R. *J. Chem. Phys.* **2013**, *138*, 024505.
(186) Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, K. M., Jr. *J. Chem. Theory Comput.* **2013**, *9*, 2733−2748.
(187) Orabi, E. A.; Lamoureux, G. *J. Chem. Theory Comput.* **2013**, *9*, 2324−2338.
(188) Sedlmeier, F.; Netz, R. R. *J. Chem. Phys.* **2013**, *138*, 115101.
(189) Saxena, A.; Sept, D. *J. Chem. Theory Comput.* **2013**, *9*, 3538−3542.
(190) Moučka, F.; Nezbeda, I.; Smith, W. R. *J. Chem. Phys.* **2013**, *138*, 154102.
(191) Moučka, F.; Nezbeda, I.; Smith, W. R. *J. Chem. Theory Comput.* **2013**, *9*, 5076−5085.
(192) Reif, M. M.; Oostenbrink, C. *Personal communication*, May 13, 2014.
(193) Eisenberg, B. Life's solutions are complex fluids. 2012; http://arxiv.org/abs/1207.4737.
(194) Eisenberg, B. *Biophys. J.* **2013**, *104*, 1849−1866.
(195) Leontyev, I. V.; Stuchebrukhov, A. A. *J. Chem. Phys.* **2009**, *130*, 085102.
(196) Leontyev, I. V.; Stuchebrukhov, A. A. *J. Chem. Theory Comput.* **2010**, *6*, 3153−3161.
(197) Leontyev, I. V.; Stuchebrukhov, A. A. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2613−2626.
(198) Leontyev, I. V.; Stuchebrukhov, A. A. *J. Chem. Theory Comput.* **2012**, *8*, 3207−3216.
(199) Slater, J. C.; Kirkwood, J. G. *Phys. Rev.* **1931**, *37*, 682−697.
(200) Reif, M. M.; Winger, M.; Oostenbrink, C. *J. Chem. Theory Comput.* **2013**, *9*, 1247−1264.
(201) Steinbrecher, T.; Latzer, J.; Case, D. A. *J. Chem. Theory Comput.* **2012**, *8*, 4405−4412.
(202) Jorgensen, W. L. *Science* **2004**, *303*, 1813−1818.
(203) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047−1069.
(204) Gohlke, H.; Klebe, G. *Angew. Chem., Int. Ed.* **2002**, *41*, 2645−2676.
(205) Jorgensen, W. L. *Acc. Chem. Res.* **2009**, *42*, 724−733.
(206) Gallicchio, E.; Levy, R. M. *Adv. Prot. Chem. Struct. Biol.* **2011**, *85*, 27−80.
(207) Wereszczynski, J.; McCammon, J. A. *Q. Rev. Biophys.* **2012**, *45*, 1−25.
(208) Carlson, H. A. *Curr. Pharm. Des.* **2002**, *8*, 1571−1578.
(209) Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. *J. Mol. Biol.* **2007**, *371*, 1118−1134.
(210) Baron, R.; McCammon, J. A. *ChemPhysChem* **2008**, *9*, 983−988.
(211) Lawrenz, M.; Baron, R.; Wang, Y.; McCammon, J. A. *J. Chem. Theory Comput.* **2011**, *7*, 2224−2232.
(212) Mobley, D. L.; Dill, K. A. *Structure* **2009**, *17*, 489−498.
(213) Shirts, M. R.; Mobley, D. L.; Brown, S. P. In *Drug Design: Structure- and Ligand-Based Approaches*; Merz, K. M., Ringe, D., Reynolds, C. H., Eds.; Cambridge University Press: New York, 2010; pp 61−86.
(214) Michel, J.; Essex, J. W. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 639−658.
(215) de Ruiter, A.; Oostenbrink, C. *Curr. Opin. Chem. Biol.* **2011**, *15*, 547−552.
(216) Homeyer, N.; Golke, H. In *In Silico Drug Discovery and Design*; Lill, M. A., Ed.; Future Science: London, U.K., 2013; pp 50−63.
(217) Baron, R.; McCammon, J. A. *Annu. Rev. Phys. Chem.* **2013**, *64*, 151−175.
(218) Westermaier, Y.; Hubbard, R. E. In *De Novo Molecular Design*; Schneider, G., Ed.; Wiley-VCH: Weinheim, Germany, 2014; pp 373−415.
(219) General, I. J. *J. Chem. Theory Comput.* **2010**, *6*, 2520−2524.
(220) Srinivasan, J.; Cheatham, T. E., III; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 9401−9409.
(221) Homeyer, N.; Gohlke, H. *Mol. Inf.* **2012**, *31*, 114−122.
(222) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III *Acc. Chem. Res.* **2000**, *33*, 889−897.
(223) Sadiq, S. K.; Wright, D. W.; Kenway, O. A.; Coveney, P. V. *J. Chem. Inf. Model.* **2010**, *50*, 890−905.
(224) Hu, G.; Wang, D.; Liu, X.; Zhang, Q. *J. Comput.-Aided. Mol. Des.* **2010**, *24*, 687−697.
(225) Singh, N.; Warshel, A. *Proteins* **2010**, *78*, 1724−1735.
(226) Hou, T.; Wang, J.; Li, Y.; Wang, W. *J. Chem. Inf. Model.* **2011**, *51*, 69−82.
(227) Woo, H.-J.; Roux, B. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6825−6830.
(228) Strajbl, M.; Sham, Y. Y.; Villa, J.; Chu, Z.-T.; Warshel, A. *J. Phys. Chem. B* **2000**, *104*, 4578−4584.
(229) Singh, N.; Warshel, A. *Proteins* **2010**, *78*, 1705−1723.
(230) Chong, L. T.; Pitera, J. W.; Swope, W. C.; Pande, V. S. *J. Mol. Graph. Modell.* **2009**, *27*, 978−982.
(231) Hu, R.; Barbault, F.; Maurel, F.; Delamar, M.; Zhang, R. *Chem. Biol. Drug Des.* **2010**, *76*, 518−526.
(232) Vorontsov, I. I.; Miyashita, O. *J. Comput. Chem.* **2011**, *32*, 1043−1053.
(233) Weis, A.; Katebzadeh, K.; Söderhjelm, P.; Nilsson, I.; Ryhde, U. *J. Med. Chem.* **2006**, *49*, 6596−6606.
(234) Mitomo, D.; Fukunishi, Y.; Higo, J.; Nakamura, H. *Genome Inform.* **2009**, *23*, 85−97.
(235) Godschalk, F.; Genheden, S.; Söderhjelm, P.; Ryde, U. *Phys. Chem. Chem. Phys.* **2013**, *15*, 7731−7739.
(236) Guimarães, C. R. W.; Mathiowetz, A. M. *J. Chem. Inf. Model.* **2010**, *50*, 547−559.
(237) Barril, X.; Gelpi, J. L.; Lopez, J. M.; Orozco, M.; Luque, F. J. *Theor. Chem. Acc.* **2001**, *106*, 2−9.
(238) Silver, N. W.; King, B. M.; Nalam, M. N. L.; Cao, H.; Ali, A.; Reddy, G. S. K. K.; Rana, T. M.; Schiffer, C. A.; Tidor, B. *J. Chem. Theory Comput.* **2013**, *9*, 5098−5115.
(239) Åqvist, J.; Medina, C.; Samuelsson, J.-E. *Protein Eng.* **1994**, *7*, 385−391.
(240) Chen, W.; Gilson, M. K.; Webb, S. P.; Potter, M. J. *J. Chem. Theory Comput.* **2010**, *6*, 3540−3557.
(241) Darve, E. In *New Algorithms for Macromolecular Simulation*; Leimkuhler, B., Chipot, C., Elber, R., Laaksonen, A., Mark, A. E., Schlick, T., Schütte, C., Skeel, R., Eds.; Springer: Berlin, 2006; pp 213−249.
(242) Trzesniak, D.; Kunz, A. P.; van Gunsteren, W. F. *ChemPhysChem* **2007**, *8*, 162−169.
(243) Li, W.; Ma, A. *Mol. Sim.* **2014**, DOI: 10.1080/08927022.2014.907898.
(244) Ma, A.; Dinner, A. R. *J. Phys. Chem. B* **2005**, *109*, 6769−6779.
(245) Best, R. B.; Chen, Y.-G.; Hummer, G. *Structure* **2005**, *13*, 1755−1763.
(246) E, W.; Ren, W.; Vanden-Eijnden, E. *Chem. Phys. Lett.* **2005**, *413*, 242−247.
(247) Du, R.; Pande, V. S.; Grosberg, A. Y.; Tanaka, T.; Shaknovich, E. *J. Chem. Phys.* **1998**, *108*, 334−350.
(248) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.
(249) Dellago, C.; Bolhuis, P. G.; Geissler, P. L. *Adv. Chem. Phys.* **2002**, *123*, 1−78.
(250) Mullen, R. G.; Shea, J.-E.; Peters, B. *J. Chem. Theory Comput.* **2014**, *10*, 659−667.
(251) Kopelevich, D. I. *J. Chem. Phys.* **2013**, *139*, 134906.

(252) Doudou, S.; Burton, N. A.; Henchman, R. H. *J. Chem. Theory Comput.* **2009**, *5*, 909−918.
(253) Shan, Y.; Kim, E. T.; Eastwood, M. P.; Dror, R. O.; Seeliger, M. A.; Shaw, D. E. *J. Am. Chem. Soc.* **2011**, *133*, 9181−9183.
(254) de Ruiter, A.; Oostenbrink, C. *J. Chem. Theory Comput.* **2013**, *9*, 883−892.
(255) Christ, C. D.; van Gunsteren, W. F. *J. Chem. Phys.* **2007**, *126*, 184110.
(256) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105.
(257) Christ, C. D.; Fox, T. *J. Chem. Inf. Model.* **2014**, *54*, 108−120.
(258) Deng, Y.; Roux, B. *J. Phys. Chem. B* **2009**, *113*, 2234−2246.
(259) Boresch, S.; Tettinger, F.; Karplus, M. *J. Phys. Chem. B* **2003**, *107*, 9535−9551.
(260) Gallicchio, E.; Lapelosa, M.; Levy, R. M. *J. Chem. Theory Comput.* **2010**, *6*, 2961−2977.
(261) Tamamis, P.; Floudas, C. A. *Biophys. J.* **2013**, *105*, 1502−1514.
(262) de Ruiter, A.; Oostenbrink, C. *J. Chem. Theory Comput.* **2012**, *8*, 3686−3695.
(263) Liu, P.; Dehez, F.; Cai, W.; Chipot, C. *J. Chem. Theory Comput.* **2012**, *8*, 26062616.
(264) Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. *J. Chem. Theory Comput.* **2012**, *8*, 3314−3321.
(265) Homeyer, N.; Gohlke, H. *J. Comput. Chem.* **2013**, *34*, 965−973.
(266) Liu, S.; Wu, Y.; Lin, T.; Abel, R.; Redmann, J. P.; Summa, C. M.; Jaber, V. R.; Lim, N. M.; Mobley, D. L. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 755−770.
(267) Stockwell, G. R.; Thornton, J. M. J. *Mol. Biol.* **2006**, *356*, 928−944.
(268) Stjernschantz, E.; Oostenbrink, C. *Biophys. J.* **2010**, *98*, 2682−2691.
(269) Rocklin, G. J.; Mobley, D. L.; Dill, K. A. *J. Chem. Phys.* **2013**, *138*, 085104.
(270) Irudayam, S. J.; Henchman, R. H. *J. Phys. Chem. B* **2009**, *113*, 5871−5884.
(271) Zeifman, A. A.; Stroylov, V. V.; Novikov, F. N.; Stroganov, O. V.; Kulkov, V.; Chilov, G. G. *J. Chem. Theory Comput.* **2013**, *9*, 1093−1102.
(272) Hritz, J.; Läppchen, T.; Oostenbrink, C. *Eur. Biophys. J.* **2010**, *39*, 1573−1580.
(273) Mobley, D. L.; Chodera, J. D.; Dill, K. A. *J. Chem. Theory Comput.* **2007**, *3*, 1231−1235.
(274) Englert, L.; Biela, A.; Zayed, M.; Heine, A.; Hangauer, D.; Klebe, G. *Biochim. Biophys. Acta* **2010**, *1800*, 1192−1202.
(275) Baron, R.; Setny, P.; McCammon, J. A. *J. Am. Chem. Soc.* **2010**, *132*, 12091−12097.
(276) Hummer, G. *Nat. Chem.* **2010**, *2*, 906−907.
(277) Setny, P. *J. Chem. Phys.* **2007**, *127*, 054505.
(278) Setny, P. *J. Chem. Phys.* **2008**, *128*, 125105.
(279) Setny, P.; Baron, R.; McCammon, J. A. *J. Chem. Theory Comput.* **2010**, *6*, 2866−2871.
(280) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 808−813.
(281) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2008**, *130*, 2817−2831.
(282) Genheden, S.; Mikulskis, P.; Hu, L. H.; Kongsted, J.; Söderhjelm, P.; Ryde, U. *J. Am. Chem. Soc.* **2011**, *133*, 13081−13092.
(283) Zhang, H.; Tan, T.; Hetényi, C.; van der Spoel, D. *J. Chem. Theory Comput.* **2013**, *9*, 4542−4551.
(284) Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2009**, *131*, 15403−15411.
(285) Baron, R.; Setny, P.; McCammon, J. A. In *Protein-Ligand Interactions*; Gohlke, H., Ed.; Wiley: New York, 2012; pp 145−170.
(286) Riniker, S.; Barandun, L. J.; Diederich, F.; Krämer, O.; Steffen, A.; van Gunsteren, W. F. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 1293−1309.
(287) Luo, H.; Sharp, K. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10399−10404.
(288) Mihailescu, M.; Gilson, M. K. *Biophys. J.* **2004**, *87*, 23−36.
(289) Rogers, K. E.; Ortiz-Sánchez, J. M.; Baron, R.; Fajer, M.; de Oliveira, C. A. F.; McCammon, J. A. *J. Chem. Theory Comput.* **2013**, *9*, 46−53.
(290) Hünenberger, P. H.; McCammon, J. A. *J. Chem. Phys.* **1999**, *110*, 1856−1872.
(291) Reif, M. M.; Oostenbrink, C. *J. Comput. Chem.* **2014**, *35*, 227−243.
(292) Lau, A. Y.; Roux, B. *Nat. Struct. Biol.* **2011**, *18*, 283−288.
(293) Dixit, S. B.; Chipot, C. *J. Phys. Chem. A* **2001**, *105*, 9795−9799.
(294) Donnini, S.; Mark, A. E.; Juffer, A. H.; Villa, A. *J. Comput. Chem.* **2005**, *26*, 115−122.
(295) Zhou, R.; Das, P.; Royyuru, A. K. *J. Phys. Chem. B* **2008**, *112*, 15813−15820.
(296) Morgan, B. R.; Massi, F. *J. Chem. Theory Comput.* **2010**, *6*, 1884−1893.
(297) Rocklin, G. J.; Mobley, D. L.; Dill, K. A. *J. Chem. Theory Comput.* **2013**, *9*, 3072−3083.
(298) Best, R. B.; Hummer, G. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6732−6737.
(299) Peters, B.; Trout, B. L. *J. Chem. Phys.* **2006**, *125*, 054108.
(300) Rohrdanz, M. A.; Zheng, W.; Clementi, C. *Annu. Rev. Phys. Chem.* **2013**, *64*, 295−316.
(301) Hansen, H. S.; Hünenberger, P. H. *J. Comput. Chem.* **2010**, *31*, 1−23.
(302) Mishra, S. K.; Kara, M.; Zacharias, M.; Koča, J. *Glycobiology* **2014**, *24*, 70−84.
(303) Hansen, H. S.; Hünenberger, P. H. *J. Comput. Chem.* **2011**, *32*, 998−1032.
(304) Kästner, J. *J. Chem. Phys.* **2012**, *136*, 234102.
(305) Fraternali, F.; van Gunsteren, W. F. *Biopolymers* **1994**, *34*, 347−355.
(306) Hansen, H. S.; Hünenberger, P. H. *J. Chem. Theory Comput.* **2010**, *6*, 2622−2646.
(307) Lin, Z.; Liu, H.; Riniker, S.; van Gunsteren, W. F. *J. Chem. Theory Comput.* **2011**, *7*, 3884−3897.
(308) Nagy, G.; Oostenbrink, C. *J. Chem. Inf. Model.* **2014**, *54*, 266−277.
(309) Nagy, G.; Oostenbrink, C. *J. Chem. Inf. Model.* **2014**, *54*, 278−288.
(310) van Gunsteren, W. F.; Bürgi, R.; Peter, C.; Daura, X. *Angew. Chem., Int. Ed.* **2001**, *40*, 351−355.
(311) Christen, M.; Kunz, A. P. E.; van Gunsteren, W. F. *J. Phys. Chem. B* **2006**, *110*, 8488−8498.
(312) Christen, M.; Kunz, A. P. E.; van Gunsteren, W. F. *J. Phys. Chem. B* **2008**, *112*, 11446.
(313) Zuckerman, D. M. *Annu. Rev. Biophys.* **2011**, *40*, 41−62.
(314) Chipot, C.; Lelièvre, T. *SIAM J. Appl. Math.* **2011**, *71*, 1673−1695.
(315) Chen, M.; Cuendet, M. A.; Tuckerman, M. E. *J. Chem. Phys.* **2012**, *137*, 024102.
(316) Zheng, L.; Yang, W. *J. Chem. Theory Comput.* **2012**, *8*, 810−823.
(317) Higo, J.; Ikebe, J.; Kamiya, N.; Nakamura, H. *Biophys. Rev.* **2012**, *4*, 27−44.
(318) Lin, Z.; van Gunsteren, W. F. *J. Chem. Phys.* **2013**, *139*, 144105.
(319) Roux, B. *Comput. Phys. Commun.* **1995**, *91*, 275−282.
(320) Shen, T.; Hamelberg, D. *J. Chem. Phys.* **2008**, *129*, 034103.
(321) Boned, R.; van Gunsteren, W. F.; Daura, X. *Chem.—Eur. J.* **2008**, *14*, 5039−5046.
(322) Lin, Z.; Necula, C.; van Gunsteren, W. F. *Chem. Phys.* **2014**, *428*, 156−163.
(323) Shirts, M. R. In *Computational Drug Discovery and Design*; Baron, R., Ed.; Springer: New York, 2012; pp 425−467.
(324) Brown, S. P.; Muchmore, S. W.; Hajduk, P. J. *Drug Discovery Today* **2009**, *14*, 420−427.
(325) Reynolds, C. H.; Holloway, M. K. *ACS Med. Chem. Lett.* **2011**, *2*, 433−437.

(326) Paliwal, H.; Shirts, M. R. *J. Chem. Theory Comput.* **2011**, *7*, 4115−4134.
(327) The Binding Database. http://www.bindingdb.org/bind/index.jsp (accessed Feb. 12, 2014).
(328) Moult, J.; Fidelis, K.; Kryshtafovych, A.; Tramontano, A. *Proteins: Struc., Funct., Bioinf.* **2011**, *79*, 1−10.
(329) Riniker, S.; Christ, C. D.; Hansen, N.; Mark, A. E.; Nair, P. C.; van Gunsteren, W. F. *J. Chem. Phys.* **2011**, *135*, 024105.
(330) Filippini, G.; Bonal, C.; Malfreyt, P. *Mol. Phys.* **2012**, *110*, 1081−1095.
(331) de Ruiter, A.; Boresch, S.; Oostenbrink, C. *J. Comput. Chem.* **2013**, *34*, 1024−1034.
(332) Gumbart, J. C.; Roux, B.; Chipot, C. *J. Chem. Theory Comput.* **2013**, *9*, 794−802.
(333) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. *Comput. Phys. Commun.* **2014**, *185*, 604−613.
(334) van Gunsteren, W. F.; Dolenc, J.; Mark, A. E. *Curr. Opin. Struct. Biol.* **2008**, *18*, 149−153.
(335) van Gunsteren, W. F.; Mark, A. E. *J. Chem. Phys.* **1998**, *108*, 6109−6116.
(336) Davis, A. M.; St-Gallay, S. A.; Kleywegt, G. J. *Drug Discovery Today* **2008**, *13*, 831−841.
(337) Nicholls, A.; McGaughey, G. B.; Sheridan, R. P.; Good, A. C.; Warren, G.; Mathieu, M.; Muchmore, S. W.; Brown, S. P.; Grant, J. A.; Haigh, J. A.; Nevins, N.; Jain, A. N.; Kelley, B. *J. Med. Chem.* **2010**, *53*, 3862−3886.
(338) van Gunsteren, W. F. In *Modelling of Molecular Structures and Properties*; Rivail, J. L., Ed.; Elsevier: Amsterdam, The Netherlands, 1990; pp 463−478.