HHS Public Access

Author manuscript

*J Chem Inf Model*. Author manuscript; available in PMC 2021 November 23.

Published in final edited form as:

*J Chem Inf Model*. 2020 November 23; 60(11): 5595–5623. doi:10.1021/acs.jcim.0c00613.

# Alchemical Binding Free Energy Calculations in AMBER20: Advances and Best Practices for Drug Discovery

**Tai-Sung Lee**[1], **Bryce K. Allen**[2], **Timothy J. Giese**[1], **Zhenyu Guo**[2], **Pengfei Li**[2], **Charles Lin**[2], **T. Dwight McGee Jr.**[2], **David A. Pearlman**[3], **Brian K. Radak**[2], **Yujun Tao**[1], **Hsu-Chun Tsai**[1], **Huafeng Xu**[2], **Woody Sherman**[2,*], **Darrin M. York**[1,**]

[1]Rutgers, the State University of New Jersey, Laboratory for Biomolecular Simulation Research, and Department of Chemistry and Chemical Biology, United States

[2]Silicon Therapeutics, Boston, Massachusetts 02210, United States

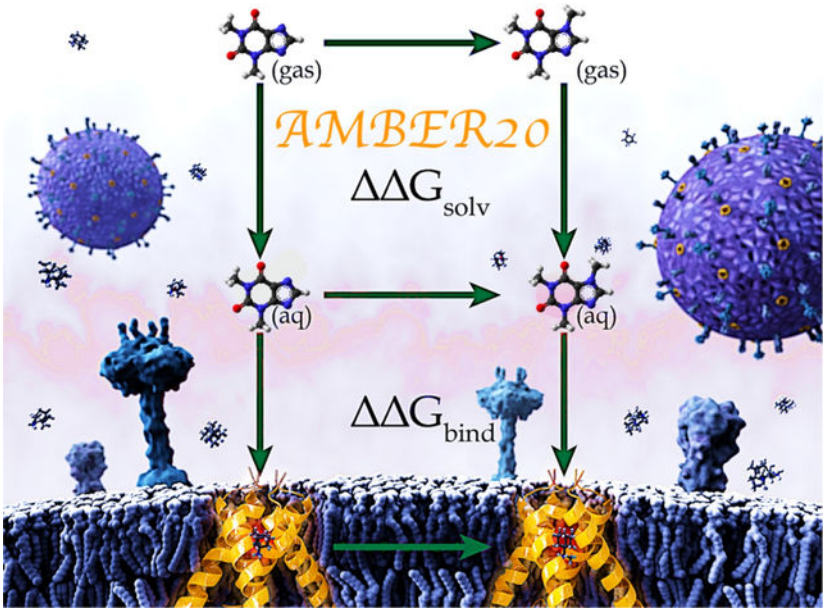
[3]QSimulate Incorporated, Cambridge, Massachusetts 02139, United States

## Abstract

Predicting protein-ligand binding affinities and the associated thermodynamics of biomolecular recognition is a primary objective of structure-based drug design. Alchemical free energy simulations offer a highly accurate and computationally efficient route to achieving this goal. While the AMBER molecular dynamics package has successfully been used for alchemical free energy simulations in academic research groups for decades, widespread impact in industrial drug discovery settings has been minimal due to previous limitations within the AMBER alchemical code, coupled with challenges in system setup and post-processing workflows. Through a close academia-industry collaboration we have addressed many of the previous limitations with an aim to improve accuracy, efficiency and robustness of alchemical binding free energy simulations in industrial drug discovery applications. Here, we highlight some of the recent advances in AMBER20 with a focus on alchemical binding free energy (BFE) calculations, which are less computationally intensive than alternative binding free energy methods where full binding/unbinding paths are explored. In addition to scientific and technical advances in AMBER20, we also describe the essential practical aspects associated with running relative alchemical BFE calculations along with recommendations for best practices, highlighting the importance not only of the alchemical simulation code, but also the auxiliary functionalities and expertise required to obtain accurate and reliable results. This work is intended to provide a contemporary overview of the scientific, technical, and practical issues associated with running relative BFE simulations in AMBER20, with a focus on real-world drug discovery applications.

## Graphical Abstract

*Corresponding author: Woody@silicontx.com. **Corresponding author: Darrin.York@rutgers.edu.




## 1 Introduction

Accurate, robust prediction of the relative binding free energy (BFE) of ligands to a target protein is of tremendous value in drug discovery, serving as an *in silico* assay and a way to gain deeper insights into the origin of biomolecular recognition.[1–4] Rigorous free energy simulations of ligand-protein binding yield both thermodynamic and kinetic information, but can be extremely computationally intensive to converge to high precision due to the need to explore and sufficiently sample the minimum free energy pathway that connect bound and unbound states (including often starkly different entropic differences). Alchemical BFE simulations, on the other hand, can be engineered to be much more tractable owing to the property that the free energy is a state function from which thermodynamic end states (bound and unbound) can be connected by any pathway. In practice, thermodynamic cycles can be constructed that utilize ”alchemical” pathways between end states that can be optimally computed. Whereas alchemical BFE simulations do not provide a complete mechanistic and kinetic characterization of the binding process, they provide a highly efficient and practical approach to predict the binding affinities of lead compounds important in drug discovery.

While alchemical free energy simulation capability has been in AMBER since the 1980s, a number of technical and scientific challenges have impeded progress toward the broader adoption and higher impact of AMBER in drug discovery projects. This work provides a modern update of advances in BFE simulations in AMBER20 and a description of current guidelines and best practices in the context of real-world drug discovery applications. The manuscript is organized as follows. The remainder of this section describes the history and origin of free energy simulations in AMBER leading up to the latest developments in AMBER20 that enable large-scale application in drug discovery projects. Section 2 that follows provides an overview of the background formalism for alchemical free energy

simulations, with extended discussion of transformation pathways and protocols using so-called ”softcore potentials”. Sections 3 and 4 discuss performance and feature advances, respectively, in AMBER20 for drug discovery. Section 5 reviews practical considerations and provides guidance to achieving robust and reliable BFE results, including system preparation, docking, atom mapping and $\lambda$ scheduling, use of restraints and estimation of confidence and errors. Section 6 presents our perspective about important forthcoming and future work to advance the state-of-the-art. The final section provides brief concluding remarks that emphasize the purpose and main points of the manuscript.

### 1.1 Historical Overview of Alchemical Binding Free Energy Simulations

Alchemical binding free energy simulations on computers have been performed since the 1980s, although the theoretical foundation began decades earlier with studies of nonpolar gases by Zwanzig, where he derived the master equation for free energy perturbation (FEP) to compute thermodynamic differences between two states A and B.[5] In complementary (and earlier) work, Kirkwood described a coupling parameter, typically called lambda ($\lambda$), that has since been used to improve the accuracy of FEP calculations for meaningful chemical transformations by making neighboring states much closer (and therefore having a smoother thermodynamic path) as one moves from state A to B.[6] Later, Bennett introduced an alternative approach based on minimizing the expected squared error (known as Bennett Acceptance Ratio, BAR),[7] which was further improved based on a statistically optimal analysis of samples (multistate BAR or MBAR).[8–11] Thermodynamic integration (TI), an alternative approach to FEP/BAR based on Kirkwood’s work on the theory of liquids, requires the calculation of the Boltzmann averaged potential energy derivative at each intermediate state $\lambda$.[12] AMBER20 now has functionality for these multiple approaches of FEP (BAR, MBAR, and TI), which can be used in tandem with minimal computational overhead in order to gain confidence in free energy estimations.

The first published free energy methods applied to chemical systems came from Postma, Berendsen, and Haak in 1982, where the authors reported the free energy cost associated with the formation of a cavity in water[13] followed by Jorgensen’s seminal work in 1985 computing the hydration free energy difference of ethane to methanol, demarking the first true alchemical transformation.[14] Alchemical free energy simulations were made more efficient by Tembe and McCammon, who noted the concept of the thermodynamic cycle and designed a model system to compute the $\Delta\Delta G$ between atoms.[15] This approach was applied to compute the free energies for model systems, such as ligand binding in a host-guest system[16,17] and hydration of noble gases.[17] While these early works showed the promise of alchemical free energy simulations in drug discovery, it took years for the first prospective applications to appear in the literature[18] and over a decade for the first published industry application to appear,[19] yet these studies involved only single heavy atom changes. Eventually, larger and more pharmaceutically relevant chemical transformations were shown to be tractable with alchemical free energy simulations.[20] Details about the history, theory, methods, and applications of alchemical simulations can be found in a number of excellent reviews.[3,4,21–27]

### 1.2 The Origin of Free Energy Simulations in AMBER

The first implementation of free energy calculations within the AMBER suite came in 1986. Singh implemented and tested the software, which he built upon the previously developed AMBER molecular dynamics (MD) code base that had been published a year earlier as AMBER 2.0.[28] The newly christened "Gibbs" module of AMBER was subsequently applied to several systems in a collaboration between Singh and Kollman with Bash of the Langridge Laboratory at UCSF. Together, they published the first papers describing the application of AMBER for free energy calculations, and the first computational free energy paper to appear in Science Magazine.[29–31] While this was not the first application of such free energy calculations to be published (see section 1.1), in many ways it exploded interest in the field, owing to the broad journal readership, the pharmaceutically relevant test systems (nucleic acid bases, amino acid side chains, organic small molecules, and protein-ligand interactions), and the asserted high quality of the results.

Upon release, this first implementation of free energy calculations within AMBER supported three free energy protocols: Free Energy Perturbation (FEP), Thermodynamic Integration (TI), and "Slow Growth," which represented the limiting case of TI where (it was asserted) if you used a very large number of $\lambda$ windows, you could evaluate each window with exactly one sample point. It was later demonstrated that the slow growth approximation was unreliable in practice due to a "Hamiltonian lag"[32] and therefore this approach was not pursued for long (although there has been a recent resurgence due to foundational work by Jarzynski,[33] as demonstrated by Gapsys *et al.*[34] and others). Free energy-specific options in this first release were limited to setting the number of $\lambda$ windows, the durations of equilibration, the amount of data collection at each window, and the ability to "decouple" the vdW and electrostatic contributions. At this time, calculations were limited to the single topology approach (wherein only a single geometry for the molecule exists at any time, and changes with lambda are reflected by modifications to the target values of the internal coordinates and modifications of atom types)[35] and the integration required for TI was performed using the trapezoidal rule.

At this early stage in the development and application of free energy methods, Gibbs in AMBER was one of only a few software packages broadly available. The primary molecular simulation packages at this time were research software packages from the academic groups of Kollman (AMBER),[36] Jorgensen (MCPRO),[14] Karplus (CHARMM),[37] and van Gunsteren (GROMOS).[38] Shortly after their initial publications, both Singh and Bash left UCSF for other positions. A second generation of development of the Gibbs free energy module was carried out primarily by Pearlman and Kollman, where they focused on 1) addressing shortfalls in the first implementation (e.g. the contribution from bond constraints),[39,40] 2) dynamically changing the $\lambda$ schedule to reflect the evolution of the system as it progressed,[41] 3) validating/improving the intermediate mixing rules for non-physical $\lambda$ states,[42] 4) developing, characterizing and implementing best practices,[35] and 5) integrating error propagation.[43] During this period, the Gibbs module was also updated to reflect improvements to the base molecular dynamics methods, including the development of the particle mesh Ewald (PME) method[44,45] for efficient treatment of long-range electrostatic interactions in simulations of proteins[46] and nucleic acids,[47,48] and their

parallel implementation to accommodate the supercomputing platforms of the era. In the early 2000s, reflecting a desire to lower the maintenance overhead in light of a rapidly increasing number of modifications to the base AMBER molecular dynamics platform, the fundamental Gibbs functionality was reimplemented into the Sander module of AMBER, which also serves as the (non-free energy) molecular dynamics platform, and the Gibbs module was retired. More recently, the free energy methods have been implemented in the AMBER PMEMD program, which generally replicates the functionality of Sander but provides appreciably better efficiency for highly parallel CPU platforms.

Free energy-specific improvements since integration into Sander (and then PMEMD) have included methods that can improve the efficiency of sampling (e.g. Replica exchange), more control over the $\lambda$ scheduling, improved methods for integrating the TI curve, and tools to support absolute binding free energy calculations. Broadly, changes to free energy methodologies in AMBER since the first implementations have been evolutionary. The fundamental advancement that has had the greatest effect on the ability to obtain reliable results from free energy calculations has been an increase by more than 6 orders of magnitude in compute power since the first AMBER free energy simulations, coupled with better force fields and auxiliary tools for facilitating control over advanced simulation options.

A major performance enhancement introduced in AMBER11 was the ability to use graphical processing units (GPUs) to massively accelerate PMEMD for both explicit solvent PME and implicit solvent/Generalized Born (GB) simulations.[49,50] The performance envelope was pushed even further with AMBER14 and AMBER16. Those releases represented leaps in both performance and functionality through the full utilization of the single-precision floating-point format (SPFP), which significantly boosted performance on GPUs without sacrificing numerical accuracy.[51] Although, the GPU accelerated version of PMEMD, namely PMEMD.cuda, has been designed to support as many of the standard PMEMD features as possible, there were some limitations, such as the inability to perform alchemical free energy simulations on GPUs. Giese and York[52] recognized that certain types of alchemical transformations that involved only the interpolation of force field parameters representing the two end states (rather than mixing of their Hamiltonians) could be achieved without modification of the PMEMD.cuda engine. By bringing the Gibbs functionality in Sander out of retirement and making minor extensions to work with PME, some alchemical transformations could be achieved with a post-processing tool. Around the same time, the GPU-accelerated alchemical free energy module was first implemented as a patch of AMBER16[53] and later incorporated into the official AMBER18 release.[54] Since then, the free energy methods in AMBER have been carefully validated[55] and applied,[56,57] and many advances for alchemical free energy calculations have been actively developed, such as a novel soft-core potential,[58] various types of restraints,[59] and robust analysis methods.[52,60]

### 1.3 AMBER20 for Drug Discovery Applications

Despite the advances and applications described above, the impact of alchemical binding free energy simulations has been limited in drug discovery for a number of reasons, most commonly noted as inaccurate force fields, insufficient sampling, and ease of use. However,

with the many force field advances in the past decades (discussed below in sections 5.1, 5.9, and 6.5) and increased computational throughput via GPUs (discussed in section 3, many of the remaining issues involve the balance between more advanced controls for optimal performance and simplified interfaces to improve usability. Much work has been done to expose BFE calculations to a broader audience through graphical user interfaces,[61,62] workflow tools,[63,64] and integration with powerful molecular operating environments.[65–68] While ease of use has expanded the user base for BFE methods, it has also reduced the degree of expertise that can be inserted by the user to optimize performance and reliability. Indeed, performing BFE simulations is still an expert process where experience plays an important role, especially for challenging targets where the timescales of important degrees of freedom and conformational flexibility might be unknown. As such, in this work we also highlight the practical considerations for reliable predictions in real-world applications. In some cases robust automated programs are available, but as will be described, there are many subtle details that require close attention by expert users to obtain optimal results.

Through an academia-industry collaboration, we have addressed some of the primary issues that have in the past limited AMBER utilization in drug discovery efforts. Most recently, we have improved the softcore potential to ensure more reliable simulations across a broad set of diverse alchemical transformations. We have also implemented restraints for absolute binding free energy (ABFE) simulations and finer control of the bonded terms in relative binding free energy (RBFE) simulations. These advances, coupled with the high performance of AMBER on GPUs and the practical considerations outlined herein, should facilitate the broader adoption of alchemical free energy simulations in drug discovery. We hope that these advances, coupled with the great work by others in the field, will aid researchers in drug discovery to more efficiently design medicines to treat diseases with unmet medical needs.

## 2 Background Formalism for Alchemical Free Energy Simulations

The change in free energy between two thermodynamic states can be computed from equilibrium simulations using a free energy perturbation (FEP)[5] (sometimes referred to as “thermodynamic perturbation”) or thermodynamic integration (TI)[6,69] formulations, or through non-equilibrium ensemble simulations using the Jarzynski equality and its equation variations.[33,70–74] For the purposes of the current work, we will focus on the calculation of relative binding free energies from equilibrium simulations using TI and FEP formulations with Bennett Acceptance Ratio[7,75] (BAR) and its multistate generalization (MBAR).[9,76] For additional discussion of factors that influence accuracy and robustness of free energy simulations, we refer the reader to several excellent examples.[2,3,71,77–85]

Consider the transformation of a system of $N$ particles in an initial state “0” characterized by potential energy function $U_0(\mathbf{r}^N)$, where $\mathbf{r}^N = \mathbf{r}_1,\mathbf{r}_2 \cdots \mathbf{r}_N$ represents the Cartesian positions of each particle, to a final state “1” characterized by potential energy function $U_1(\mathbf{r}^N)$ having the same degrees of freedom. The potential energy functions can, for example, represent different molecular species or environments. In general we will refer to this type of transformation as an “alchemical transformation” from which differences in thermodynamic end states can be determined, to distinguish it from a physical or chemical transformation

that involves a real mechanistic pathway that contains both thermodynamic and kinetic information. The change in free energy of the alchemical transformation between states 0 and 1 can be computed from the ratio of configurational integrals $Z_0$ and $Z_1$ as

$$\Delta A_{0\to 1} = A_1 - A_0 = -\beta^{-1}\ln\left(Z_1/Z_0\right) \tag{1}$$

where $\beta^{-1} = k_B T$, $k_B$ is the Boltzmann constant and $T$ the absolute temperature, and

$$Z_s = \int_V e^{-\beta U_s\left(\mathbf{r}^N\right)} d\mathbf{r}^N \tag{2}$$

where $s$ is the state of the system (0 or 1) and $V$ is the volume of the configurational space. Here we use the Helmholtz free energy, $\Delta A$, in the $N,V,T$ ensemble to motivate discussion, whereas extension to the Gibbs free energy and the $N,P,T$ ensemble is straight forward.

In the FEP formulation, substitution of Eqn. 2 into Eqn. 1 leads to the so-called Zwanzig, or "exponential average" relationship:

$$\Delta A_{0\to 1} = -\beta^{-1}\ln\left\langle e^{-\beta\Delta U}\right\rangle_0 \tag{3}$$

where $\Delta U = U_1 - U_0$ and the average $\langle\cdots\rangle_0$ involves integration over the configurational space of the Boltzmann probability for state 0, $P_0\left(\mathbf{r}^N\right) = e^{-\beta U_0\left(\mathbf{r}^N\right)}/Z_0$, or equivalently, Boltzmann sampling from this ensemble from a molecular simulation using the forces derived from the potential energy $U_0(\mathbf{r}^N)$. This expression allows the free energy of the transformation to be computed while requiring sampling only at one thermodynamic end state. The above relation has many useful variants that consider the other or both end states:

$$\begin{aligned}\Delta A_{0\to 1} &= -\beta^{-1}\ln\left\langle e^{-\beta\Delta U}\right\rangle_0 \\ &= -\beta^{-1}\ln\left\langle e^{\beta\Delta U}\right\rangle_1^{-1} \\ &= -\beta^{-1}\ln\left[\frac{\langle f(\beta(\Delta U - C))\rangle_0}{\langle f(\beta(C-\Delta U))\rangle_1}\right] + C\end{aligned} \tag{4}$$

where $f(x) = 1/[1 + \exp(x)]$ is the Fermi function, and $C$ is a constant with units of energy. If $C$ is set to zero, one can recover the original formula of Eqn 3, or if one solves for $C$ such that the numerator and denominator of the logarithmic term are equal (making this term vanish), then one obtains an optimal statistical estimate using the BAR method.[7,75] One can further generalize this expression to consider non-Boltzmann sampling.[86]

In principle, the above FEP equations only require sampling at the thermodynamic end states. However, the statistical precision requires that there is sufficient phase space overlap,[60,87,88] which typically necessitates stratifying the transformation into smaller steps along a pathway. In theory, the free energy is a state function, and thus the free energy difference between states is independent of the path that connects them. In practice, the choice of this pathway is of immense importance, as it can be extremely challenging to converge sampling along the pathway itself. Let us then define a thermodynamic parameter $\lambda$ that smoothly connects states 0 and 1 through a $\lambda$-dependent potential $U(\mathbf{r}^N;\lambda)$ such that $U(\mathbf{r}^N;0) = U_0(\mathbf{r}^N)$

and $U(\mathbf{r}^N;1) = U_1(\mathbf{r}^N)$. Within the FEP formulation, the transformation can be broken down into a series of $M$ steps corresponding to a set of $\lambda$ values $\lambda_1,\lambda_2 \cdots \lambda_M$ ranging from 0 to 1 such that there is sufficient phase space overlap between neighboring intermediate $\lambda$ states. This requires a separate simulation for each "$\lambda$ window" that corresponds to a specific value of $\lambda$ and using forces derived from the potential energy $U(\mathbf{r}^N;\lambda)$. These simulations can then be analyzed using the BAR[7,75] or MBAR[9,76] methods.

Alternatively, with the introduction of a defined pathway between states, the change in free energy can be equated to the reversible work of conducting the transformation between states, and this gives rise to the TI formulation,[6,69] which is characterized by the formula and numerical quadrature estimate

$$\begin{aligned} \Delta A_{0\to 1} &= \int_0^1 d\lambda \left\langle \frac{\partial U(\mathbf{r}^N;\lambda)}{\partial \lambda} \right\rangle_\lambda \\ &\approx \sum_{k=1}^{M} w_k \left\langle \frac{\partial U(\mathbf{r}^N;\lambda)}{\partial \lambda} \right\rangle_{\lambda_k} \end{aligned} \tag{5}$$

where the second sum indicates numerical integration over $M$ quadrature points ($\lambda_k$, for $k = 1, \cdots M$) with associated weights $w_k$.

## 2.1 Alchemical transformation pathways and softcore potentials

The simplest way in which the $U(\mathbf{r}^N;\lambda)$ can be constructed is to use a linear interpolation between states, which we will designate as $U^L(\mathbf{r}^N;\lambda)$:

$$U^L(\mathbf{r}^N;\lambda) = U_0(\mathbf{r}^N) + \lambda \Delta U(\mathbf{r}^N) \tag{6}$$

where the endpoint difference $\Delta U(\mathbf{r}^N) \equiv U_1(\mathbf{r}^N) - U_0(\mathbf{r}^N)$ is also, by coincidence, the thermodynamic derivative with respect to $\lambda$. Hence, the common energy components that are identical between $U_1(\mathbf{r}^N)$ and $U_0(\mathbf{r}^N)$ need not be explicitly considered as the corresponding difference is zero. As has been well established, the linear alchemical transformation pathway leads to practical problems that can be partially overcome by the use of so-called "softcore" potentials for non-bonded Lennard-Jones (LJ) and Coulombic electrostatic (Coul) interactions.[89,90] A commonly used softcore potential transformation pathway[90] originally implemented in AMBER is

$$U_0^{SC}(\mathbf{r}^N;\lambda) + \lambda \Delta U^{SC}(\mathbf{r}^N;\lambda) \tag{7}$$

where $\Delta U^{SC}(\mathbf{r}^N;\lambda) \equiv U_1^{SC}(\mathbf{r}^N;1-\lambda) - U_0^{SC}(\mathbf{r}^N;\lambda)$ as before.

There have been many different proposed softcore potential forms that modify, or "soften", these interactions. In the following sections, to be more clear, we only show the softcore potential corresponding to one end state and the system total potential should be written as the properly weighted sum of the two end states. The LJ and Coul interactions for a set of interacting point particles $i$ and $j$ separated by a distance $r_{ij}$ are given by

$$U_{\mathrm{LJ}}(r_{ij}) = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{8}$$

and

$$U_{\mathrm{Coul}}(r_{ij}) = \left(\frac{q_i q_j}{4\pi\epsilon_0}\right)\frac{1}{r_{ij}} \tag{9}$$

where $\sigma_{ij}$ and $\epsilon_{ij}$ are the pairwise LJ contact distance and well depth, respectively, and $q_i$ and $q_j$ are the partial charges of particles $i$ and $j$. In order to "soften" these pairwise interactions with particles contained within the selected softcore region, one can modify the effective interaction distance by introducing a parametric form for separation-shifted scaling with an adjustable parameter. A commonly used form of these modifications is[89,90]

$$r_{ij}^{\mathrm{LJ}}(\lambda;\alpha) = \left[r_{ij}^{n} + \lambda\alpha\sigma_{ij}^{n}\right]^{1/n} \tag{10}$$

and

$$r_{ij}^{\mathrm{Coul}}(\lambda;\beta) = \left[r_{ij}^{m} + \lambda\beta\right]^{1/m} \tag{11}$$

where $n$ and $m$ are positive integers and $\alpha$ and $\beta$ are adjustable positive semidefinite parameters for the LJ and Coul softcore interactions, respectively, with values of zero corresponding to no softcore modification for any $\lambda$ value. In several molecular simulation software suites, including the default in AMBER, the values of $n = 6$ and $m = 2$ are used, although other values have also been suggested,[90] and as will be discussed below, combined with new smoothstep softcore potentials, considerable improvements can be made to stabilize sampling of the transformations.

The LJ and Coul softcore potentials are defined from these scaled effective interaction distances as

$$U_{\mathrm{LJ}}^{SC}(r_{ij};\lambda) = U_{\mathrm{LJ}}\left[r_{ij}^{\mathrm{LJ}}(\lambda;\alpha)\right] \tag{12}$$

and

$$U_{\mathrm{Coul}}^{SC}(r_{ij};\lambda) = U_{\mathrm{Coul}}\left[r_{ij}^{\mathrm{Coul}}(\lambda;\beta)\right] \tag{13}$$

The thermodynamic derivatives with respect to $\lambda$ can be obtained using the chain relation. Recently, we developed a new smoothstep softcore potential for nonbonded LJ and Coul interactions, implemented in AMBER20 and demonstrated below, that further improves the stability in practical calculations.[58] We introduce a non-linear $\lambda$ scaling function by replacing $\lambda$ in Eqn. 7 with a so-called second-order smoothstep function, $S_2(\lambda)$, defined as

$$S_2(\lambda) = 6\lambda^5 - 15\lambda^4 + 10\lambda^3 \tag{14}$$

Note that $S_2(\lambda)$ varies smoothly from 0 to 1 and has vanishing derivatives at $\lambda$=0 and 1. Details of the implementation and testing of the second-order smoothstep softcore potential in AMBER20 can be found elsewhere.[58] Similar in spirit, but slightly different in details, is closely related work first introduced by Hritz and Oostenbrink[91] and described in further detail by Riniker and co-workers[92] where the use of third-order polynomials enable different $\lambda$-dependency (referred to a "individual Lambdas") for calculation of relative free energies. This form of the softcore potential also has been shown to have impact on the ability to predict $\lambda$ derivatives at non-simulated points in extended TI methods.[93]

Further, as will be demonstrated below, a promising new form of the effective interaction distance with separation-shifted scaling is given as

$$r_{ij}^{X}(\lambda;\alpha^{X}) = \left[r_{ij}^{n} + \alpha^{X} W(r_{ij}) S_2(\lambda) \sigma_{ij}^{n}\right]^{1/n} \tag{15}$$

where $X$ generically represents either LJ or Coul, $\alpha^X$ is the corresponding unitless parameter, and the weight function of the softcore potential $W(r_{ij})$ is designed to smoothly return to the normal $r_{ij}$ value by the end of the cutoff:

$$W(r_{ij}) \equiv 1 - S_2\left(\frac{r_{ij} - R_{cut,i}}{R_{cut,f} - R_{cut,i}}\right) \tag{16}$$

where $R_{cut,i}$ is the onset distance where the weight function becomes effective and $R_{cut,f}$ is the final distance of the weight function where the softcore potential completely diminishes, and is set to the same as the non-bonded cutoff distance.

## 2.2 Common problems with softcore potentials and their solutions

We call specific attention to three problems that commonly occur in simulations of alchemical transformations, and in particular for "concerted transformations" that involve simultaneous changes in both non-bonded LJ and Coul terms. These are referred to as the "endpoint catastrophe", the "particle collapse problem", and the "large gradient-jump problem".

The endpoint catastrophe is well-known, and arises from a sharp divergence of the contribution to the free energy at the thermodynamic endpoints ($\lambda$ values near 0 and 1) due to poor phase space overlap, and can be avoided by the use of softcore potentials. The particle collapse problem involves the introduction of new spurious minima at intermediate $\lambda$ states, frequently manifesting in the artificial superposition of particles that can lead to large amplitude fluctuations or phase transition behavior along the $\lambda$ dimension.[94] This problem results from an imbalance of Coulomb attraction and exchange repulsion, and can be overcome by ensuring that these terms are scaled in such a way that preserves overall repulsive behavior at short distances for all $\lambda$ values (e.g., by insuring that the repulsive terms are sufficiently large to overcompensate for any attractive Coulomb interactions). Finally, the large gradient-jump problem involves sensitivity of the free energy to certain softcore parameter values that adjust the exchange repulsion and can lead to spurious jumps in the free energy near the thermodynamic endpoints. This problem can be solved through

use of a smoothstep softcore potential that has scaling weights with derivatives that vanish at the endpoints.

### 2.3 Stepwise versus concerted, and absolute versus relative protocols for alchemical transformations

Here we discuss strategies for alchemical free energy simulation protocols and parameters that will yield the best results for a given system of interest. One of the most pivotal technical issues is the choice of the alchemical path connecting the two real states (i.e., connecting the two thermodynamic endpoints). While, the free energy difference between two states is independent of the path that connects them in the regime of complete conformational sampling, in practical calculations of complex systems, the choice of the alchemical transformation path is critical to obtain stable, converged results with affordable sampling.

In the discussion that follows, we separate the atoms involved in the alchemical transformation into two regions: the softcore region, and the common core region. The common core atoms are those that transform from a “real atom” in the initial state to another real atom in the final state, and in intermediate $\lambda$ states, interact with other (non-softcore) atoms via normal Lennard-Jones (LJ) and Coulombic electrostatic (Coul) interactions. The softcore atoms, on the other hand, are those selected to interact with other atoms (including the common core atoms) via a softcore potential[90] in intermediate $\lambda$ states. Often the atoms of the softcore region are transformed from “real atoms” in the initial state to “dummy atoms” in the final state. A discussion of the requirements that the dummy atoms reproduce the ensemble and potential of mean force of the real state has been discussed extensively by Boresch and Karplus[95,96] and Roux and co-workers.[84,97]

The two most commonly applied procedures for alchemical transformations are referred to as “stepwise” and “concerted” protocols.[83,98] For the stepwise protocol, also referred to as ”split”, “multi-step” or “decoupled” procedures,[56] the transformation is carried out by scaling Coulombic and Lennard-Jones interactions separately, where the charges of the dummy atoms are scaled linearly and LJ interactions are scaled via the softcore potentials. In these procedures, all or parts of the Coulomb and LJ transformations are decoupled, and performed as separate steps. An example of a 3-step “decharge-LJ-recharge” protocol would be as follows: First, the atoms in the softcore region (those atoms that will transform into dummy atoms) are fully decharged. Next, these decharged atoms undergo a LJ transformation using softcore potentials, while at the same time the charges of the non-softcore atoms are also transformed. Last, the atoms in the softcore region are recharged to the final state. This protocol is generally quite robust, since the softcore LJ transformations occur after the partial charges of the softcore atoms have been eliminated. One caveat of the conventional stepwise procedure is that, depending on the selection of atoms in the softcore region, a non-integer charge change can be introduced at intermediate $\lambda$ states even for alchemical transformations between molecules having the same net charge. Care should be taken to include net charge corrections and appropriate sampling in these cases (see below).

Alternatively, one can use a concerted protocol (also referred to as “unified,” “single step” transformation). In this procedure, the softcore LJ and Coulomb terms are in some way

performed in concert. This procedure might have some advantages in terms of throughput performance and ease of use with advanced $\lambda$-schedule optimization and enhanced sampling schemes, such as $\lambda$ dynamics,[99–102] Hamiltonian replica exchange methods,[103–107] adaptive biasing[100,108,109] or self-adjusted mixture sampling[110,111] methods. Concerted alchemical transformations, however, are more sensitive to the treatment of softcore atoms, and are more susceptible to the endpoint, particle collapse and large-gradient jump problems discussed earlier. Consequently, it is of tremendous practical interest to work toward more robust and efficient methods to enable stable concerted alchemical transformations.

Related to these issues is the choice of the atoms in the softcore region. There are a number of strategies, methods, and software tools that have been developed to assist in defining the optimal sets of transformations for a library of compounds. This is sometimes referred to as a "perturbation map".[3] One method of common core/softcore atom selection is based on maximizing the common substructure (i.e., minimizing the number of softcore atoms that are to be transformed).[112] Alternatively, selection can be based on grouping softcore atoms into chemical functional groups.[84,113,114] Additional detail can be found in 5.5.

Alchemical transformations are most reliable when the transformations involve a short thermodynamic path (i.e. minimal perturbation to the free energy landscape). This often translates into perturbing the smallest number of atoms, although the nature of the perturbation (size, polarity, conformational preferences, etc.) can have a significant impact on reliability. Consequently, most drug discovery applications focus on computation of relative binding free energies (RBFEs), where a common core is unperturbed. Still, even with a small number of perturbed atoms, the thermodynamic path between the states may be long due to the nature of the perturbation (e.g. a small ligand change that induces a large protein conformational change or alters the ligand conformational preference). While there are no procedures to our knowledge that can *a priori* determine when a given perturbation is too large, there are many reviews of free energy methods that provide guidance and best practices for dealing with such situations.[20,22,25,115] In some cases where large perturbations lead to an exceedingly long thermodynamic path, it is recommended to insert intermediate molecules that bridge the two molecules of interest, as described in a recent application to BACE.[116] Additionally, it is possible to compute absolute binding free energies (ABFEs), whereby an entire ligand is transformed to a dummy-state that is non-interacting with its environment. While this process typically involves a much larger "perturbation", and consequently more sampling to achieve a fixed level of precision, in some cases it may be complementary or even preferable to the calculation of RBFEs alone. ABFE is particularly useful when exploring diverse ligands, such as in virtual screening, as described recently by Cournia *et al.*[4]

## 3 Performance

A critical aspect of BFE simulations is the amount of conformational sampling, which directly relates to the convergence and accuracy of the simulations. While longer simulations can be achieved with more wallclock time, there quickly comes a point where impact in drug discovery will be limited due to real-time throughput of guiding predictions in time-critical projects. Historically, compute power has been dominated by the speed of individual cores.

As single core performance stagnated in the past decade, parallel computing emerged to allow performance scaling to remain. Molecular dynamics (MD) is a problem that is inherently parallelizable, although challenging, as each atom has to compute its energy and forces relative to the current state of the system (i.e. all other atoms). This makes MD an ideal candidate for graphical processing units (GPUs),[117] and since 2012,[49,50] MD has been largely performed on GPUs. Current single GPU performance offers orders of magnitude increased performance relative to a conventional central processing unit (CPU) hardware for most common protein systems. While more sampling is generally preferred in free energy simulations, it is not always the case that more sampling affords better results. The disconnect between sampling and accuracy can broadly be attributed to: 1) poor force field (sampling cannot help), 2) local minima, where sampling in the local minimum of interest is sufficient to attain converged free energy results but additional sampling opens new regions of phase space, thereby resulting in poorer apparent convergence, and 3) poor system setup, where a longer simulation may result in propagation of errors that increasingly degrade the results over time, such as protein unfolding events.

With the AMBER18 release, a GPU-accelerated Thermodynamic Integration (GTI) method was implemented.[53,54,118] The key technical challenge overcome in AMBER18 GTI involved cleverly enabling TI-based calculations without compromising the optimized AMBER energy kernels. This was accomplished by using a streaming kernel to filter and separately process alchemical atoms and their interactions. Thus, the GTI code has a slight performance dip in comparison to standard MD simulations in AMBER but still a tremendous speedup relative to CPU implementations and other GPU codes. For example, a TI calculation on cyclin-dependent kinase 2 (CDK2) with approximately 54,000 atoms takes approximately 4.5 GPU-hours using a GTX 1080Ti GPU with 24 $\lambda$ windows split between a complex stage and a solvated stage (contains 4,500 atoms) using 2 ns simulation time per $\lambda$, with a 4 fs time step facilitated with hydrogen mass repartitioning (HMR).[119] These simulations can also be run at a 2 fs time step without HMR with shake, and 1 fs time steps without shake. This same simulation takes 2.5 GPU-hours on the more recent RTX 2080Ti GPUs. As each $\lambda$ window is independent, these calculations can be run in parallel across multiple GPUs with no hit to performance and can be done in less than 20 minutes on a GTX 1080Ti across 24 GPUs, or 12 minutes across 24 RTX 2080Ti GPUs. Figure 1 summarizes the results of AMBER20 on three targets of different size with GeForce 1080Ti and 2080Ti graphics cards using standard MD and GTI.

## 4 Advances in AMBER20

A number of important improvements were introduced in AMBER20 to facilitate large-scale RBFE and ABFE simulations. Specifically, the softcore potential was improved using a smoothstep function,[58] which significantly reduces a number of known issues in previous versions of AMBER (namely the end-point catastrophe, particle collapse, and large gradient jumps in the $dU/d\lambda$ curve). Additionally, Boresch restraints[59] have been implemented, which can be used in an automated fashion for ABFE simulations with many diverse ligands. Boresch bonded terms[95,96,120] were also implemented, which can be used to control which energy terms are included in the softcore region. Below is a more detailed description of the advances in AMBER20.

### 4.1 Smoothstep Softcore Potentials

With the newly developed class of smoothstep softcore potentials described in 2.1, we were able to demonstrate that, unlike the conventional softcore potential in previous versions of AMBER, there is a single set of $\alpha$ and $\beta$ values that can be utilized for reliable and accurate simulations across a wide range of diverse molecular systems.[58] The key characteristic of the smoothstep softcore function is that the weights used in the alchemical transformation have derivatives that vanish at the transformation endpoints ($\lambda$ = 0 and 1) and enable smooth adjustment of the $\lambda$-dependent terms in the potential. The second-order smoothstep softcore potential, SSC(2), with $\alpha$ = 0.2 and $\beta$ = 50 Å$^2$ has been demonstrated to overcome all three problems for a broad set of alchemical transformations used in the calculation of hydration free energies and RBFEs. Results are examined for edge cases where the original AMBER softcore potential is observed to fail – the SSC(2) smoothstep softcore potential was demonstrated to remain accurate. The SSC(2) potential has been further tested against a broad set of hydration free energy and RBFEs for a commonly used FEP validation dataset containing 200 ligands and spanning 8 protein targets.[121] The SSC(2) potential has the advantage that it can be used in concerted transformations and is better suited for enhanced sampling methods with more advanced, adaptive $\lambda$ scheduling requirements, which is part of our ongoing research collaboration and intended to be in upcoming AMBER releases (see section 6.3.

In AMBER20, the $\lambda$-dependence of individual interactions (e.g., bonded, Coulombic and Lennard-Jones) now can be controlled by the user, including both linear and smoothstep functional forms, and advanced $\lambda$-scheduling within the $\lambda$ interval [0,1]. This "$\lambda$-scheduling" can be applied to individual interactions independently and gives the users a very flexible way to control the mixing scheme of the softcore potentials. For example, one can utilize a smoothstep function with boundaries at [0.0,0.5] for Coulombic (Coul) interaction and a smoothstep function with boundaries at [0.5,1.0] for Lennard-Jones (LJ) potential, which will execute a stepwise alchemical transformation with the Coulombic interactions being transformed in the $\lambda$ interval [0.0,0.5] and the Lennard-Jones being transformed in the $\lambda$ interval [0.5,1.0]. Similar $\lambda$-scheduling features have been reported and implemented in other simulation packages such as GROMOS[92] and NAMD.[84,122] Application of these $\lambda$ scheduling features in combination with the new smoothstep softcore potentials are discussed in more detail in Future Work (section 6) below.

### 4.2 Pose Restraints for ABFE

Accurate predictions of absolute binding free energies of small organic molecules from MD simulations offer significant value in drug discovery and design. In particular, ABFE (as opposed to RBFE) is not restricted to perturbations on a common core and is thus amenable for use in virtual screening,[4] selectivity screening,[123] and core hopping.[124] However, this added flexibility also introduces new challenges and uncertainties, which may explain why ABFE has seen minimal use in actual drug discovery projects (in addition to the additional sampling requirements).

There are multiple ways to realize a valid thermodynamic cycle that is compatible with the aims of ABFE, but most schemes employ a set of restraints to restrict the ligand to remain

near the binding site.[59,125–128] The necessary (or allowable) extent of this restriction is rather dependent on the system. For example, some strategies employ loose “flat-bottom” restraints that only restrict the ligand center of mass motion,[128] while others accommodate more elaborate restraints on the ligand conformation, translation, and rotation,[125] possibly in multiple unique poses.[126,127]

In our drug discovery efforts, we have worked under the assumption that the unrestrained ligand is fairly strongly bound and therefore can be assumed to occupy a single pose with high occupancy. This assumption is generally safe because we are primarily interested in the identification of tight binding species and we can tolerate bias or inaccuracy in weakly bound species – other use cases may have different needs, such as identifying weakly bound fragments[129] or molecules that stabilize intrinsically disordered proteins.[130] Working under the assumption of a single well-defined binding mode permits the rather simple restraint framework described by Boresch *et al.*,[59] which only places harmonic translational and rotational restraints on the ligand in a local coordinate frame via one distance, two angles, and three dihedrals (see Section 5.7 for further details). The implementation in AMBER20 also permits these restraints to be included in the overall alchemical transformation such that the component of the free energy arising from the restraints in the bound state can be computed in the same way as other force field terms. The absence of restraints when simulating the unbound state can be accounted for using a simple analytic formula in the limit that the harmonic restraints are relatively stiff.[59,131] The implementation in AMBER20 has been validated on virtual screens with thousands of diverse compounds run through ABFE. The results demonstrate the usefulness of the approach and confirm that the calculated binding free energy is independent of the details of the restraints, as determined by comparing results from multiple runs with randomized restraint combinations.[4]

### 4.3 Handling Interactions Involving Softcore Atoms for RBFE

In AMBER, the TI region is defined as the part of the system to undergo alchemical transformation from one end state to another one; hence there are two regions representing two end states. There are two parts for each TI region: the common region to both TI regions and the softcore region unique to each TI region. The softcore potential is utilized to treat the interactions between the softcore regions and other parts of the system. Atoms which are growing or disappearing during the alchemical transformation must be included in the softcore region. An atom included in a softcore region is defined as a softcore atom. Previous versions of AMBER have not allowed for detailed control over the interactions between the common and softcore regions. While most of the time treatment of these terms will not cause significant deviations of the calculated free energy differences, theoretically it should be treated more rigorously when applicable.

**Bonded terms between the common and softcore regions—**A key advantage of RBFE simulations is that the ligand scaffold (common core) is always present and interacts with the receptor binding site, thus obviating (or at least greatly reducing) the need for orientational restraints as with ABFE. However, a similar issue is encountered when chemical groups extending off of the common core are created or annihilated (transforming from, or into ”dummy atoms”) – similar to the ligand drift problem in ABFE, the chemical

group must be tethered to the common core. While this may seem readily accomplished by retaining bonded terms between the disappearing group (softcore region) and the common core, care should be taken that these retained bonded terms obey certain constraints and conditions. These conditions require the ensembles generated in the state with “dummy atoms” in the softcore region that have ”disappeared” to reproduce the same potential of mean force on the real atoms as the real system without the dummy atoms. Extensive discussion of these conditions, including ”rules” of how to select retained bonded terms between the softcore and common core regions, are provided by Boresch[95,96,120] and Roux and co-workers.[84,132]

AMBER20 now includes a general facility for selectively retaining bonded interactions with non-interacting softcore atoms. The non-retained terms are then decoupled using the usual scheduling strategies for non-bonded interactions. Importantly, the simulation efficiency can be highly affected by which terms are selected for retention and which are not – poorly chosen terms can lead to high variability or even non-ergodicity. Unfortunately, there does not appear to be a general solution to this issue. Theoretically rigorous results can only be obtained by retaining terms that involve not more than three atoms in the common core. However, for efficiency, the retained terms must also keep the softcore atoms in or near a physically relevant geometry and not hinder rotameric transitions.

**Non-bonded terms between the common and softcore regions—**The non-bonded terms between the common region and the softcore regions should be always scaled with the alchemical variable $\lambda$. Nevertheless, the 1–4 non-bonded terms across the softcore boundary were not treated properly in some previous versions of AMBER.[98] A fix has been implemented and verified in AMBER20, resulting in much improved relative hydration free energies of 9 benchmark molecules using the concerted transformation protocol.[55]

**Interactions within the softcore region—**In AMBER20, both bonded and non-bonded interaction terms within the softcore regions can be either scaled with $\lambda$ or not. Either implementation is theoretically correct, provided that the conformational sampling of the softcore regions at the end point states are sufficient. Users can control how the interactions within the softcore regions are treated. For recommended guidelines, refer to recent validation studies of free energy methods in AMBER.[55]

### 4.4 RBFE Accuracy on Drug Targets

The GPU-accelerated free energy simulation methods in AMBER have been validated in an Application Note appearing in the current special issue.[55] Although the methods discussed here are quite new, they have already seen a number of applications,[56–58] particularly against a well-studied data set that includes $\tilde{2}00$ ligand mutations spanning 8 protein targets (Bace, CDK2, Jnk1, MCL1, p38, PTP1B, Thrombin and Tyk2).[121] This data set serves as a tractable benchmark set for RBFE calculations because there are no known significant conformational changes or other challenging scenarios such as ambiguities in tautomer/ionization states or buried waters. Results using the protocols described in this work with the GAFF2 force field[133] and TIP3P[14] water model are on par with other recent RBFE

publications on the same data set, as seen in Table 1, and stand to improve with the forthcoming release of new MM and QM force fields for ligand-binding predictions.

## 5 Practical Considerations

The aforementioned topics (force field, sampling, and alchemical parameters) are critical to achieving robust and reliable BFE results, yet many publications under-emphasize the importance of other considerations. Items such as system preparation, docking, and confidence estimates can be just as important as force field, sampling, and alchemical parameters for obtaining robust BFE predictions. In some cases, it is possible to define best practices and even automate the process to some extent. However, other instances may require expert decisions on a case by case basis. Below, we highlight practical considerations found to be most important in prospective applications of free energy simulations in AMBER. When possible, we provide guidance for best practices, yet in other cases we simply highlight the challenges and leave it to the reader to further explore these areas. More details about practical considerations in alchemical binding free energy simulations can be found in a Perspective by Cournia *et al.*[3]

### 5.1 Force Field

A force field is used to model the interactions between atoms in the molecular system of interest. The force field allows determination of potential energy as a function of configuration and is used along with the kinetic energy to calculate the Hamiltonian for molecular dynamics simulations and binding free energy calculations. The accuracy of the force field may limit that of the binding free energy predictions, but not all inaccuracies in the predictions should be blamed on force field problems: poor quality in the initial structure, erroneous protonation or tautomeric states, and inadequate sampling should first be inspected. One lesson that we learned from the past decade is that the force field accuracy can be substantially improved by simply avoiding the obvious mistakes in the parameterization of ligand molecules.

Generalized force field models such as GAFF,[133,134] CGenFF,[135,136] and OPLS[137–140] represent efforts to provide force field parameters for any molecule at a small computational cost, using look-up tables for parameters predetermined for different bond, angle, and torsion types. Such models and the associated software tools are good starting points for parameterizing molecules for binding free energy calculations. They provide reasonable parameters for molecules consisting of chemical structures similar to those in the training set. Yet in real-world applications, it is not uncommon to encounter molecules that these generalized force fields have not been tuned for and thus do not yield accurate results.

With advances in computational hardware and GPU-enabled quantum chemistry software such as TeraChem,[141] it has become feasible to parameterize hundreds of small molecules–as commonly required each week in drug discovery programs–individually based on detailed quantum chemistry calculations: a complete force field parameterization of a small molecule may be performed in approximately 1 GPU-hour. Such bespoke molecular force fields help to avoid gross parameterization errors, and often lead to improved free energy results.

One common type of error in force field parameterization is in the torsional parameters that determine the potential energies at different torsional angles of a rotatable bond in the molecule. For example, a biphenyl system with substitutions at the ortho-, meta-, and para-positions can substantially perturb the torsional energy profile around the bond connecting the phenyl rings, and the perturbation depends strongly on the position and the moiety of the substitution. As a demonstration of the benefit of refitting the torsional parameters to the quantum chemistry calculations of the torsional energy profile, Figure 2 shows the hydration free energies computed for a set of alcohol molecules, comparing the generalized GAFF2 and the bespoke force field in which the torsional parameters are refit; the latter significantly improves the agreement between the predicted hydration free energies and the experimental measurements.

Generalized force fields sometimes fail to capture the electrostatic potential around the molecule. A well-known example is the $\sigma$-hole in aromatic halogens,[143] in which a "hole" of positive potential along the carbon-halogen bond cannot be reproduced by the common atom-centered charges. Inclusion of off-atom-center charges, or virtual sites, is an effective approach to resolving such discrepancies. Such virtual sites for select functional groups are now finding their way into generalized force fields, but transferable parameters take onerous efforts to derive. In contrast, they are straightforward to parameterize for bespoke molecular force fields. Figure 3 shows how their inclusion in our bespoke force fields for aromatic halogens and aromatic nitrogens improves the fit for the electrostatic potential computed by DFT and for the predicted hydration free energy.

Parameterizing bespoke molecular force field is associated with a smaller computational cost than the binding free energy calculations, yet they may significantly improve the predictive accuracy. A number of automated tools for parameterizing small molecules, including several in the public domain, have been developed, such as CGenFF,[135,136] GAAMP (https://gaamp.lcrc.anl.gov/index.html), FFTK,[144] and the tools developed by the OpenFF Initiative.[145] We believe that automated programs for bespoke parameterization will become the default option in future applications of binding free energy calculations.

### 5.2 Protein Preparation

Protein structures must be prepared prior to running MD free energy simulations. While X-ray structures are the most common source of atomic-resolution structures, the following guidelines also apply to structures obtained by other means, such as NMR or Cryo-EM. In general, structures must be prepared to add hydrogen atoms, optimize hydrogen bond networks, remove atomic clashes, in some cases insert regions missing from the refinement such as disordered loops, and perform other operations that are not part of the experimental structure refinement process. While the prerequisite for good system preparation is generally accepted in the field, the specific steps are not well defined. Fortunately, many of the considerations for BFE simulations are similar to those for other structure-based approaches, like docking, and have been described in detail in other works.[146–148] Nonetheless, docking for BFE simulations may entail additional considerations beyond standard docking calculations for pose prediction or virtual screening, as described in the Docking section below.

It should be noted that protein preparation can have a significant impact on the quality of results and can introduce artificial biases, especially in the case of retrospective validation studies, as has been demonstrated for docking studies[149] and likely has similar issues in BFE calculations. After protein preparation, including the following steps, it is recommended to manually inspect the structure, run protein analysis programs (e.g. PROCHECK,[150] WHATCHECK,[151] MolProbity,[152] and SurVol[153]), and perform MD simulations[148,154] to ensure stability of the system before running computationally costly BFE simulations. For example, multiple short simulations on different protein preparation states can reveal problematic cases where there are large structural fluctuations, degradation in secondary structural elements, or loss of key binding site interactions.[155] Additionally, such MD simulations can be used to improve the overall protein structure.[156–158] Details of the protein preparation capabilities and options for AMBER20 can be found in the user manual (https://ambermd.org/doc12/Amber20.pdf).

**Hydrogen bonds—**Hydrogen atoms are not typically present in experimentally-determined structures (other than those at resolution better than ~1.0 Å) and therefore need to be added computationally. The initial coordinates of hydrogen atoms are inconsequential, as long as proper valences are satisfied and subsequent sampling is performed. The protonation state of titratable residues should be determined for the pH of interest (typically this involves His, Asp, and Glu, although this could be expanded to Lys and Cys). Additionally, the two His tautomers should be sampled (proton on the N$\delta$, N$\epsilon$, or both). Programs such as WHATIF[159] can be used for this step, which can be augmented with $pK_a$ predictions programs such as PROPKA.[160]

Once hydrogen atoms are added, the H-bond network should be optimized by sampling 180° flips of the terminal chi angle for Asn, Gln, and His, which significantly changes the spatial H-bonding capabilities of the side chains, but does not appreciably change the fit to the electron density. In addition, hydrogens on hydroxyls and thiols should be sampled to optimize the H-bond network. After the above steps are completed, it is recommended to perform an analysis of the structure to ensure a viable state has been generated. Automated programs such as WHATIF, PROCHECK, and MolProbity[152] are useful for the analysis, although manually inspecting changes in the atomic fit to electron density with programs such as Coot[161] is strongly recommended. If it is unclear which state(s) are correct, it is recommended to perform modest MD simulations (on the order of 100 ns) and structural analysis to determine the stability of the structures. The aforementioned protocol for H-bond optimization is necessary because sampling times required to overcome sampling barriers needed to rectify an incorrect initial state could be prohibitively long, and local denaturing of the protein can occur in the process, which would take even more sampling time to correct, if it could be corrected at all. Nonetheless, if sufficient sampling time is achievable or enhanced sampling approaches are available to solve this problem, then simply adding hydrogen atoms as needed to satisfy valencies for the pH of interest may be sufficient.

**Waters—**The treatment of explicit water molecules can influence docking accuracy and enrichment results, as has been extensively demonstrated by other works.[162–167] Indeed, water is the source of the hydrophobic effect[168] – subtle changes in waters can impact ligand

binding energetics[169] and even reverse the thermodynamic signature of ligands binding to a protein.[170] The treatment of water molecules in BFE calculations should be considered during the initial system setup and during the alchemical simulations themselves. The determination of which waters to retain during the setup of BFE calculations is often unclear, primarily because not all waters are present in crystal structures and even when there are many waters, the free energy of a water molecule is not directly related to the crystallographic occupancy. Furthermore, the crystal structure being used might not correspond to the specific ligand or ligand series being explored and therefore the water pattern may be inaccurate.

In most cases, it is recommended that all crystallographic waters are retained for the system setup, although the electron density should be inspected to ensure that there is confidence in the water presence. Even in cases where there are many water molecules in the experimentally determined structure, it is generally necessary to add additional water molecules before MD simulations. Programs such as 3D-RISM,[171] GCMC,[172] JAWS,[173] WaterMap,[174] and other approaches[175–178] can be useful for this step, since it is generally a fast calculation relative to the BFE simulations. Importantly, the method to place water should be capable of solvating buried pockets that are challenging to sample during the simulation time of a BFE run due to large energetic barrier for entering/exiting the binding pocket.

Once waters have been placed for the initial system setup, it still might be necessary to explicitly sample waters (beyond MD sampling) during the alchemical simulations. This is especially important when dealing with regions of the binding site that are occluded from exchange with bulk solvent, such as fully buried binding sites or subpockets that are blocked from bulk solvent exchange due to parts of the ligand that are not being perturbed (in the case of RBFE). A combined MC/MD method has recently been described and is available in AMBER20,[179,180] which allows water to equilibrate between bulk and buried cavities. This method allows for partial water densities during the BFE calculation by allowing the locations and occupancy of buried sites to vary with $\lambda$ in the course of alchemical calculations. Other approaches, such as Grand Canonical Monte Carlo (GCMC), have been proposed to address buried water sampling in the context of alchemical free energy calculations.[181–183]

### 5.3 Ligand Preparation

All-atom three-dimensional (3D) ligands are required for RBFE and ABFE simulations. As such, a critical issue to investigate before embarking on computationally expensive free energy simulations is to generate the correct ligand state (ionization, tautomers, stereochemistry, etc.). Incorrect states could result in false negatives (*e.g.*, where a favorable H-bond cannot be made) or false positives (e.g. where an incorrect H-bond is made). In addition to calculating reasonable ligand states, it is ideal to predict an energetic penalty associated with each state to account for the energetic cost it takes to generate each state in solution. This energetic penalty should then be added to a computed free energy to get a final binding prediction.

Significantly, ionization and tautomerization energies are absent from traditional molecular mechanics force fields, since there are no terms for bond making/breaking. Some empirical corrections are possible via methods like constant-pH MD, but these may not be cost-effective for the large number of ligands studied in a drug design setting. Although largely speculative at this point, we suppose that advances rooted in quantum mechanics and/or machine learning will be necessary for progress in this area.

### 5.4 Docking

Docking is an essential part of binding free energy simulations, while in theory the binding free energy results should be independent of input pose, that assumes sufficient sampling to explore all accessible poses with MD, which would be prohibitively computationally expensive. As such, it is critical to obtain a reasonable initial pose and in cases where the bet pose is ambiguous, then multiple poses should be pursued. The nature of the docking problem is different between RBFE and ABFE (and different from docking as a final calculation): For RBFE calculations a reference pose is typically known and can be used to constrain the docking whereas with ABFE there is typically no reference molecule and therefore unconstrained docking is required.

**Docking for ABFE—**Docking for an ABFE calculation can be quite challenging, especially if one does not have any prior knowledge that can be employed when evaluating docked poses. In addition to sampling the ligand conformation/orientation, the receptor might undergo induced-fit.[184] As such, in order to generate a reasonable starting pose, it may be necessary to induce the site.[185] Numerous approaches have been developed in order to address this issue ranging from employing a softened non-bonded potential in order to alleviate the penalty of protein-ligand clashes, followed by a robust protocol that incorporates sampling different side-chain rotamers of the receptor and redocking the compound to multiple receptors (ensemble docking).[186–188] Cases where large-scale backbone motions are required to generate the correct binding pose still remains a challenge for the field even when incorporating enhanced sampling methods.

Fortunately, it has been shown that combining docking with molecular dynamics (MD) to further refine the pose can be beneficial.[186,189] One attractive feature of coupling MD with docking is that the receptor and ligand are sampled simultaneously in the presence of explicit water, allowing for the receptor to become induced in a physically meaningful way. For example, one might generate N docked poses and run a MD simulation in replicate varying the random seed to assess pose stability. As an additional example, to prioritize poses for more rigorous free energy simulations (e.g. ABFE), multiple short MD simulations can be performed and the RMSD from the docked pose can be utilized as a metric to assess pose stability. Generally, low RMSDs are attributed to the ligand making energetically favorable interactions within a targeted site.[190–192] It is important not only to consider the averaged RMSD value but also evaluate the RMSD versus time as a ligand could have adopted a stable conformation that is substantially different from (e.g., larger than 2.5 Å RMSD) from its docked pose and remained there for the duration of the simulation. Excluding atoms of the ligand that are very solvent accessible from the RMSD calculation may also be required. While MD refinement offers advantages over traditional docking

(protein flexibility, explicit water, etc.), it does not contain a rigorous treatment of the binding thermodynamics (e.g. no unbound ligand calculation) and therefore should not be used as a final scoring estimator.

**Docking for RBFE—**Docking for a RBFE calculation is theoretically simpler when compared to ABFE due to the conserved binding mode of most congeneric molecules. The technical challenges of constraining ligand atoms, especially when done in a high-throughput automated fashion for drug discovery, presents challenges. Core-constrained docking is generally the most effective way to generate poses, which requires a definition of the core atoms (either manually or based on a maximum common substructure, MCS). As seen in Figure 4, core constrained poses produce much cleaner alignments, which facilitate the atom mapping and stability of the perturbations in RBFE simulations. Open source docking programs such as rDOCK[193] support core constraints, although many commercial solutions are also available. Common atoms shared between the lead and the candidate ligand are constrained while the degrees of freedom of other atoms are sampled during the docking calculation. It should be noted that using an MCS is not always optimal, as the 2D mapping does not ensure the correct 3D characteristics (see section 5.5).

Steric clashes present another challenge when docking ligands that do not fit into the rigid receptor and therefore should be handled carefully. As a consequence of the core constraints, significant protein-ligand clashes might be unavoidable while still satisfying the core positional constraints. Generally, these types of issues can be resolved through an energy minimization and/or short restrained MD equilibration only allowing key atoms to move, or redocking the compound with a reduced number of constrained atoms. Another important item to note is that if the protein residue(s) involved in the steric clash have to move significantly in order to alleviate the clash during the minimization or MD equilibration prior to running the RBFE calculation, then the energy required to adopt this new protein conformation will not be accounted for during the RBFE simulation, which could possibly lead to erroneous results. Therefore it may be advantageous to make a series of smaller perturbations or instead run ABFE.

### 5.5 Atom Mapping

For RBFE calculations, a critical step is determining the relationship between atoms of the reference and perturbed structure such that the common atoms (“mapped” atoms) are linearly interpolated with $\lambda$ and the unmapped atoms are treated with a softcore functional form to allow for their insertion or deletion. Theoretically, the best atom mapping scheme is one that minimizes the thermodynamic path between the two molecules, however there are many factors to consider in practice such as atom type, bond order, ring membership, chirality, and binding conformation.

Generally, topological similarity is assessed computationally using a maximum common substructure (MCS) algorithm, which aims to maximize the number of mapped atoms between two molecules from a congeneric series.[112] Many MCS algorithms require specification of atom type, bond order, and ring membership considerations to define the maximum atomic overlap between molecules. Assuming perfect geometric complementarity

between molecules, mapping of atoms that differ in type should maximize phase space overlap between states by decoupling as few atoms as possible. In practice, many times this is not the case due to conformational differences between molecules that can lead to convergence issues and large errors between neighboring $\lambda$ values along a thermodynamic path.

In some cases the mappings are clear, such as the substitution of an aromatic para-fluorine for a para-methoxy, where the fluorine and methoxy are the only unmapped atoms. However, in other cases the mapping can be less clear (or even ambiguous), such as bulky ortho/meta substitutions to a similarly substituted phenyl ring. In addition, mapping of atoms with different bond orders can be problematic as atomic torsional preferences change between atomic environments and such mappings should be avoided whenever possible.

The mapping of atoms within ring systems requires special consideration and introduces a potential source of error propagation if mapped inappropriately, such as allowing for ring breaking/forming.[3] As such, most atom mapping protocols avoid ring breaking when possible. Previous literature has demonstrated that bonded term contributions from dummy atoms should cancel in RBFE simulations of the bound and unbound states.[194] Yet if the conformational ensemble of the molecule is significantly affected by the remaining core atoms, as is the case for members of a ring, the cancellation of error is no longer valid. It is for this reason that large errors are often observed in RBFE calculations involving ring breaking/forming, as the free energies are only collected from a restricted and inaccurate conformational ensemble. To address these conformational restrictions, recently a “soft bond” potential has been added to the softcore functional form and suggests that improvements to core hopping transformations can be made. Still, more work is necessary to demonstrate its utility across broad ring breaking/forming scenarios.[124]

When performing manual RBFE calculations it is often straightforward to determine the correct mapping between ligands, especially if the binding poses are well determined. However, manual mapping is tedious, time consuming, and error prone, especially when processing hundreds of molecules on a weekly basis. As such, manual mapping is impractical in drug discovery applications, where hundreds of molecules will be explored on a weekly basis. There are programs that perform automatic mapping, such as LOMAP,[195] which operates on 2D structures (typically using a 2D method such as maximum common substructure). It should be noted that in some cases the mapping is ambiguous based solely on the 2D information, such as ortho substitutions to a similarly substituted phenyl ring, as seen in Figure 5A. In this example, the preferred conformation of the ortho methoxy substituted molecule is forming an intermolecular hydrogen bond between the amide nitrogen and the methoxy oxygen. However, the bulky chloro substitution prefers an alternate conformation, and using only 2D topological information the oxygen is oriented towards the same substitution vector. Mapping issues such as this will often lead to incorrect results due to the unreasonably long thermodynamic path between the states (a high-energy conformational transition would be required to interconvert between the two states). Fortunately, mapping based on 3D poses would yield the correct result, as seen in Figure 5B. As such, it is highly recommended to perform atom mapping using accurate 3D poses when possible.

To our knowledge, there is currently no widely adopted tool for atom mapping based on 3D poses. A sensible approach is to align the two molecules–say, $A$ and $B$–in their binding poses and preferentially map each atom in $A$ to an atom in $B$ that is spatially close. This may be formulated as a discrete optimization problem: one can define a quantitative measure of spatial overlap between each pair of atoms, and then find the graph-isomorphic mapping that maximizes the total overlap of the mapped atom pairs. We anticipate that such 3D atom mapping tools will eventually replace the current 2D atom mapping tools.

### 5.6 $\lambda$ Schedule

Alchemical BFE simulations are performed by defining a transformation (e.g. between two different bound ligands). The extent of the transformation is defined in terms of a coupling parameter, usually denoted as a value $\lambda$ between zero and one. As such, the intermediate steps involved in the perturbation are often referred to as "$\lambda$ values" or "$\lambda$ windows". We use the term "$\lambda$ schedule" to refer to 1) the number and placement of the specific values included in the simulations and 2) the functional form of the coupling in terms of $\lambda$ (e.g. use of a softcore potential). One would like to choose the $\lambda$ schedule in an optimal way. The answer to this problem is strongly dependant on the methodology being used and decisions to be made by the practitioner. In what follows we will assume that a conventional alchemical approach is being used and that multiple $\lambda$ values will be chosen with simulations carried out at each value with $\lambda$ held fixed. Other simulation approaches may permit variation in $\lambda$ either as a discrete[196] or continuous[197] quantity, but these are currently outside the scope of what constitutes a best practice in AMBER.

In general, the aim is to have enough, but not too many, $\lambda$ windows in order to obtain sufficient accuracy at the lowest possible cost. From the perspective of TI, this means sampling the integrand more densely in regions where the curvature changes rapidly and possibly spacing the values so as to abide by a numerical quadrature rule. In the context of FEP-like protocols (e.g. MBAR), this means choosing neighboring sampling distributions to achieve minimum variance behavior with respect to a set of Monte Carlo moves (see, for example, the overlap metric introduced by Bennett[7]). Optimizing according to either of these schemes requires *a priori* information. Lacking this advantage, the most straightforward approach is to use equally spaced values and a generic quadrature scheme such as the trapezoidal rule. In this case, TI essentially reduces to a piece-wise linear approximation of the integrand and is roughly equivalent to approximating the neighboring sampling distributions as Gaussians.[198] Using this approximation can still provide good results for calculations where the lambda spacing is small enough to capture the essence of the variations in the integrand with lambda throughout the full $\lambda$=[0,1] trajectory.

Interestingly, for simple transformations that only add, remove, or change a charge distribution in a limited volume the TI integrand tends to be approximately linear or perhaps cubic[199] (Figure 6, top row). This follows from a Born-like linear response model where a charge or point dipole is introduced into a spherical cavity in a homogeneous dielectric environment.[198] In this case, the integrand may be extremely linear and well-behaved and only a few $\lambda$ values could be required (only two points are needed to accurately integrate a line). Any additional curvature tends to occur near the endpoints, although the general

prescription is to focus $\lambda$ values where the slope is largest (usually at intermediate values).[198] Strong deviation from linearity could simply indicate that the charge change is occurring over a large extent and the linear response character could be breaking down. An approach that still samples intermediate values is thus recommended in general, rather than assuming linearity. We have found that as few as five $\lambda$ values can give reliable results for small perturbations (assuming other sources of error are considered, as discussed below). Indeed, it has been shown that in some scenarios a one-step $\lambda$ schedule can be sufficient to achieve accurate binding free energy predictions,[200–202] although such cases of very small perturbations provide insufficient coverage of chemical space to have high impact for most drug discovery applications.

More complicated transformations, especially those that introduce short-range repulsive interactions, are by far the most difficult to handle (Figure 6, bottom row). The Lennard-Jones potential is widely understood to introduce large variance and/or singularities which were originally overcome by introducing *many* finely spaced $\lambda$ values near (but not at) the endpoint. This approach was supplanted by the introduction of softcore potentials which tend to redistribute, but not entirely eliminate, the higher variance across the more intermediate points. The variance of the result is generally proportional to the size of the chemical group being introduced. For example, one should expect higher uncertainty from an ABFE calculation of a drug-like molecule compared to removing/inserting a small chemical moiety onto a ligand scaffold. Unfortunately, introducing many additional $\lambda$ values does not seem to mitigate this issue beyond a point – one only needs enough values to capture the shape of the integrand. This is because the insertion of uncharged, repulsive interactions generally leads to configurations of low physical relevance and so the variance is inherently a sampling issue. When time and resources are available, then more $\lambda$ values can be added to enhance overlap between adjacent windows and thereby improve reliability of results.[83,202]

### 5.7 ABFE Pose Restraints

The purpose of pose restraints in ABFE is to hold the ligand in the binding pocket when the interactions are scaled to extremely small values (or zero). At the same time, the restraints can also be interpreted as *defining* the bound microscopic state.[125,128] Therefore, a reasonable criterion is to require that the restraints impose an orientation that is similar to the fully interacting ligand. Put another way, the restraints should approximate the potential of mean force of the physical system. A similar perspective has been offered in several theoretical frameworks[125–128] and more elaborate choices than the one described here could also be employed using AMBER. In general, the assumptions here hold for relatively strongly bound compounds and different restraint protocols may work better in other regimes.

A procedure that we have found effective is to first perform a relatively short (~5 ns) non-alchemical simulation of the ligand-protein complex. We then search for relatively stationary points on both the ligand and receptor that can be used to define both their relative orientation as well as the internal conformation of the ligand, as described by Kim, *et al.*[131] For proteins, we look for low-mobility, buried residues by searching for minimal solvent-

exposed surface area over the course of the trajectory – the $\alpha$-carbons of these residues are then considered good candidates, but other choices are certainly possible (e.g. the backbone center of mass). For ligands, we look for heavy atoms within rigid scaffold motifs such as fused ring systems or the central core of the ligand. One should be careful not to select multiple atoms within a rotatable torsion, otherwise one might "lock-in" the configuration and induce non-ergodicity. Candidate coordinates can then be created from both groups and the six terms (one distance, two angles, and three dihedrals) can be tracked over the trajectory.

The ideal coordinates are unimodally (perhaps Gaussian) distributed and have low variance. In AMBER it is also useful to avoid overly long distances ($<30$ Å, say), noncolinear angles (far from 0° or 180°), and dihedrals that are from the periodic boundary (i.e. not near ±180°). Any combination of atom or point selections that fit these criteria should constitute a reasonable set of restraints.

### 5.8 Periodicity and Charge Corrections

The use of periodic boundary conditions (PBCs) has long been known to introduce subtle artifacts in MD simulations. However, the alternative of no boundary conditions is generally not preferred because it would induce artifacts of a different nature (and larger magnitude). While some methodologies may avoid different issues, the general philosophy in AMBER is that the PME scheme for PBCs is the best compromise between accuracy and efficiency.[44,45] In current AMBER20 implementation, the thermodynamic derivative of the PME reciprocal part is calculated in the linear way (,**eq:UL** i.e. the PME reciprocal calculations are only done on the end states and the $\left\langle \frac{dU}{d\lambda} \right\rangle$ is the difference of the PME reciprocal energies of the end states. This approach is simple but requires two PME calculations per MD iteration.

Nonetheless, specific care must be taken in alchemical simulations and the issues are unusually pronounced for alchemical transformations that do not conserve the net charge. Rocklin, *et al.*[203] recently catalogued these issues with an eye towards alchemical ABFE calculations and proposed specific approximation schemes for correcting them. A follow up work by Chen, *et al.*[204] also examined the Rocklin corrections in the case of RBFE, along with other possible solutions. These issues are briefly described here.

The most significant artifacts due to PBCs arise from the mean electrostatic potential definition imposed by PME. An extensive review has been supplied by Lin, *et al.*[205] This term is normally innocuous, as it amounts to a simple shift in the zero of energy and does not affect forces. However, alchemical simulations are extremely sensitive to arbitrary shifts in the zero of energy as this effectively shifts the binding free energy of a ligand based purely on the system charge. Following Lin, *et al.*, the artifact can be considered as the work required to move a charged species across the boundary between a solvent and vacuum – clearly no such boundary exists under PBCs and so this contribution is missing. Rocklin, *et al.* refer to this shift as arising from a residual integrated potential (RIP), as it corresponds to the energy "left over" when a charge species is removed from a PBC box. The RIP energy is proportional to the integral of the mean potential over the whole volume. Lin, *et al.* describe how a correction can be approximated *a posteriori* from the spatial electrostatic potential

averaged over a trajectory and then integrated over the box. Fortuitously, along with others[204] we find that the approximate Poisson-Boltzmann scheme proposed by Rocklin *et al.* provides reasonable corrections based on only a single structure. In practice we see minimal statistical noise of ~0.1–0.2 kcal/mol, which is quite acceptable given the expected accuracy of ABFE.

Other PBC-based corrections are possible but generally much smaller in magnitude than the RIP correction.[203] The net charge interaction and solvation correction terms are generally quite small and only depend on the magnitude of the charge change and box size/shape. Another term unrelated to PME but dependent on charge is the discrete solvation correction, which is meant to compensate for the distortion in solvent structure when, for example, water molecules interact with their overly ordered image counterpart. Interestingly, this term can be quite large for a given ligand-receptor complex (~ 10 − 20 kcal/mol), but it is generally much less sensitive to the differences between the ligand bound and unbound states and so tends to cancel extensively. Note, some care must be taken because the correction depends on the specific nature of the solvent model (namely the charge distribution).

### 5.9 Confidence and Error Analysis

As has been amply discussed above, there are numerous reasons that a BFE prediction may be incorrect. False positive predictions (compounds predicted to be good that are not) can be costly in drug design projects, where synthesizing an inactive compound can cost thousands of dollars and weeks of lab work. Therefore one should leverage the relatively low cost of additional simulations in order to establish confidence (or lack thereof) in a prediction. Here we suggest several best practices for appropriately assessing the quality of binding free energy predictions as well for building more robust hypotheses around the chemical matter surrounding a target. Other works have discussed techniques for confidence assessment and error analysis in binding free energy calculations,[206] including ways to improve free energy predictions by meaningful error estimates.[79] Sources of errors can also be attributed to validation data sets and data set biases, as examined in detail elsewhere.[207] Below we discuss ways to assess some of the most common sources of errors in BFE calculations, including statistical analysis, structural analysis, binding pose uncertainty, sampling, and force fields.

**Statistical Errors—**There are two main sources of random error in free energy estimates in addition to the systematic error related to the quality of the model potential energy function. These contributions are:

1. The uncertainty in the ensemble average quantities due to fluctuations in the observed sample distribution.
2. The uncertainty caused by approximating the true distribution via finite-length simulations.

The uncertainty of the ensemble averages within the observed distribution can be estimated by considering only the statistically significant samples to calculate the standard error of the mean. The statistically significant samples are those data points separated in time by at least

the autocorrelation time of the timeseries. Upon pruning the correlated data of observable $x$, the sample standard deviation $\sigma_x$ is calculated, and the standard error of the mean is then $\sigma_{\langle x \rangle} = \sqrt{\sigma_x^2 / N_x}$, where $N_x$ is the number of statistically independent samples of $x$. For TI calculations, the observables are the timeseries values of $\partial U/\partial\lambda$. The estimated error in the free energy due to the uncertainty within the observed distribution requires the calculation of the standard errors $\sigma_{\langle \partial U/\partial\lambda \rangle_\lambda}$ and propagation of these errors through the quadrature formula; that is,

$$\sigma_{\Delta G} = \sqrt{\sum_i w_i^2 \sigma^2_{\langle \partial U/\partial\lambda \rangle_{\lambda_i}}} \tag{17}$$

Shirts and Chodera developed an analytic expression for the (large sample size) free energy errors calculated from the MBAR method, which is more complicated due to MBARs use of coupled equations.[9]

An alternative approach for calculating the standard errors is *bootstrapping*. The bootstrap algorithm requires one to generate many estimates of the free energy and then calculate the standard deviation of those estimates, which is the standard error of the mean. To generate many estimates of the free energy, new time series values are artificially created by resampling the observed distribution with replacement. That is, if the observed distribution contains $N$ data points, then a bootstrap distribution containing $N$ points is created by randomly selecting samples from the observed distribution. In the case that the observed distribution contains correlated data, a block bootstrap algorithm can be used. The block bootstrap algorithm differs only by grouping the observed data into consecutive segments, such that the length of each segment is at least as long as the autocorrelation time of the data. The bootstrap distributions are then generated by randomly selecting blocks from the observed distribution. In general a block bootstrap error estimate will be greater than that from the standard algorithm and the ratio of the two estimates can be a used as a rough estimate of the autocorrelation time.[10]

The second source of error in free energy estimates arises from insufficient sampling, such that the observed distribution is not reflective of the true distribution.[208] To estimate the magnitude of this error, one can repeat the simulations with different initial conditions and calculate the standard error across independent trials. This strategy has been called the "ensemble average approach".[209–212] It has been suggested that the length of each simulation should be at least 50 times the length of the autocorrelation time of the data. Unfortunately, autocorrelation times are largely system-dependent and so this is difficult to verify in practice. For example, it was found that $\partial U/\partial\lambda$ had an autocorrelation time of up to 3 ns in charge-changing $p\mathrm{K}_a$ simulations of base pairs,[52] whereas autocorrelation times between 1 and 2 ps in solvation free energy bookending simulations.[60] Finally, it is worth noting that Hamiltonian replica exchange between $\lambda$-simulations has been shown to reduce autocorrelation times and improve the reproducibility of free energy estimates between independent trials.[52]

Perhaps the simplest (although not always sufficient) way to create an ensemble of simulations is to repeat a calculation with the same initial structure but different random

seed. This behavior has long been the default in AMBER when repeating a simulation with a stochastic integration scheme. If one or more simulations initiated from the same configuration but using different seeds do not return similar answers, then it is likely that the simulations are of insufficient length. As always, one should be cautious of the false negative rate of this approach and collect as many repeats of appropriate length as can be afforded in order to obtain the necessary accuracy. For example, in the limit of many infinitesimally short simulations this approach would (very probably incorrectly) suggest that no convergence problems exist.

**Structural Analysis—**BFE simulations are only accurate if the calculations sample important states and transitions. If one could sample infinitely, then it would be relatively simple to calculate a binding coefficient by directly counting transitions from a bound state to an unbound state. The difficulty is that the transitions themselves are quite rare and therefore require exceedingly long simulations to observe sufficient transitions to accurately calculate the binding coefficient. Indeed, the real timescale of binding events is generally of the order of microseconds or longer. This is currently inaccessible on commodity computing architectures and can only be accomplished with modest throughput on purpose-built or leadership computing platforms.[213,214] Fortunately, the thermodynamics of binding can be recapitulated by exclusively sampling the alchemical (as opposed to conformational) transition between the states. The problem then becomes a much more tractable issue of correctly characterizing these two specific states.[215]

It is important to ensure, as a basic test, that the protein and ligand are bound for the duration of the complex phase of simulation. This can be done by computing the root-mean-square deviation (RMSD) or center of mass (COM) motion of a ligand in a pocket relative to that of the protein. That is, these quantities should be computed once the translational and rotational motion of the protein have been minimized via rigid transformations of all coordinates. If the RMSD or COM motion of the ligand deviates significantly then the ligand has likely fallen out of the pocket.

Another test that can be done to ensure the consistency of a binding calculation is to examine torsional profiles of the ligand and perhaps even proximal protein sidechains. As a general rule, one expects more flexibility from a ligand in solution. As such, comparisons between bound and unbound simulations can easily expose obvious under-sampling when the torsional populations appear to be completely uncorrelated. Of course, even when sampling is sufficient the overlap need not be exactly identical because the protein binding pocket will impact the relevant configurations in the bound state relative to the unbound state. If there is reasonable confidence in the sampling, then differences between the torsional profiles can also be used to identify specific interactions in the protein.

Additionally, it is important to check dihedral profiles between independent runs, since rare transitions involving hidden high energy barriers may need to be captured to get a truly accurate binding free energy calculation. This can simply be done by extending your simulation time or, in difficult cases, applying enhanced sampling techniques. A number of enhanced sampling methods are available in AMBER20, such as Gaussian Accelerated MD (GaMD),[216] the replica exchange version of GaMD (rex-GaMD),,[217] and the recently

introduced Ligand Gaussian Accelerated Molecular Dynamics (LiGaMD).[218] Other enhanced sampling methods, such as replica exchange with solute tempering,[219,220] will be available in future versions of AMBER. Other methods for enhanced sampling, such as umbrella sampling,[221,222] metadynamics[223,224] and adaptive biasing force,[109,225] could be employed, although such methods rely on the definition a collective variable (CV) prior to simulating the system.

**Multiple Poses—**Similar to using multiple random seeds, it is also possible to use variations of the input poses. In cases where there is an unambiguous pose for each ligand, small variations in the poses can still provide insights regarding the local convergence of the BFE simulations. Multiple poses can be generated many ways, such as saving multiple poses from the docking program, using different docking programs, subjecting a pose to different minimization routines, or minimizing with different force fields. Such deviations act in a similar way to random seeds, although due to the slight variation in initial coordinates, simulations can sample significantly different portions of phase space.

In cases where the poses are ambiguous, it is necessary to run BFE simulations in each of the viable poses and combine the results. Multiple possible ligand poses may be encountered throughout a drug design project, especially during hit finding or early lead optimization when detailed and/or trustworthy structural data may not be available via experimental or hybrid means (*e.g.* homology modeling). Since the docking scores are often only weakly correlated with the true binding affinities,[226] a subsequent binding free energy assessment can provide valuable information. After an initial test that the binding poses are in fact stable (see Section 5.9), free energy simulations can be launched from each of the viable poses.

A few scenarios are possible when running simulations from multiple poses:

1. the poses interconvert and yield the same free energy,
2. the poses interconvert and do not yield the same freeenergy,
3. the poses do not interconvert and do not yield thesame free energy.
4. the poses do not interconvert but still yield the samefree energy.

If interconversion does occur, then the presence of multiple poses could imply complex dynamics that may offer a useful guide for assessing sampling in other ligands. If one gets the same free energies from different poses (scenario #1), then sufficient sampling can generally be assumed. If different free energy predictions are produced in the separate interconverting simulations (scenario #2), then likely a hysteresis issue has been uncovered, suggesting insufficient sampling.

If interconversion does not occur then it is possible that a particular pose is not physically relevant or that the energetic barrier between poses is too large to overcome within the simulation timescale. Poses like this can be identified based on having significantly different binding affinity predictions between them (scenario #3). In this case the more favorable free energy state is most probably the correct (or most relevant) state (assuming that the pose for the reference ligand is correct, in the case of RBFE simulations). Finally, if the poses do not interconvert but the free energy estimate is the same or similar (scenario #4), then a

correction should be made to account for the multiple states. For poses giving rise to separated states one can use a simple discrete model (more sophisticated alternatives have also been reported[126,128]). For $N$ different poses with binding free energies $\Delta G_i$ ($i = 1,\ldots,N$), the corrected binding free energy $\Delta G_{corr}$ across all states is:

$$\Delta G_{corr} = \sum_{i=1}^{N} p_i \Delta G_i + \beta^{-1} \sum_{i=1}^{N} p_i \ln p_i \,. \tag{18}$$

where $\beta$ is $1/kT$ and the probabilities $p_i$ are the normalized Boltzmann weights for the $i$th pose:

$$p_i \equiv \frac{e^{-\beta \Delta G_i}}{\sum_{i=1}^{N} e^{-\beta \Delta G_i}} \,. \tag{19}$$

The first term is just a weighted average of the binding free energies for each state (pose) while the second term is the entropic contribution. In the special case that all the $\Delta G_i$ are the same, one obtains $p_i = 1/N$ and the second term is simply $\beta^{-1} \ln N$. In the case that one pose is considerably more favorable than the others, it will dominate the first term but *not* the second. A more detailed discussion, including expanded variations to the expression for $\Delta G_{corr}$, is provided elsewhere.[126,128]

In the case where interconversion between poses takes place in a subset of the $\lambda$ windows, special care should be taken to understand the nature of the poses and the relationship between them. While differential interconversion is expected in different $\lambda$ windows due to differences in the Hamiltonian, it may also indicate insufficient sampling. In such cases, employing a Hamiltonian replica exchange approach may improve results (or indicate that the differential interconversion is not an issue).[104] Alternatively, methods can be employed to directly sample the transition between the states using an alchemical[227] or Monte Carlo approach.[228]

**Reversibility and Hysteresis—**A related concept to using multiple poses arises in RBFE calculations when a reference compound (e.g. one observed in a crystallographic structure) is used to dock a candidate compound but multiple orientations are returned. In theory, a free energy simulation should yield the same result no matter which compound is alchemically morphed into which – any discrepancy likely implies a sampling issue. The two may also differ if the candidate is suitably different in character from the reference as to take on a completely different pose and/or induce a conformational change in (part of) the protein. These two scenarios can be difficult to tell apart when the ligand perturbation is spatially quite large.

**Force Field Variations—**Unfortunately, sampling issues, whatever the cause, are not the only concerns when assessing confidence and error estimation – the details of the force field can pre-ordain a calculation to yield an inaccurate outcome, no matter how much care is placed on the other steps (garbage in, garbage out). Ideally, one would be able to probe specific characteristics of a model. What is the extent of charge “prepolarization”? Which torsional populations are dominant? How close are specific van der Waals contacts?

Unfortunately these are rarely clear cut knobs that the user can dial up and down, but it may be possible to correlate them roughly with a family of force fields. While it may not be possible in all cases, it can be informative to repeat a simulation with a different force field model of the ligand and/or protein (our experience is that solvent models display less sensitivity, at least on short time scales). A useful practice is to reserve this strategy for *extreme* predictions (e.g. ligand modifications $>2$ kcal/mol more or less favorable than the reference). If such outcomes are reproducible, the result may be due to a fundamental bias in the force field. For example, excessive polarization may lead to over-stabilized hydrogen or halide bonds or an unusual torsional profile might favor an implausible conformation. Rough consensus across multiple force field models can be a source of higher confidence or else an indication that a trend is qualitatively, but not quantitatively correct.

There are a number of force fields available in AMBER for proteins, nucleic acids, carbohydrates, lipids, solvents, and ions (for details about recommended available force fields, see the AMBER20 manual http://ambermd.org/doc12/Amber20.pdf). Additionally, variants of the general AMBER force field (GAFF[133] and GAFF2[57]) are available for non-standard residues including drug-like molecules and modified amino acids. For protein systems, the Stony Brook (SB) family of protein force fields (ff19SB,[229] ff14SB,[230] and ff99SB[231]) are the most commonly used in AMBER. For organic drug-like molecules, GAFF2[57] is the latest version of the generalized AMBER force field. Additional details can be found in the AMBER20 manual.

Here, we specifically suggest that lack of consensus in BFE calculations among different force fields is a red flag and should be investigated further. In some cases a particular force field may be significantly better than another for the ligands of interest, in which case differences in results would be expected (and results from the better force field should be more trusted). However, when results vary and it is not clear which force field is better, it is not obvious to us that there is a general course of action to take other than to increase scrutiny of the results. Recently, Gapsys, *et al.*[34] proposed a consensus method using multiple force fields that improved results in certain situations. This may be a profitable avenue for future research.

## 6 Future Work

There are a number of areas that will be the focus of new free energy developments for drug discovery in AMBER driven by academic-industry partnership. These include the development of new force fields (QM, MM, and machine learning), enhanced sampling methods (both in the $\lambda$ dimension as well as conformational degrees of freedom), improved alchemical transformation pathways, and optimization of RBFE networks (including integration of experimental constraints). These are briefly summarized below.

### 6.1 MM→QM Book-ending Approaches

The robust prediction afforded by alchemical free energy methods in drug discovery presents considerable challenges for conventional molecular mechanical (MM) force fields.[232] This is due, in part, to the need to test chemically diverse molecules for which tested parameters may not exist.[3] Modeling certain types of electrostatic interactions, such as sigma holes and

cation-pi interactions are especially challenging for point charge and multipole models. Further, the process of drug binding involves a considerable change in the molecular electrostatic environment that requires explicit consideration of electronic response for high accuracy. Finally, the modeling of complex interactions of metal ions and formation/cleavage of chemical bonds for covalent inhibitors demands a more sophisticated quantum electronic structure treatment. Quantum mechanical (QM) methods, if made sufficiently fast to be computationally tractable, offer a potentially transformative solution to these problems.

In this regard, one could argue that practically every alchemical free energy prediction used in drug discovery could potentially benefit from accurate QM methods. In addition to the general cases described above, particularly prominent examples that demand QM methods include systems such as metalloproteins where drugs target inner-sphere coordination to the metal centers, and in general highly charged systems (including RNA targets) or systems that involve charge-changing transformations (including protonation/deprotonation events) that exacerbate the need for many-body polarization and charge transfer effects. As these QM methods are still in early stages with respect to their application to drug discovery in alchemical free energy simulations, it remains to be seen the degree to which they may have impact and over what range of targets and drugs.

Free energy simulations with combined quantum mechanical/molecular mechanical (QM/MM) potentials or fully quantum mechanical force fields (QMFFs), if made sufficiently fast to be computationally tractable, offer a potentially transformative solution to these problems.[233–238] Quantum models, if made affordable, are thus highly attractive for drug design applications owing to their accuracy, robustness and lack of adjustable free parameters relative to MM force fields.[239]

A common strategy to efficiently correct is to perform the alchemical transformation with a MM method, and then apply MM→QM/MM free energy corrections to the end-states. This is referred to as a “book-ending”, “indirect” or “reference potential” approach[60,240–252] The primary goal of these methods is to indirectly estimate a free energy difference between states $A$ and $B$, $\Delta G_{A\rightarrow B}$, using a computationally demanding Hamiltonian by evaluating the free energy change at a low-level of theory and then correcting for the free energy difference associated with changing the Hamiltonian.

Book-ending method development was motivated by the idea of using QM and QM/MM Hamiltonians to improve the accuracy of solvation and RBFE predictions.[239] Book-ending methods circumvent a number of obstacles. First, the alchemical transformation step often involves simulation of several nonphysical intermediate states connecting the end-states, and it may not be obvious how the nonphysical states might be modeled with a QM Hamiltonian. For example, in the case where atoms are deleted or inserted into the system, one might have to contend with the idea of having a noninteger number of electrons and partial nuclear charges. Furthermore, alchemical thermodynamic pathways often require a number of intermediate states, each of which requiring a significant amount of sampling to converge the free energy result, which would be prohibitively expensive with a QM or QM/MM Hamiltonian. A bookending method instead computes the alchemical transformation with a molecular mechanical Hamiltonian method, where softcore potentials have been developed

and which are inexpensive enough to be sampled. The estimation of the MM→QM/MM free energy changes does not require simulation of alchemical systems; one simulates the two end-states (*A* and *B*) using one-or-more Hamiltonians that connect the low- and high-level Hamiltonians. It is important to choose the most compatible reference (MM) potential for the particular high-level Hamiltonian to avoid slow convergence of the Hamiltonian free energy correction estimate.[253–256] This has led to work that sought to increase the distribution overlap between the reference and high-level Hamiltonians,[244,245,254,257–261] including methods that perform *ad hoc* parameterization of the MM reference potential via "force matching" to the QM/MM potential.[245,255,260–267]

It is worthwhile to note that other methods have been explored to reduce the number of energy and force evaluations necessary to converge QM or QM/MM free energy estimates. These include trajectory reweighting,[251,268–271] the use of frozen density functional approximations,[272,273] integrated Hamiltonian sampling,[274] orthogonal space random walk strategies,[275] and paradynamics.[258,276] Collectively, these methods offer considerable promise to greatly improve the accuracy and predictive capability of alchemical free energy simulations with practical computational resources.

### 6.2 Further Softcore Improvements

**Modify the exponents in the softcore potential—**We are exploring different forms of the softcore potentials in order to obtain more numerically stable and smooth results in TI, BAR and MBAR calculations. The general forms of "effective diatomic distances" (Eq.(15)) are:

$$r_{ij}^{\mathrm{LJ}}\left(\lambda;\alpha^{LJ}\right)=\left[r_{ij}^{n}+\alpha^{\mathrm{LJ}}W\left(r_{ij}\right)S_2(\lambda)\sigma_{ij}^{n}\right]^{1/n}$$
$$r_{ij}^{\mathrm{Coul}}\left(\lambda;\alpha^{\mathrm{Coul}}\right)=\left[r_{ij}^{m}+\alpha^{\mathrm{Coul}}W\left(r_{ij}\right)S_2(\lambda)\sigma_{ij}^{m}\right]^{1/m} \quad (20)$$

For the Lennard-Jones and the Coulombic interactions, respectively. In the current default AMBER softcore form, $m$ is 2 and $n$ is 6. As mentioned earlier, now both $r_{ij}^{\mathrm{LJ}}\left(\lambda;\alpha^{\mathrm{LJ}}\right)$ and $r_{ij}^{\mathrm{Coul}}\left(\lambda;\alpha^{\mathrm{Coul}}\right)$ are in exactly the same form and $\alpha^{\mathrm{LJ}}$ and $\alpha^{\mathrm{Coul}}$ are unitless. (Note that $\alpha^{\mathrm{LJ}}$ is the same as the original $\alpha$ in AMBER) Hence it is easier to explore and directly compare different parameters. One obvious observation is that, in the current form, a larger $m$ or $n$ will imply a shorter range of the softcore effect. The current setting ($m$=2, $n$=6) will have stronger short range softness in LJ than in Coul interactions, which could be one of the reasons that it is difficult to balance them by only modifying $\alpha$ and $\alpha^{\mathrm{Coul}}$.

We explored different combinations of ($m$, $n$), including ($m$=1, $n$=1) and ($m$=2, $n$=2), incorporated with the smoothstep function SSC(2) and the new softcore parameter $\alpha^{\mathrm{Coul}}$ and the preliminary results are shown in Figure 7 and Figure 8. Figure 7 compares the current default scheme ($m$=2, $n$=6) with ($m$=1, $n$=1) and ($m$=2, $n$=2) schemes with the AMBER18 default $\beta$ value (12 Å$^2$) for ($m$=2, $n$=6) and $\alpha^{\mathrm{LJ}}$ = 1 for ($m$=1, $n$=1) and ($m$=2, $n$=2), with various $\alpha$ (the same as $\alpha^{\mathrm{LJ}}$ in the new form) values. Figure 8 shows the same comparison except $\beta$ is the AMBER20 default $\beta$ value for SSC(2) scheme (50 Å$^2$) and $\alpha^{\mathrm{Coul}}$ = 4. The preliminary results shown in Figure 7 and Figure 8 suggest that by modifying the softcore exponents $m$ and $n$, the imbalance between the LJ and Coul interactions can be significantly

reduced. We are currently extending the verification to other more realistic molecular systems.

### 6.3 $\lambda$-scheduling with smoothstep functions

As mentioned earlier, AMBER20 now provides $\lambda$-scheduling for flexible turning on and off for individual interactions at different stages along the alchemical $\lambda$-axis, similar to other simulation packages such as GROMOS[92] and NAMD.[84,122] The commonly-used "stepwise" scheme is equivalent to "schedule" the LJ and Coul interactions at different stages (*e.g.*, in NAMD). Furthermore, the $\lambda$-scheduling can be applied to bonded terms, i.e., bond length, bond angle and torsion terms, so that the internal conformation of a disappearing softcore region can be kept until very late in the $\lambda$ transformation, which could prevent or reduce the conformation sampling problems when the internal bonded terms need to be scaled with $\lambda$. We are currently exploring different $\lambda$-scheduling for different interactions to identify the best scheduling that will deliver the most smooth and stable $\left\langle \frac{dU}{d\lambda} \right\rangle$ curves.

### 6.4 Enhanced sampling

In the past few decades methods have been developed that address the sampling problem, such as replica-exchange molecular dynamics, metadynamics, simulated annealing, and orthonomal space methods. Major focus has been on enhanced sampling in the conformational spaces. Nevertheless, with the rapid developing advances in hardware and software, the MD-based free energy methods become feasible and hence emerges the importance of enhanced sampling in the alchemical space. On one hand, enhanced sampling methods in the alchemical space concern similarly as the counter methods in the conformational space. They both need to have adequate space coverage in order to obtain proper statistically meaningful ensembles and at the same time need to reduce the time spent on the spaces that are not critical for the desired properties. On the other hand, a fundamental difference is that the free energy is a state function hence theoretically it is possible to reduce alchemical sampling through efficient and theoretically robust methods.

With the advanced methods recently implemented in AMBER20 reported here, we are in a much better position to explore various advanced enhanced sampling methods in the alchemical space. For example, the developed SSC(2) scheme is well suited for advanced $\lambda$-scheduling optimization and enhanced sampling schemes in the alchemical space where a single-pass concerted $\lambda$ transformation is desirable, including $\lambda$ dynamics,[99–102] Hamiltonian replica exchange methods,[91,103–107] adaptive biasing,[100,108,109] and self-adjusted mixture sampling[110,111] methods. For the conformational enhanced sampling at a given $\lambda$, the REST/REST2 methods[277,278] have been shown to be very successful. We are actively investigating possible incorporation of the SSC(2) potential with these techniques.

Another approach is the use of Gaussian accelerated molecular dynamics (GaMD) and its more recent derivatives, particularly LiGaMD, as an enhanced sampling methodology.[216,218] This methodology allows for a Gaussian boost potential to be selectively applied to either the bonded terms, nonbonded terms, and/or the potential energy. These can be applied to the entire system or selectively to the ligand and its contacts. This amount of control allows for faster sampling of ligand binding states and to overcome high energy barriers. As the

methodology currently does not support $\lambda$ scaling simulations, we are actively investigating incorporating these methodologies for TI and MBAR.

### 6.5 Force Field

Substantial progress has been made in the parameterization of empirical molecular force fields, as manifest in the consistent improvements in the agreement between the predictions of molecular dynamics simulations and the experimental measurements,[279] including in the binding free energies.[116,140] The remaining journey to experimental-level accuracy, however, is almost surely no shorter and no less arduous than the road that has led us to where we are now. Although further improvements in predictive accuracy are expected from fine-tuning the parameters for the current functional forms in the force fields, more substantive modifications may be required to reach experimental level accuracy.

It is now widely recognized that off-atom-center partial charges are necessary to accurately capture the electrostatic potential around a molecule.[140,280–282] These off-atom-center charges, referred to as virtual sites, are placed at a predetermined position within a molecular frame defined by a parent atom and up to three neighboring atoms that are covalently bonded to the parent atom. The current AMBER code only supports a limited number of ways for placing the virtual sites, and refer to them as extra points (EP). Generalized virtual sites, however, have been implemented in a developmental branch of AMBER. Some new types of virtual sites are shown in Figure 9 and will likely become available in the next official release. In order to make the virtual sites truly useful, however, substantial work is required to optimize their locations and the methods to parameterize their charge values.

The short-range repulsive interactions between atoms have commonly been modeled by a $1/r^{12}$ potential, which was originally proposed for computational efficiency rather than for physical accuracy. Buckingham suggested that the repulsive potential due to the Pauli exclusion principle should resemble an exponential, and he proposed the following potential function as a substitute for the prevalent Lennard-Jones 12–6 potential:

$$V(r) = Ae^{-Br} - \frac{C}{r^6} \tag{21}$$

The Buckingham potential, first published in an article authored by Buckingham and communicated by no other than Lennard-Jones himself to the Proceedings of the Royal Society of London,[283–285] has found widespread use in material science simulations. But its adoption in biomolecular simulations is so far limited. It will be interesting to see whether its more realistic description of the repulsive potential can lead to more accurate binding free energy results.

In protein-ligand binding, the ligand molecule transfers from the high dielectric environment of water to a different dielectric environment of the protein binding pocket. This typically induces a redistribution of the electrons in the ligand due to molecular polarizability. Thus, at least in theory, a polarizable force field should improve the accuracy of binding free energy calculations.[286] Polarizability can be modeled by either inducible dipoles[287] or Drude oscillators.[288] So far, however, polarizable force field models have not been widely

used in binding free energy calculations, due to both the roughly 10- to 20-fold increase in the computational cost and the difficulty in parameterizing the models.[289] Nevertheless, it will be worthwhile to keep an eye out for the development of polarizable models,[290] and to implement them in AMBER when the time comes.

### 6.6 Network RBFE

In a typical drug discovery project, BFE is routinely used to compute the binding free energies of a large number–usually 10s to 100s in each batch–of candidate molecules against the same protein target of interest. There are many ways to collectively compute these binding free energies. Take a simple example of 3 molecules: A, B, and C. One can compute the individual binding free energy for A ($\Delta G_A$) and the relative binding free energies between A and B ($\Delta\Delta G_{AB}$) and between A and C ($\Delta\Delta G_{AC}$), and then estimate the binding free energies for B and C by $\Delta G_B = \Delta G_A + \Delta\Delta G_{AB}$ and $\Delta G_C = \Delta G_A + \Delta\Delta G_{AC}$. Alternatively, one can compute the individual binding free energies $\Delta G_A$ and $\Delta G_B$, as well as the relative binding free energies $\Delta\Delta G_{BC}$ and $\Delta\Delta G_{AC}$, and estimate the binding free energies for C by $\Delta G_C = (\Delta G_A + \Delta\Delta G_{AC} + \Delta G_B + \Delta\Delta G_{BC})/2$. Moreover, some of the calculations can be allocated more simulation time than others, resulting in different statistical errors in different calculations. Given a fixed computational cost, there are infinite numbers of ways to allocate them to the calculations of different individual and relative binding free energy calculations. This poses an interesting problem of experimental design: how to best allocate the computational resources to the calculations so as to minimize the overall statistical error in the estimated binding free energies?

A number of approaches have been developed to address this problem and to plan the binding free energy calculations for a set of many (between 10 and 100) compounds against the same target. The earlier approaches such as LOMAP[195] aim to construct a network of relative binding free energy calculations so that any pair of molecules can be computed by combining the results of a small number of pairs of structurally similar molecules and the binding free energy difference between any pair can be computed by at least two different combinations above. Subsequent works introduced rigorous mathematical frameworks to minimize the overall statistical error with respect to either the selection of computed pairs[291] or the allocation of computational resources to each pair.[292] Using these optimal allocations, the same level of statistical precision may be achieved at half of the computational cost compared to commonly used *ad hoc* allocations. Such optimized network of binding free energy calculations may help expand the number of molecules that can be characterized by binding free energy calculations in each round of drug design.

### 6.7 Network-wide Free Energy Analysis

In the future, methods will be explored that can incorporate known experimental binding energies into RBFE networks to improve the quality of the remaining free energy estimates. One idea for achieving this is to incorporate constraints into the RBFE calculations. Recently, the solution of the MBAR equations has been re-expressed in terms of a nonlinear minimization of an objective function, rather than having to solve a self-referential set of equations.[293] This re-expression of the MBAR solution allows one to use nonlinear parameter optimization software to minimize the objective function, where the parameters of

the optimization are the simulation free energies (to within a constant). When a network of RBFEs are considered, one can expand the optimization method by constructing a new objective function that is a sum of objectives corresponding to each edge of the network. Optimization of the objective function sum is equivalent to the independent optimization of each objective *unless* constraints are introduced that couple the free energies between edges. An obvious set of constraints are *cycle closures* that force the sum of free energies along a closed path to be zero. Furthermore, if a partial list of known RBFEs are experimentally known, then those free energies can be included as affine constraints.

Giese and York have performed a preliminary, proof-of-concept implementation of the constrained network MBAR approach described above, and an example of its use is shown in Figure 10 which displays RBFEs of ligands to the CDK2 protein. There are 16 ligands, and the ligand with the highest experimental binding free energy was chosen to define the zero of free energy. There are a total 22 cycle closure constraints (11 constraints for the solution-phase network and 11 constraints for the CDK2-bound network). The image shows that enforcing cycle closure constraints within the MBAR optimization has little effect on the RBFEs unless an additional experimental RBFE is included. In this example, the calculated RBFE of “ligand 28” was constrained to match experiment. It is noteworthy that the inclusion of the experimental RBFE improves the prediction of most ligands because the calculated RBFEs are highly coupled by cycle closures. If the cycle closures were not included as constraints, then the experimental constraint on “ligand 28” would only change the result of “ligand 28”. Improvement to the correlation coefficient (R) would still be made if a different ligand’s RBFE was constrained to match experiment. The average correlation coefficient of the 15 possible constraints (the “ligand 28” constraint illustrated in Figure 10 is only one of the 15 cases) is 0.84 ± 0.02, which is a significant improvement relative to the 0.69 correlation coefficient obtained when cycle-closure and experimental constraints are not enforced. Furthermore, the average mean unsigned error is reduced from 1.0 kcal/mol to 0.6 ± 0.1 kcal/mol, and the average mean signed error is improved from 0.9 kcal/mol to 0.3 ± 0.2 kcal/mol.

## 7 Conclusion

In this work we describe new features in AMBER20 for performing GPU-accelerated alchemical binding free energy simulations. We focus on features and functionality related to drug discovery effort that arose from an ongoing collaboration between the York Group, the Laboratory for Biomolecular Simulation Research at Rutgers University, and Silicon Therapeutics. We also describe the ancillary tools outside of AMBER needed for preparing and analyzing alchemical binding free energy simulations. We have attempted to note the nuances associated with free energy simulations in the context of drug discovery, especially regarding the balance between ease-of-use and expert control. While automated protocols can work, especially for highly validated protein targets, each target presents unique challenges with respect to free energy landscapes and intrinsic timescales for sampling relevant protein and solvent motions. As such, currently the best results can be obtained by experienced users with fine-tuned control of the software packages that they use. Indeed, while alchemical free energy simulations offer great value even with current automated packages, it is important to note that significant challenges still exist related to obtaining

accurate and robust binding free energy predictions for drug discovery applications. We hope that this work has illustrated some of the critical issues that should be considered and will help educate the broader population of researchers engaged in the use of binding free energy simulations in drug discovery and related applications to emerging areas such as precision medicine.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors are grateful for financial support provided by the National Institutes of Health (No. GM107485 to DMY). Computational resources were provided by the Office of Advanced Research Computing (OARC) at Rutgers, The State University of New Jersey (specifically, the Amarel cluster and associated research computing resources), the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number No. ACI-1548562[294] (specifically, the resources COMET and COMET GPU at SDSC through allocation TG-CHE190067), and the Texas Advanced Computing Center (TACC) at the University of Texas at Austin (specifically, the Frontera Supercomputer through allocation CHE20002). We also gratefully acknowledge the support of the NVIDIA Corporation with the donation of several Pascal, Volta, and Turing GPUs for testing.

## References

(1). Jorgensen WL Efficient Drug Lead Discovery and Optimization. Acc. Chem. Res 2009, 42, 724–733. [PubMed: 19317443]

(2). Abel R; Wang L; Harder ED; Berne BJ; Friesner RA Advancing Drug Discovery through Enhanced Free Energy Calculations. Acc. Chem. Res 2017, 50, 1625–1632. [PubMed: 28677954]

(3). Cournia Z; Allen B; Sherman W Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. J. Chem. Inf. Model 2017, 57, 2911–2937. [PubMed: 29243483]

(4). Cournia Z; Allen BK; Beuming T; Pearlman DA; Radak BK; Sherman W Role of Rigorous Free Energy Simulations in Virtual Screening. J. Chem. Inf. Model 2020, DOI: 10.1021/acs.jcim.0c00116.

(5). Zwanzig RW High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. J. Chem. Phys 1954, 22, 1420–1426.

(6). Kirkwood JG Statistical Mechanics of Fluid Mixtures. J. Chem. Phys 1935, 3, 300–313.

(7). Bennett CH Efficient Estimation of Free Energy Differences from Monte Carlo Data. J. Comput. Phys 1976, 22, 245–268.

(8). Souaille M; Roux B Extension to the Weighted Histogram Analysis Method: Combining Umbrella Sampling with Free Energy Calculations. Comput. Phys. Commun 2001, 135, 40–57.

(9). Shirts MR; Chodera JD Statistically Optimal Analysis of Samples from Multiple Equilibrium States. J. Chem. Phys 2008, 129, 124105. [PubMed: 19045004]

(10). Tan Z; Gallicchio E; Lapelosa M; Levy RM Theory of Binless Multi-state Free Energy Estimation with Applications to Protein-Ligand Binding. J. Chem. Phys 2012, 136, 144102. [PubMed: 22502496]

(11). Shirts MR Reweighting from the Mixture Distribution as a Better Way to Describe the Multistate Bennett Acceptance Ratio. arXiv preprint arXiv:1704.00891 2017.

(12). Kirkwood JG, Theory of Liquids; Gordon and Breach: 1968; Vol. 2.

(13). Postma JP; Berendsen HJ; Haak JR Thermodynamics of Cavity Formation in Water. A Molecular Dynamics Study. Faraday Symposia of the Chemical Society 1982, 17, 55–67.

(14). Jorgensen WL; Ravimohan C Monte Carlo Simulation of Differences in Free Energies of Hydration. J. Chem. Phys 1985, 83, 3050–3054.

(15). Tembe BL; Mccammon JA Ligand-Receptor Interactions. Comp. Chem 1984, 8, 281–283.

(16). Lybrand TP; McCammon JA; Wipff G Theoretical Calculation of Relative Binding Affinity in Host-Guest Systems. Proc. Natl. Acad. Sci. USA 1986, 83, 833–835. [PubMed: 3456569]

(17). Straatsma T; Berendsen H; Postma J Free Energy of Hydrophobic Hydration: A Molecular Dynamics Study of Noble Gases in Water. J. Chem. Phys 1986, 85, 6720–6727.

(18). Merz KM Jr.; Kollman PA Free Energy Perturbation Simulations of the Inhibition of Thermolysin: Prediction of the Free Energy of Binding of A New Inhibitor. J. Am. Chem. Soc 1989, 111, 5649–5658.

(19). Pearlman DA; Connelly PR Determination of the Differential Effects of Hydrogen Bonding and Water Release on the Binding of FK506 to Native and Tyr82 → Phe82 FKBP-12 Proteins Using Free Energy Simulations. J. Mol. Biol 1995, 248, 696–717. [PubMed: 7538591]

(20). Helms V; Wade RC Computational Alchemy to Calculate Absolute Protein-Ligand Binding Free Energy. J. Am. Chem. Soc 1998, 120, 2710–2713.

(21). Mobley DL; Klimovich PV Perspective: Alchemical Free Energy Calculations for Drug Discovery. J. Chem. Phys 2012, 137, 230901. [PubMed: 23267463]

(22). Chodera JD; Mobley DL; Shirts MR; Dixon RW; Branson K; Pande VS Alchemical Free Energy Methods for Drug Discovery: Progress and Challenges. Curr. Opin. Struc. Biol 2011, 21, 150–160.

(23). Shirts MR; Mobley DL; Chodera JD Alchemical Free Energy Calculations: Ready for Prime Time?Annual Reports in Computational Chemistry 2007, 3, 41–59.

(24). Deng Y; Roux B Computations of Standard Binding Free Energies with Molecular Dynamics Simulations. J. Phys. Chem. B 2009, 113, 2234–2246. [PubMed: 19146384]

(25). Gilson MK; Zhou H-X Calculation of Protein-Ligand Binding Affinities. Annu. Rev. Biophys. Biomol. Struct 2007, 36, 21–42. [PubMed: 17201676]

(26). Gilson MK; Given JA; Bush BL; McCammon JA The Statistical-Thermodynamic Basis for Computation of Binding Affinities: A Critical Review. Biophys. J 1997, 72, 1047–1069. [PubMed: 9138555]

(27). Woo H-J; Roux B Calculation of Absolute Protein-Ligand Binding Free Energy from Computer Simulations. Proc. Natl. Acad. Sci. USA 2005, 102, 6825–6830. [PubMed: 15867154]

(28). Singh U; Weiner SJ; Kollman P Molecular Dynamics Simulations of d(CGCGA) X d(TCGCG) With and Without ”Hydrated” Counterions. Proc. Natl. Acad. Sci. USA 1985, 82, 755–759. [PubMed: 3856228]

(29). Bash P; Singh U; Langridge R; Kollman P Free Energy Calculations by Computer Simulation. Science 1987, 236, 564–568. [PubMed: 3576184]

(30). Singh U; Brown FK; Bash PA; Kollman PA An Approach to The Application of Free Energy Perturbation Methods Using Molecular Dynamics: Applications to the Transformations of Methanol→Ethane, Oxonium→Ammonium, Glycine→Alanine, and Alanine→Phenylalanine in Aqueous Solution and to H3O+ (H2O) 3→NH4+ (H2O) 3 in the Gas Phase. J. Am. Chem. Soc 1987, 109, 1607–1614.

(31). Bash PA; Singh UC; Brown FK; Langridge R; Kollman PA Calculation of the Relative Change in Binding Free Energy of A Protein-Inhibitor Complex. Science 1987, 235, 574–576. [PubMed: 3810157]

(32). Pearlman DA; Kollman PA The Lag Between the Hamiltonian and The system Configuration in Free Energy Perturbation Calculations. J. Chem. Phys 1989, 91, 7831–7839.

(33). Jarzynski C Nonequilibrium Equality for Free Energy Differences. Phys. Rev. Lett 1997, 78, 2690–2693.

(34). Gapsys V; Pérez-Benito L; Aldeghi M; Seeliger D; van Vlijmen H; Tresadern G; de Groot BL Large Scale Relative Protein Ligand Binding Affinities Using Non-Equilibrium Alchemy. Chem. Sci 2020, 11, 1140–1152.

(35). Pearlman DA A Comparison of Alternative Approaches to Free Energy Calculations. J. Phys. Chem 1994, 98, 1487–1493.

(36). Pearlman DA; Case DA; Caldwell JW; Ross WS; Cheatham III TE; DeBolt S; Ferguson D; Seibel G; Kollman P AMBER, A Package of Computer Programs for Applying Molecular Mechanics, Normal Mode Analysis, Molecular Dynamics and Free Energy Calculations to

Simulate The Structural and Energetic Properties of Molecules. Comput. Phys. Commun 1995, 91, 1–41.

(37). Brooks BR; Bruccoleri RE; Olafson BD; States DJ; Swaminathan S. a.; Karplus M CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. J. Comp. Chem 1983, 4, 187–217.

(38). Fujinaga M; Gros P; Van Gunsteren W Testing the Method of Crystallographic Refinement using Molecular Dynamics. J. Appl. Crystallogr 1989, 22, 1–8.

(39). Pearlman DA; Kollman PA The Overlooked Bond-Stretching Contribution in Free Energy Perturbation Calculations. J. Chem. Phys 1991, 94, 4532–4545.

(40). Sun Y; Spellmeyer D; Pearlman DA; Kollman P Simulation of the Solvation Free Energies for Methane, Ethane, and Propane and Corresponding Amino Acid Dipeptides: A Critical Test of the Bond-PMF Correction, A New Set of Hydrocarbon Parameters, and the Gas Phase-Water Hydrophobicity Scale. J. Am. Chem. Soc 1992, 114, 6798–6801.

(41). Pearlman DA; Kollman PA A New Method for Carrying out Free Energy Perturbation Calculations: Dynamically Modified Windows. J. Chem. Phys 1989, 90, 2460–2470.

(42). Ferguson DM; Pearlman DA; Swope WC; Kollman PA Free Energy Perturbation Calculations Involving Potential Function Changes. J. Comp. Chem 1992, 13, 362–370.

(43). Pearlman DA; Rao BG In; Wiley Online Library: 2002; Chapter Free Energy Calculations: Methods and Applications

(44). Darden T; York D; Pedersen L Particle Mesh Ewald: An N log (N) Method for Ewald Sums in Large Systems. J. Chem. Phys 1993, 98, 10089–10092.

(45). Essmann U; Perera L; Berkowitz ML; Darden T; Hsing L; Pedersen LG A Smooth Particle Mesh Ewald Method. J. Chem. Phys 1995, 103, 8577–8593.

(46). York DM; Wlodawer A; Pedersen LG; Darden TA Atomic-Level Accuracy in Simulations of Large Protein Crystals. Proc. Natl. Acad. Sci. USA 1994, 91, 8715–8718. [PubMed: 7521533]

(47). York DM; Yang W; Lee H; Darden T; Pedersen LG Toward the Accurate Modeling of DNA: The Importance of Long-Range Electrostatics. J. Am. Chem. Soc 1995, 117, 5001–5002.

(48). Cheatham III TE; Miller JL; Fox T; Darden TA; Kollman PA Molecular Dynamics Simulations on Solvated Biomolecular Systems: The Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA, and Proteins. J. Am. Chem. Soc 1995, 117, 4193–4194.

(49). Götz A; Williamson MJ; Xu D; Poole D; Le Grand S; Walker RC Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. J. Chem. Theory Comput 2012, 8, 1542–1555. [PubMed: 22582031]

(50). Salomon-Ferrer R; Götz AW; Poole D; Le Grand S; Walker RC Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. J. Chem. Theory Comput 2013, 9, 3878–3888. [PubMed: 26592383]

(51). Le Grand S; Göetz AW; Walker RC SPFP: Speed without Compromise–A Mixed Precision Model for GPU Accelerated Molecular Dynamics Simulations. Comput. Phys. Commun 2013, 184, 374–380.

(52). Giese TJ; York DM A GPU-Accelerated Parameter Interpolation Thermodynamic Integration Free Energy Method. J. Chem. Theory Comput 2018, 14, 1564–1582. [PubMed: 29357243]

(53). Lee T-S; Hu Y; Sherborne B; Guo Z; York DM Toward Fast and Accurate Binding Affinity Prediction with pmemdGTI: An Efficient Implementation of GPU-Accelerated Thermodynamic Integration. J. Chem. Theory Comput 2017, 13, 3077–3084. [PubMed: 28618232]

(54). Lee T-S; Cerutti DS; Mermelstein D; Lin C; LeGrand S; Giese TJ; Roitberg A; Case DA; Walker RC; York DM GPU-Accelerated Molecular Dynamics and Free Energy Methods in Amber18: Performance Enhancements and New Features. J. Chem. Inf. Model 2018, 58, 2043–2050. [PubMed: 30199633]

(55). Tsai H-C; Tao Y; Lee T-S; Merz KM Jr.; York DM Validation of Free Energy Methods in AMBER. J. Chem. Inf. Model 2020, in press, 0–0.

(56). Song LF; Lee T-S; Zhu C; York DM; Merz KM Jr. Using AMBER18 for Relative Free Energy Calculations. J. Chem. Inf. Model 2019, 59, 3128–3135. [PubMed: 31244091]

(57). He X; Liu S; Lee T-S; Ji B; Man VH; York DM; Wang J Fast, Accurate, and Reliable Protocols for Routine Calculations of Protein-Ligand Binding Affinities in Drug Design Projects Using AMBER GPU-TI with ff14SB/GAFF. ACS Omega 2020, 5, 4611–4619. [PubMed: 32175507]

(58). Lee T-S; Lin Z; Allen BK; Lin C; Radak BK; Tao Y; Tsai H-C; Sherman W; York DM Improved Alchemical Free Energy Calculations with Optimized Smoothstep Softcore Potentials. J. Chem. Theory Comput 2020, In press.

(59). Boresch S; Tettinger F; Leitgeb M Absolute Binding Free Energies: A Quantitative Approach for Their Calculation. J. Phys. Chem. B 2003, 107, 9535–9551.

(60). Giese TJ; York DM Development of a Robust Indirect Approach for MM → QM Free Energy Calculations That Combines Force-Matched Reference Potential and Bennett's Acceptance Ratio Methods. J. Chem. Theory Comput 2019, 15, 5543–5562. [PubMed: 31507179]

(61). Humphrey W; Dalke A; Schulten K VMD – Visual Molecular Dynamics. J. Mol. Graphics 1996, 14, 33–38.

(62). Pettersen EF; Goddard TD; Huang CC; Couch GS; Greenblatt DM; Meng EC; Ferrin TE UCSF Chimera–A Visualization System for Exploratory Research and Analysis. J. Comput. Chem 2004, 25.

(63). Purawat S; Ieong PU; Malmstrom RD; Chan GJ; Yeung AK; Walker RC; Altintas I; Amaro RE A Kepler Workflow Tool for Reproducible AMBER GPU Molecular Dynamics. Biophys. J 2017, 112, 2469–2474. [PubMed: 28636905]

(64). Doerr S; Harvey MJ; Noé F; De Fabritiis G HTMD: High-Throughput Molecular Dynamics for Molecular Discovery. J. Chem. Theory Comput 2016, 12, 1845–1852. [PubMed: 26949976]

(65). Chemical Computing Group ULC Molecular Operating Environment(MOE),2019.01, 1010 Sherbrooke St. West, Suite 910, Montreal, QC, Canada, H3A 2R7, 2019.

(66). Homer RW; Swanson J; Jilek RJ; Hurst T; Clark RD SYBYL Line Notation (SLN): A Single Notation To Represent Chemical Structures, Queries, Reactions, and Virtual Libraries. J. Chem. Inf. Model 2008, 48, 2294–2307. [PubMed: 18998666]

(67). Release S Maestro. Schrödinger, LLC, New York, NY 2017, 2017.

(68). Scheen J; Wu W; Mey A; Tosco P; Mackey M; Michel J A Hybrid Alchemical Free Energy and Machine Learning Methodology for the Calculation of Absolute Hydration Free Energies of Small Molecules. ChemRxiv 2020, DOI: 10.26434/chemrxiv.12380612.v1.

(69). Straatsma TP; Berendsen HJ Free Energy of Ionic Hydration: Analysis of a Thermodynamic Integration Technique to Evaluate Free Energy Differences by Molecular Dynamics Simulations. J. Chem. Phys 1988, 89, 5876–5886.

(70). Crooks GE Path-Ensemble Averages in Systems Driven Far from Equilibrium. Phys. Rev. E 2000, 61, 2361–2366.

(71). Boresch S; Woodcock HL Convergence of Single-Step Free Energy Perturbation. Mol. Phys 2017, 115, 1200–1213.

(72). Gapsys V; Michielssens S; Peters JH; de Groot BL; Leonov H Calculation of Binding Free Energies. Methods Mol. Biol 2015, 1215, 173–209. [PubMed: 25330964]

(73). Jeong D; Andricioaei I Reconstructing Equilibrium Entropy and Enthalpy Profiles from Non-Equilibrium Pulling. J. Chem. Phys 2013, 138, 114110. [PubMed: 23534630]

(74). Wei D; Song Y; Wang F A Simple Molecular Mechanics Potential for μm Scale Graphene Simulations from the Adaptive Force Matching Method. J. Chem. Phys 2011, 134, 184704. [PubMed: 21568526]

(75). Torrie GM; Valleau JP Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. J. Comput. Phys 1977, 23, 187–199.

(76). Shirts MR; Pande VS Comparison of Efficiency and Bias of Free Energies Computed by Exponential Averaging, the Bennett Acceptance Ratio, and Thermodynamic Integration. J Chem Phys 2005, 122, 144107. [PubMed: 15847516]

(77). Lu N; Kofke DA Accuracy of Free-Energy Perturbation Calculations in Molecular Simulation. I. Modeling. J. Chem. Phys 2001, 114, 7303–7311.

(78). Lu N; Kofke DA Accuracy of Free-Energy Perturbation Calculations in Molecular Simulation. II. Heuristics. J. Chem. Phys 2001, 115, 6866–6875.

(79). Pohorille A; Jarzynski C; Chipot C Good Practices in Free-Energy Calculations. J. Phys. Chem. B 2010, 114, 10235–10253. [PubMed: 20701361]

(80). Bruckner S; Boresch S Efficiency of Alchemical Free Energy Simulations. I. A Practical Comparison of the Exponential Formula, Thermodynamic Integration, and Bennett's Acceptance Ratio Method. J. Comput. Chem 2011, 32, 1303–1319. [PubMed: 21425288]

(81). Bruckner S; Boresch S Efficiency of Alchemical Free Energy Simulations. II. Improvements for Thermodynamic Integration. J. Comput. Chem 2011, 32, 1320–1333. [PubMed: 21425289]

(82). De Ruiter A; Boresch S; Oostenbrink C Comparison of Thermodynamic Integration and Bennett Acceptance Ratio for Calculating Relative Protein-Ligand Binding Free Energies. J. Comput. Chem 2013, 34, 1024–1034. [PubMed: 23335287]

(83). Klimovich PV; Shirts MR; Mobley DL Guidelines for the Analysis of Free Energy Calculations.J. Comput.-Aided Mol. Des 2015, 29, 397–411.

(84). Jiang W; Chipot C; Roux B Computing Relative Binding Affinity of Ligands to Receptor: An Effective Hybrid Single-Dual-Topology Free-Energy Perturbation Approach in NAMD. J. Chem. Inf. Model 2019, 59, 3794–3802. [PubMed: 31411473]

(85). Gumbart JC; Roux B; Chipot C Standard Binding Free Energies from Computer Simulations: What Is the Best Strategy?J. Chem. Theory Comput 2013, 9, 794–802. [PubMed: 23794960]

(86). König G; Boresch S Non-Boltzmann Sampling and Bennett's Acceptance Ratio Method: How to Profit from Bending the Rules. J. Comput. Chem 2011, 32, 1082–1090. [PubMed: 21387335]

(87). Wu D; Kofke DA Phase-Space Overlap Measures. I. Fail-Safe Bias Detection in Free Energies Calculated by Molecular Simulation. J. Chem. Phys 2005, 123, 54103.

(88). Wu D; Kofke DA Phase-Space Overlap Measures. II. Design and Implementation of Staging Methods for Free-Energy Calculations. J. Chem. Phys 2005, 123, 84109.

(89). Beutler TC; Mark AE; René C van Schaik and Paul R. Gerber and Wilfred F. van Gunsteren Avoiding Singularities and Numerical Instabilities in Free Energy Calculations Based on Molecular Simulations. Chem. Phys. Lett 1994, 222, 529–539.

(90). Steinbrecher T; Joung I; Case DA Soft-Core Potentials in Thermodynamic Integration: Comparing One- and Two-Step Transformations. J. Comput. Chem 2011, 32, 3253–3263. [PubMed: 21953558]

(91). Hritz J; Oostenbrink C Hamiltonian Replica Exchange Molecular Dynamics Using Soft-Core Interactions. J. Chem. Phys 2008, 128, 144121. [PubMed: 18412437]

(92). Riniker S; Christ CD; Hansen HS; Hünenberger PH; Oostenbrink C; Steiner D; van Gunsteren WF Calculation of Relative Free Energies for Ligand-Protein Binding, Solvation, and Conformational Transitions Using the GROMOS Software. J. Phys. Chem. B 2011, 115, 13570–13577. [PubMed: 22039957]

(93). Ruiter A. d.; Oostenbrink C. Extended Thermodynamic Integration: Efficient Prediction of Lambda Derivatives at Nonsimulated Points. J. Chem. Theory Comput 2016, 12, 4476–4486, ppublish. [PubMed: 27494138]

(94). Pal RK; Gallicchio E Perturbation Potentials to Overcome Order/Disorder Transitions in Alchemical Binding Free Energy Calculations. J. Chem. Phys 2019, 151, 124116. [PubMed: 31575187]

(95). Boresch S; Karplus M The Role of Bonded Terms in Free Energy Simulations. 2. Calculation of Their Influence on Free Energy Differences of Solvation. J. Phys. Chem. A 1999, 103, 119–136.

(96). Boresch S; Karplus M The Role of Bonded Terms in Free Energy Simulations: 1. Theoretical Analysis. J. Phys. Chem. A 1999, 103, 103–118.

(97). Shobana S; Roux B; Andersen OS Free Energy Simulations: Thermodynamic Reversibility and Variability. J. Phys. Chem. B 2000, 104, 5179–5190.

(98). Loeffler HH; Bosisio S; Duarte Ramos Matos G; Suh D; Roux B; Mobley DL; Michel J Reproducibility of Free Energy Calculations across Different Molecular Simulation Software Packages. J. Chem. Theory Comput 2018, 14, 5567–5582. [PubMed: 30289712]

(99). Ding X; Vilseck JZ; Hayes RL; Brooks CL Gibbs Sampler-Based $\lambda$-Dynamics and Rao-Blackwell Estimator for Alchemical Free Energy Calculation. J. Chem. Theory Comput 2017, 13, 2501–2510. [PubMed: 28510433]

(100). Hayes RL; Armacost KA; Vilseck JZ; Brooks CL Adaptive Landscape Flattening Accelerates Sampling of Alchemical Space in Multisite λ Dynamics. J. Phys. Chem. B 2017, 121, 3626–3635. [PubMed: 28112940]

(101). Guo Z; Brooks CL Rapid Screening of Binding Affinities: Application of the λ-Dynamics Method to a Trypsin-Inhibitor System. J. Am. Chem. Soc 1998, 120, 1920–1921.

(102). Guo Z; Brooks CL; Kong X Efficient and Flexible Algorithm for Free Energy Calculations Using the λ-Dynamics Approach. J. Phys. Chem. B 1998, 102, 2032–2036.

(103). Jiang W; Roux B Free Energy Perturbation Hamiltonian Replica-Exchange Molecular Dynamics (FEP/H-REMD) for Absolute Ligand Binding Free Energy Calculations. J. Chem. Theory Comput 2010, 6, 2559–2565. [PubMed: 21857813]

(104). Arrar M; de Oliveira CAF; Fajer M; Sinko W; McCammon JA w-REXAMD: A Hamiltonian Replica Exchange Approach to Improve Free Energy Calculations for Systems with Kinetically Trapped Conformations. J. Chem. Theory Comput 2013, 9, 18–23. [PubMed: 23316122]

(105). Itoh SG; Okumura H Hamiltonian Replica-Permutation Method and Its Applications to an Alanine Dipeptide and Amyloid-β(29–42) Peptides. J. Comput. Chem 2013, 34, 2493–2497. [PubMed: 23925979]

(106). Armacost KA; Goh GB; Brooks CL Biasing Potential Replica Exchange Multisite λ-Dynamics for Efficient Free Energy Calculations. J. Chem. Theory Comput 2015, 11, 1267–1277. [PubMed: 26579773]

(107). Yang M; Huang J; MacKerell AD Jr. Enhanced Conformational Sampling Using Replica Exchange with Concurrent Solute Scaling and Hamiltonian Biasing Realized in One Dimension. J. Chem. Theory Comput 2015, 11, 2855–2867. [PubMed: 26082676]

(108). Babin V; Roland C; Sagui C Adaptively Biased Molecular Dynamics for Free Energy Calculations. J. Chem. Phys 2008, 128, 134101. [PubMed: 18397047]

(109). Darve E; Rodríguez-Gómez D; Pohorille A Adaptive Biasing Force Method for Scalar and Vector Free Energy Calculations. J. Chem. Phys 2008, 128, 144120. [PubMed: 18412436]

(110). Carlson DE; Stinson P; Pakman A; Paninski L In Proceedings of the 33rd International Conference on Machine Learning, JMLR.org: New York, NY, USA, 2016; Vol. 48, pp 2896–2905.

(111). Tan Z Optimally Adjusted Mixture Sampling and Locally Weighted Histogram Analysis. J. Comput. Graph. Stat 2017, 26, 54–65.

(112). Liu S; Wu Y; Lin T; Abel R; Redmann JP; Summa CM; Jaber VR; Lim NM; Mobley DL Lead Optimization Mapper: Automating Free Energy Calculations for Lead Optimization. J. Comput. Aided Mol. Des 2013, 27, 755–770. [PubMed: 24072356]

(113). Loeffler HH; Michel J; Woods C FESetup: Automating Setup for Alchemical Free Energy Simulations. J. Chem. Inf. Model 2015, 55, 2485–2490. [PubMed: 26544598]

(114). Homeyer N; Gohlke H FEW: A Workflow Tool for Free Energy Calculations of Ligand Binding. J. Comput. Chem 2013, 34, 965–973. [PubMed: 23288722]

(115). Shirts MR; Mobley DL In; Humana Press: 2013, pp 271–311.

(116). Kuhn M; Firth-Clark S; Tosco P; Mey ASJS; Mackey MD; Michel J Assessment of Binding Affinity via Alchemical Free Energy Calculations. J. Chem. Inf. Model 2020, 3120–3130. [PubMed: 32437145]

(117). Stone JE; Hardy DJ; Ufimtsev IS; Schulten K GPU-Accelerated Molecular Modeling Coming of Age. J. Mol. Graphics Model 2010, 29, 116–125.

(118). Mermelstein DJ; Lin C; Nelson G; Kretsch R; McCammon JA; Walker RC Fast and Flexible GPU Accelerated Binding Free Energy Calculations within The AMBER Molecular Dynamics Package. J. Comp. Chem 2018, 39, 1354–1358. [PubMed: 29532496]

(119). Hopkins CW; Le Grand S; Walker RC; Roitberg AE Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. J. Chem. Theory Comput 2015, 11, 1864–1874. [PubMed: 26574392]

(120). Boresch S The Role of Bonded Energy Terms in Free Energy Simulations - Insights from Analytical Results. Mol. Simul 2002, 28, 13–37.

(121). Wang L; Wu Y; Deng Y; Kim B; Pierce L; Krilov G; Lupyan D; Robinson S; Dahlgren MK; Greenwood J; Romero DL; Masse C; Knight JL; Steinbrecher T; Beuming T; Damm W; Harder

E; Sherman W; Brewer M; Wester R; Murcko M; Frye L; Farid R; Lin T; Mobley DL; Jorgensen WL; Berne BJ; Friesner RA; Abel R Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-energy Calculation Protocol and Force Field. J. Am. Chem. Soc 2015, 137, 2695–2703. [PubMed: 25625324]

(122). Phillips JC; Braun R; Wang W; Gumbart J; Tajkhorshid E; Villa E; Chipot C; Skeel RD; Kalé L; Schulten K Scalable Molecular Dynamics with NAMD. J. Comput. Chem 2005, 26, 1781–1802. [PubMed: 16222654]

(123). Huggins DJ; Sherman W; Tidor B Rational Approaches to Improving Selectivity in Drug Design. Journal of Medicinal Chemistry 2012, 55, 1424–1444. [PubMed: 22239221]

(124). Wang L; Deng Y; Wu Y; Kim B; LeBard DN; Wandschneider D; Beachy M; Friesner RA; Abel R Accurate Modeling of Scaffold Hopping Transformations in Drug Discovery. J. Chem. Theory Comput 2017, 13, 42–54. [PubMed: 27933808]

(125). Wang J; Deng Y; Roux B Absolute Binding Free Energy Calculations Using Molecular Dynamics with Restraining Potentials. Biophys. J 2006, 91, 2798–2814. [PubMed: 16844742]

(126). Jayachandran G; Shirts MR; Park S; Pande VS Parallelized-Over-Parts Computation of Absolute Binding Free Energy with Docking and Molecular Dynamics. J. Chem. Phys 2006, 125, 084902. [PubMed: 16965052]

(127). Mobley DL; Chodera JD; Dill KA Confine-And-Release Method: Obtaining Correct Binding Free Energies in the Presence of Protein Conformational Change. J. Chem. Theory Comput 2007, 3, 1231–1235. [PubMed: 18843379]

(128). Gallicchio E; Levy RM Recent Theoretical and Computational Advances in Modeling Protein-Ligand Binding Affinities. Adv. Protein Chem. Struct. Biol 2011, 85, 27–80. [PubMed: 21920321]

(129). Steinbrecher TB; Dahlgren M; Cappel D; Lin T; Wang L; Krilov G; Abel R; Friesner R; Sherman W Accurate Binding Free Energy Predictions in Fragment Optimization. J. Chem. Inf. Model 2015, 55, 2411–2420. [PubMed: 26457994]

(130). Chen J; Kriwacki RW Intrinsically Disordered Proteins: Structure, Function and Therapeutics. J. Mol. Biol 2018, 430, 2275. [PubMed: 29906412]

(131). Kim MO; Blachly PG; Kaus JW; McCammon JAJ. Phys. Chem. B 2014, 119, 861–872. [PubMed: 25134690]

(132). Shobana S; Roux B; Andersen OS Free Energy Simulations: Thermodynamic Reversibility and Variability. J. Phys. Chem. B 2000, 104, 5179–5190.

(133). Wang J; Wolf RM; Caldwell JW; Kollman PA; Case DA Development and Testing of a General AMBER Force Field. J. Comp. Chem 2004, 25, 1157–1174. [PubMed: 15116359]

(134). Träg J; Zahn D Improved GAFF2 Parameters for Fluorinated Alkanes and Mixed Hydro-and Fluorocarbons. J. Mol. Model 2019, 25, 39. [PubMed: 30659357]

(135). Vanommeslaeghe K; MacKerell A Jr. Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. J. Chem. Inf. Model 2012, 52, 3144–3154. [PubMed: 23146088]

(136). Vanommeslaeghe K; Raman PE; MacKerell A Jr. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. J. Chem. Inf. Model 2012, 52, 3155–3168. [PubMed: 23145473]

(137). Jorgensen WL; Tirado-Rives J The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin. J. Am. Chem. Soc 1988, 110, 1657–1666. [PubMed: 27557051]

(138). Shivakumar D; Williams J; Wu Y; Damm W; Shelley J; Sherman W Prediction of Absolute Solvation Free Energies Using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. J. Chem. Theory Comput 2010, 6, 1509–1519. [PubMed: 26615687]

(139). Shivakumar D; Harder E; Damm W; Friesner RA; Sherman W Improving the Prediction of Absolute Solvation Free Energies Using the Next Generation OPLS Force Field. J. Chem. Theory Comput 2012, 8, 2553–2558. [PubMed: 26592101]

(140). Roos K; Wu C; Damm W; Reboul M; Stevenson JM; Lu C; Dahlgren MK; Mondal S; Chen W; Wang L; Abel R; Friesner RA; Harder ED OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. J. Chem. Theory Comput 2019, 1863–1874. [PubMed: 30768902]

(141). Ufimtsev IS; Martinez TJ Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. J. Chem. Theory Comput 2009, 5, 2619–2628. [PubMed: 26631777]

(142). Bayly CI; Cieplak P; Cornell W; Kollman PA A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. J. Phys. Chem 1993, 97, 10269–10280.

(143). Shinada NK; Brevern A. G. d.; Schmidtke P Halogens in Protein-Ligand Binding Mechanism: A Structural Perspective. J. Med. Chem 2019, DOI: 10.1021/acs.jmedchem.8b01453.

(144). Mayne CG; Saam J; Schulten K; Tajkhorshid E; Gumbart JC Rapid Parameterization of Small Molecules Using the Force Field Toolkit. J. Comp. Chem 2013, 34, 2757–2770. [PubMed: 24000174]

(145). Slochower DR; Henriksen NM; Wang L-P; Chodera JD; Mobley DL; Gilson MK Binding Thermodynamics of HostâGuest Systems with SMIRNOFF99Frosst 1.0.5 from the Open Force Field Initiative. J. Chem. Theory Comput 2019, 15, 6225–6242. [PubMed: 31603667]

(146). Sastry GM; Adzhigirey M; Day T; Annabhimoju R; Sherman W Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. J. Comput. Aided Mol. Des 2013, 27, 221–234. [PubMed: 23579614]

(147). Ten Brink T; Exner TE Influence of Protonation, Tautomeric, and Stereoisomeric States on Protein-ligand Docking Results. J. Chem. Inf. Model 2009, 49, 1535–1546. [PubMed: 19453150]

(148). Martínez-Rosell G; Giorgino T; De Fabritiis G PlayMolecule ProteinPrepare: A Web Application for Protein Preparation for Molecular Dynamics Simulations. J. Chem. Inf. Model 2017, 57, 1511–1516. [PubMed: 28594549]

(149). Jain AN Bias, Reporting, and Sharing: Computational Evaluations of Docking Methods. J. Comput. Aided Mol. Des 2008, 22, 201–212. [PubMed: 18075713]

(150). Laskowski RA; MacArthur MW; Moss DS; Thornton JM PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. Journal of applied crystallography 1993, 26, 283–291.

(151). Hooft RW; Vriend G; Sander C; Abola EE Errors in Protein Structures. Nature 1996, 381, 272–272. [PubMed: 8692262]

(152). Chen VB; Arendall WB; Headd JJ; Keedy DA; Immormino RM; Kapral GJ; Murray LW; Richardson JS; Richardson DC Mol-Probity: All-atom Structure Validation for Macromolecular Crystallography. Acta Cryst. D 2010, 66, 12–21. [PubMed: 20057044]

(153). Pontius J; Richelle J; Wodak SJ Deviations from Standard Atomic Volumes as a Quality Measure for Protein Crystal Structures. J. Mol. Biol 1996, 264, 121–136. [PubMed: 8950272]

(154). Zhao H; Caflisch A Molecular Dynamics in Drug Design. European J. Med. Chem 2015, 91, 4–14. [PubMed: 25108504]

(155). Wang Y; Zhu G-F; Ren S-Y; Han Y-G; Luo Y; Du L-F Insight into the Structural Stability of Wild Type and Mutants of the Tobacco Etch Virus Protease with Molecular Dynamics Simulations. Journal of molecular modeling 2013, 19, 4865–4875. [PubMed: 24043540]

(156). Mirjalili V; Noyes K; Feig M Physics-based Protein Structure Refinement through Multiple Molecular Dynamics Trajectories and Structure Averaging. Proteins 2014, 82, 196–207. [PubMed: 23737254]

(157). Feig M; Mirjalili V Protein Structure Refinement via Molecular-Dynamics Simulations: What Works and What Does Not?Proteins 2016, 84, 282–292. [PubMed: 26234208]

(158). Heo L; Feig M Experimental Accuracy in Protein Structure Refinement via Molecular Dynamics Simulations. Proc. Nat. Acad. Sci. USA 2018, 115, 13276–13281. [PubMed: 30530696]

(159). Vriend G WHAT IF: A Molecular Modeling and Drug Design Program. Journal of molecular graphics 1990, 8, 52–56. [PubMed: 2268628]

(160). Olsson MH; Søndergaard CR; Rostkowski M; Jensen JH PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical p$K_a$ Predictions. J. Chem. Theory Comput 2011, 7, 525–537. [PubMed: 26596171]

(161). Emsley P; Cowtan K Coot: Model-Building Tools for Molecular Graphics. Acta Cryst. D 2004, 60, 2126–2132. [PubMed: 15572765]

(162). Corbeil CR; Moitessier N Docking Ligands into Flexible and Solvated Macromolecules. 3. Impact of Input Ligand Conformation, Protein Flexibility, and Water Molecules on The Accuracy of Docking Programs. J. Chem. Inf. Model 2009, 49, 997–1009. [PubMed: 19391631]

(163). Verdonk ML; Chessari G; Cole JC; Hartshorn MJ; Murray CW; Nissink JWM; Taylor RD; Taylor R Modeling Water Molecules in Protein-Ligand Docking Using GOLD. J. Med. Chem 2005, 48, 6504–6515. [PubMed: 16190776]

(164). Lemmon G; Meiler J Towards Ligand Docking Including Explicit Interface Water Molecules. PloS One 2013, 8.

(165). Rarey M; Kramer B; Lengauer T The Particle Concept: Placing Discrete Water Molecules During Protein-ligand Docking Predictions. Proteins 1999, 34, 17–28. [PubMed: 10336380]

(166). Robinson DD; Sherman W; Farid R Understanding Kinase Selectivity Through Energetic Analysis of Binding Site Waters. ChemMedChem 2010, 5, 618–627. [PubMed: 20183853]

(167). Lenselink EB; Beuming T; Sherman W; van Vlijmen HW; IJzerman AP Selecting an Optimal Number of Binding Site Waters to Improve Virtual Screening Enrichments Against the Adenosine A2A Receptor. J. Chem. Inf. Model 2014, 54, 1737–1746. [PubMed: 24835542]

(168). Snyder PW; Mecinović J; Moustakas DT; Thomas SW; Harder M; Mack ET; Lockett MR; Héroux A; Sherman W; Whitesides GM Mechanism of the Hydrophobic Effect in the Biomolecular Recognition of Arylsulfonamides by Carbonic Anhydrase. Proc. Natl. Acad. Sci. USA 2011, 108, 17889–17894. [PubMed: 22011572]

(169). Breiten B; Lockett MR; Sherman W; Fujita S; Al-Sayah M; Lange H; Bowers CM; Heroux A; Krilov G; Whitesides GM Water Networks Contribute to Enthalpy/Entropy Compensation in Protein–Ligand Binding. J. Am. Chem. Soc 2013, 135, 15579–15584. [PubMed: 24044696]

(170). Fox JM; Kang K; Sastry M; Sherman W; Sankaran B; Zwart PH; Whitesides GM Water-Restructuring Mutations Can Reverse the Thermodynamic Signature of Ligand Binding to Human Carbonic Anhydrase. Angew. Chem. Int. Ed 2017, 56, 3833–3837.

(171). Imai T; Kovalenko A; Hirata F Solvation Thermodynamics of Protein Studied by the 3D-RISM Theory. Chem. Phys. Lett 2004, 395, 1–6.

(172). Ross GA; Bodnarchuk MS; Essex JW Water Sites, Networks, and Free Energies with Grand Canonical Monte Carlo. J. Am. Chem. Soc 2015, 137, 14930–14943. [PubMed: 26509924]

(173). Michel J; Tirado-Rives J; Jorgensen WL Prediction of the Water Content in Protein Binding Sites. J. Phys. Chem. B 2009, 113, 13337–13346. [PubMed: 19754086]

(174). Beuming T; Che Y; Abel R; Kim B; Shanmugasundaram V; Sherman W Thermodynamic Analysis of Water Molecules at The Surface of Proteins and Applications to Binding Site Prediction and Characterization. Proteins 2012, 80, 871–883. [PubMed: 22223256]

(175). Bodnarchuk MS Water, Water, Everywhere… It's Time to Stop and Think. Drug Discovery Today 2016, 21, 1139–1146. [PubMed: 27210724]

(176). Nittinger E; Flachsenberg F; Bietz S; Lange G; Klein R; Rarey M Placement of Water Molecules in Protein Structures: From Large-scale Evaluations to Single-case Examples. J. Chem. Inf. Model 2018, 58, 1625–1637. [PubMed: 30036062]

(177). Mayol E; García-Recio A; Tiemann JK; Hildebrand PW; Guixà-González R; Olivella M; Cordomí A HomolWat: A Web Server Tool to Incorporate 'Homologous' Water Molecules into GPCR Structures. Nucleic Acids Research 2020.

(178). Li Y; Gao YD; Holloway MK; Wang R Prediction of the Favorable Hydration Sites in a Protein Binding Pocket and Its Application to Scoring Function Formulation. J. Chem. Inf. Model 2020.

(179). Ben-Shalom IY; Lin C; Kurtzman T; Walker RC; Gilson MK Simulating Water Exchange to Buried Binding Sites. J. Chem. Theory Comput 2019, 15, 2684–2691. [PubMed: 30835999]

(180). Ben-Shalom IY; Lin C; Kurtzman T; Walker RC; Gilson MK Equilibration of Buried Water Molecules to Enhance Protein-Ligand Binding Free Energy Calculations. Biophys. J 2020, 118, 144a.

(181). Bodnarchuk MS; Packer MJ; Haywood A Utilizing Grand Canonical Monte Carlo Methods in Drug Discovery.ACS Med. Chem. Lett 2019, 11, 77–82.

(182). Deng Y; Roux B Computation of Binding Free Energy with Molecular Dynamics and Grand Canonical Monte Carlo Simulations. J. Chem. Phys 2008, 128, 03B611.

(183). Ross G; Russell E; Deng Y; Lu C; Harder E; Abel R; Wang L Enhancing Water Sampling in Free Energy Calculations with Grand Canonical Monte Carlo2020, DOI: 10.26434/chemrxiv.12595073.v1.

(184). McGovern SL; Shoichet BK Information Decay in Molecular Docking Screens against Holo, Apo, and Modeled Conformations of Enzymes. J. Med. Chem 2003, 46, 2895–2907. [PubMed: 12825931]

(185). Sherman W; Day T; Jacobson MP; Friesner RA; Farid R Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. J. Med. Chem 2006, 49, 534–553. [PubMed: 16420040]

(186). Miller E; Murphy R; Sindhikara D; Borrelli K; Grisewood M; Ranalli F; Dixon S; Jerome S; Boyles N; Day T; Ghanakota P; Mondal S; Rafi SB; Troast DM; Abel R; Friesner R A Reliable and Accurate Solution to the Induced Fit Docking Problem for Protein-Ligand Binding2020, DOI: 10.26434/chemrxiv.11983845.v1.

(187). Evangelista Falcon W; Ellingson SR; Smith JC; Baudry J Ensemble Docking in Drug Discovery: How Many Protein Configurations from Molecular Dynamics Simulations are Needed To Reproduce Known Ligand Binding?J. Phys. Chem. B 2019, 123, 5189–5195. [PubMed: 30695645]

(188). Rao S; Sanschagrin PC; Greenwood JR; Repasky MP; Sherman W; Farid R Improving Database Enrichment Through Ensemble Docking. J. Comput. Aided Mol. Des 2008, 22, 621–627. [PubMed: 18253700]

(189). Alonso H; Bliznyuk AA; Gready JE Combining Docking and Molecular Dynamic Simulations in Drug Design. Med. Res. Rev 2006, 26, 531–568. [PubMed: 16758486]

(190). Zhou X; Chou J; Wong STC Protein structure similarity from Principle Component Correlation analysis. BMC Bioinf 2006, 7, 40.

(191). Liu K; Kokubo H Exploring the Stability of Ligand Binding Modes to Proteins by Molecular Dynamics Simulations: A Cross-Docking Study.J. Chem. Inf. Model 2017, 57, 2514–2522. [PubMed: 28902511]

(192). Liu K; Watanabe E; Kokubo H Exploring the stability of ligand binding modes to proteins by molecular dynamics simulations.J. Chem. Inf. Model 2017, 31, 201–211.

(193). Ruiz-Carmona S; Alvarez-Garcia D; Foloppe N; Garmendia-Doval AB; Juhos S; Schmidtke P; Barril X; Hubbard RE; Morley SD rDock: A Fast, Versatile and Open Source Program for Docking Ligands to Proteins and Nucleic Acids. PLOS Comp. Bio 2014, 10, 1–7.

(194). Liu S; Wang L; Mobley DL Is Ring Breaking Feasible in Relative Binding Free Energy Calculations?J. of Chem. Inf. Model 2015, 55, 727–735. [PubMed: 25835054]

(195). Liu S; Wu Y; Lin T; Abel R; Redmann JP; Summa CM; Jaber VR; Lim NM; Mobley DL Lead Optimization Mapper: Automating Free Energy Calculations for Lead Optimization. J. Comput. Aided Mol. Des 2013, 27, 755–70. [PubMed: 24072356]

(196). Chodera JD; Shirts MR Replica Exchange and Expanded Ensemble Simulations as Gibbs Sampling: Simple Improvements for Enhanced Mixing. J. Chem. Phys 2011, 135, 194110. [PubMed: 22112069]

(197). Knight JL; Brooks CL III, $\lambda$ - Dynamics Free Energy Simulation Methods. J. Comp. Chem 2009, 30, 1692–1700. [PubMed: 19421993]

(198). Radak Brian K. Suh Donghyuk, R. B A Generalized Linear Response Framework for Expanded Ensemble and Replica Exchange Simulations. J. Chem. Phys 2018, 149, 072315. [PubMed: 30134700]

(199). Jandova Z; Fast D; Setz M; Pechlaner M; Oostenbrink C Saturation Mutagenesis by Efficient Free-Energy Calculation. J. Chem. Theory Comput 2018, 14, 894–904. [PubMed: 29262673]

(200). Lin Z; Van Gunsteren WF; Liu H Conformational State-specific Free Energy Differences by One-step Perturbation: Protein Secondary Structure Preferences of the GROMOS 43A1 and 53A6 Force Fields. J. Comp. Chem 2011, 32, 2290–2297. [PubMed: 21541965]

(201). Oostenbrink C; van Gunsteren WF Free Energies of Ligand Binding for Structurally Diverse Compounds. Proc. Nat. Acad. Sci. USA 2005, 102, 6750–6754. [PubMed: 15767587]

(202). Liu H; Mark AE; van Gunsteren WF Estimating the Relative Free Energy of Different Molecular States with Respect to a Single Reference State. J. Phys. Chem 1996, 100, 9485–9494.

(203). Rocklin GJ; Mobley DL; Dill KA; Hünenberger PH Calculating the Binding Free Energies of Charge Species Based on Explicit-Solvent Simulations Employing Lattice-Sum Methods: An Accurate Correction Scheme for Electrostatic Finite-Size Effects. J. Chem. Phys 2013, 139, 184103. [PubMed: 24320250]

(204). Chen W; Deng Y; Russell E; Wu Y; Abel R; Wang L Accurate Calculation of Relative Binding Free Energies between Ligands with Different Net Charges. J. Chem. Theory Comput 2018, 14, 6346–6358. [PubMed: 30375870]

(205). Lin Y-L; Aleksandrov A; Simonson T; Roux B An Overview of Electrostatic Free Energy Computations for Solutions and Proteins. J. Chem. Theory Comput 2014, 10, 2690–2709. [PubMed: 26586504]

(206). Homeyer N; Stoll F; Hillisch A; Gohlke H Binding Free Energy Calculations for Lead Optimization: Assessment of Their Accuracy in an Industrial Drug Design Context. J. Chem. Theory Comput 2014, 10, 3331–3344. [PubMed: 26588302]

(207). Abel R; Wang L; Mobley DL; Friesner RA A Critical Review of Validation, Blind Testing, and Real- World Use of Alchemical Protein-Ligand Binding Free Energy Calculations. Curr. Top. Med. Chem 2017, 17, 2577–2585. [PubMed: 28413950]

(208). Zuckerman DM; Woolf TB Overcoming Finite-Sampling Errors in Fast-Switching Free-Energy Estimates: Extrapolative Analysis of a Molecular System. Chem. Phys. Lett 2002, 351, 445–453.

(209). Caves LS; Evanseck JD; Karplus M Locally Accessible Conformations of Proteins: Multiple Molecular Dynamics Simulations of Crambin. Protein Sci 1998, 7, 649–666. [PubMed: 9541397]

(210). Loccisano AE; Acevedo O; DeChancie J; Schulze BG; Evanseck JD Enhanced Sampling by Multiple Molecular Dynamics Trajectories: Carbonmonoxy Myoglobin 10 micros A0→A(1–3) Transition from Ten 400 Picosecond Simulations. J. Mol. Graphics Modell 2004, 22, 369–376.

(211). Likic VA; Gooley PR; Speed TP; Strehler EE A Statistical Approach to the Interpretation of Molecular Dynamics Simulations of Calmodulin Equilibrium Dynamics. Protein Sci 2005, 14, 2955–2963. [PubMed: 16322577]

(212). Elofsson A; Nilsson L How Consistent are Molecular Dynamics Simulations? Comparing Structure and Dynamics in Reduced and Oxidized Escherichia Coli Thioredoxin. J. Mol. Biol 1993, 233, 766–780. [PubMed: 8411178]

(213). Pan AC; Xu H; Palpant T; Shaw DE Quantitative Characterization of the Binding and Unbinding of Millimolar Drug Fragments with Molecular Dynamics Simulations. J. Chem. Theory Comput 2017, 13, 3372–3377. [PubMed: 28582625]

(214). Shaw DE; Grossman JP; Bank JA; Batson B; Butts JA; Chao JC; Deneroff MM; Dror RO; Even A; Fenton CH; Forte A; Gagliardo J; Gill G; Greskamp B; Ho CR; Ierardi DJ; Iserovich L; Kuskin JS; Larson RH; Layman T; Lee L-S; Lerer AK; Li C; Killebrew D; Mackenzie KM; Mok SY-H; Moraes MA; Mueller R; Nociolo LJ; Peticolas JL; Quan T; Ramot D; Salmon JK; Scarpazza DP; Ben Schafer U; Siddique N; Snyder CW; Spengler J; Tang PTP; Theobald M; Toma H; Towles B; Vitale B; Wang SC; Young C Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer2014, 41–53.

(215). Gohlke H; Kiel C; Case DA Insights into Protein-Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras-Raf and Ras-RalGDS Complexes. J. Mol. Biol 2003, 330, 891–913. [PubMed: 12850155]

(216). Miao Y; Feher VA; McCammon JA Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. J. Chem. Theory Comput 2015, 11, 3584–3595. [PubMed: 26300708]

(217). Huang Y.-m. M.; McCammon JA; Miao Y Replica Exchange Gaussian Accelerated Molecular Dynamics: Improved Enhanced Sampling and Free Energy Calculation. J. Chem. Theory Comput 2018, 14, 1853–1864. [PubMed: 29489349]

(218). Miao Y; Bhattarai A; Wang J Ligand Gaussian Accelerated Molecular Dynamics (LiGaMD): Characterization of Ligand Binding Thermodynamics and Kinetics. bioRxiv 2020, DOI: 10.1101/2020.04.20.051979.

(219). Jo S; Jiang W A Generic Implementation of Replica Exchange with Solute Tempering (REST2) Algorithm in NAMD for Complex Bio-physical Simulations. Comput. Phys. Commun 2015, 197, 304–311.

(220). Wang L; Berne B; Friesner RA On Achieving High Accuracy and Reliability in the Calculation of Relative Protein–Ligand Binding Affinities. Proc. Nat. Acad. Sci. USA 2012, 109, 1937–1942. [PubMed: 22308365]

(221). Mobley DL; Chodera JD; Dill KA Confine-and-Release Method: Obtaining Correct Binding Free Energies in the Presence of Protein Conformational Change. J. Chem. Theory Comput 2007, 3, 1231–1235. [PubMed: 18843379]

(222). Leitgeb M; Schröder C; Boresch S Alchemical Free Energy Calculations and Multiple Conformational Substates. J. Chem. Phys 2005, 122, 084109.

(223). Laio A; Gervasio FL Metadynamics: A Method to Simulate Rare Events and Reconstruct the Free Energy in Biophysics, Chemistry and Material Science. Rep. Prog. Phys 2008, 71, 126601.

(224). Laio A; Parrinello M Escaping Free-energy Minima. Proc. Nat. Acad. Sci. USA 2002, 99, 12562–12566. [PubMed: 12271136]

(225). Fu H; Shao X; Chipot C; Cai W Extended Adaptive Biasing Force Algorithm. An On-The-Fly Implementation for Accurate Free-Energy Calculations. J. Chem. Theory Comput 2016, 12, 3506–3513. [PubMed: 27398726]

(226). Enyedy IJ; Egan WJ Can We Use Docking and Scoring for Hit-To-Lead Optimization?J. Comput.-Aided Mol. Des 2008, 22, 161–168. [PubMed: 18183356]

(227). Sakae Y; Zhang BW; Levy RM; Deng N Absolute Protein Binding Free Energy Simulations for Ligands with Multiple Poses, a Thermodynamic Path That Avoids Exhaustive Enumeration of the Poses. J. Comp. Chem 2020, 41, 56–68.

(228). Gill SC; Lim NM; Grinaway PB; Rustenburg AS; Fass J; Ross GA; Chodera JD; Mobley DL Binding Modes of Ligands Using Enhanced Sampling (BLUES): Rapid Decorrelation of Ligand Binding Modes via Nonequilibrium Candidate Monte Carlo. J. Phys. Chem. B 2018, 122, 5579–5598. [PubMed: 29486559]

(229). Tian C; Kasavajhala K; Belfon KA; Raguette L; Huang H; Migues AN; Bickel J; Wang Y; Pincay J; Wu Q, et al. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. J. Chem. Theory Comput 2019, 16, 528–552. [PubMed: 31714766]

(230). Maier JA; Martinez C; Kasavajhala K; Wickstrom L; Hauser KE; Simmerling C ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J. Chem. Theory Comput 2015, 11, 3696–3713. [PubMed: 26574453]

(231). Hornak V; Abel R; Okur A; Strockbine B; Roitberg A; Simmerling C Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. Proteins 2006, 65, 712–725. [PubMed: 16981200]

(232). Cavasotto CN; Adler NS; Aucar MG Quantum Chemical Approaches in Structure-Based Virtual Screening and Lead Optimization. Front. Chem 2018, 6, 188. [PubMed: 29896472]

(233). Giese TJ; Chen H; Huang M; York DM Parametrization of an Orbital-Based Linear-Scaling Quantum Force Field for Noncovalent Interactions. J. Chem. Theory Comput 2014, 10, 1086–1098. [PubMed: 24803856]

(234). Giese TJ; Huang M; Chen H; York DM Recent Advances toward a General Purpose Linear-Scaling Quantum Force Field. Acc. Chem. Res 2014, 47, 2812–20. [PubMed: 24937206]

(235). Giese TJ; York DM Quantum mechanical force fields for condensed phase molecular simulations. J. Phys. Condens. Matter 2017, 29, 383002. [PubMed: 28817382]

(236). Lu X; Fang D; Ito S; Okamoto Y; Ovchinnikov V; Cui Q QM/MM Free Energy Simulations: Recent Progress and Challenges. Mol. Simul 2016, 42, 1056–1078. [PubMed: 27563170]

(237). Kamerlin SCL; Haranczyk M; Warshel A Progress in Ab Initio QM/MM Free-Energy Simulations of Electrostatic Energies in Proteins: Accelerated QM/MM Studies of pKa, Redox Reactions and Solvation Free Energies. J. Phys. Chem. B 2009, 113, 1253–1272. [PubMed: 19055405]

(238). Duarte F; Amrein BA; Blaha-Nelson D; Kamerlin SC Recent Advances in QM/MM Free Energy Calculations Using Reference Potentials. Biochim. Biophys. Acta 2015, 1850, 954–965. [PubMed: 25038480]

(239). Rathore RS; Sumakanth M; Siva Reddy M; Reddanna P; Rao AA; Erion MD; Reddy MR Advances in Binding Free Energies Calculations: QM/MM-Based Free Energy Perturbation Method for Drug Design. Curr. Pharm. Des 2013, 19, 4674–4686. [PubMed: 23260025]

(240). Gao J Absolute Free Energy of Solvation from Monte Carlo Simulations Using Combined Quantum and Molecular Mechanical Potentials. J. Phys. Chem 1992, 96, 537–540.

(241). Gao J; Xia X A priori Evaluation of Aqueous Polarization Effects Through Monte Carlo QM-MM Simulations. Science 1992, 258, 631–635. [PubMed: 1411573]

(242). Luzhkov V; Warshel A Microscopic Models for Quantum Mechanical Calculations of Chemical Processes in Solutions: LD/AMPAC and SCAAS/AMPAC Calculations of Solvation Energies. J. Comput. Chem 1992, 13, 199–213.

(243). Kearns FL; Warrensford L; Boresch S; Woodcock HL The Good, the Bad, and the Ugly: "HiPen", a New Dataset for Validating (S)QM/MM Free Energy Simulations. Molecules 2019, 24, 681.

(244). Hudson PS; Han K; Woodcock HL; Brooks BR Force Matching as a Stepping Stone to QM/MM CB[8] Host/Guest Binding Free Energies: A SAMPL6 Cautionary Tale. J. Comput. Aided Mol. Des 2018, 32, 983–999. [PubMed: 30276502]

(245). Hudson PS; Boresch S; Rogers DM; Woodcock HL Accelerating QM/MM Free Energy Computations via Intramolecular Force Matching. J. Chem. Theory Comput 2018, 14, 6327–6335. [PubMed: 30300543]

(246). Kearns FL; Hudson PS; Woodcock HL; Boresch S Computing Converged Free Energy Differences between Levels of Theory via Nonequilibrium Work Methods: Challenges and Opportunities. J. Comput. Chem 2017, 38, 1376–1388. [PubMed: 28272811]

(247). Hudson PS; Woodcock HL; Boresch S Use of Nonequilibrium Work Methods to Compute Free Energy Differences Between Molecular Mechanical and Quantum Mechanical Representations of Molecular Systems. J. Phys. Chem. Lett 2015, 6, 4850–4856. [PubMed: 26539729]

(248). König G; Pickard FC; Huang J; Thiel W; MacKerell AD Jr.; Brooks BR; York DM A Comparison of QM/MM Simulations with and without the Drude Oscillator Model Based on Hydration Free Energies of Simple Solutes. Molecules 2018, 23, 2695–2720.

(249). König G; Brooks BR; Thiel W; York DM On the Convergence of Multi-Scale Free Energy Simulations. Mol. Simul 2018, 44, 1062–1081. [PubMed: 30581251]

(250). König G; Brooks BR Correcting for the Free Energy Costs of Bond or Angle Constraints in Molecular Dynamics Simulations. Biochim. Biophys. Acta 2015, 1850, 932–943. [PubMed: 25218695]

(251). König G; Hudson PS; Boresch S; Woodcock HL Multiscale Free Energy Simulations: An Efficient Method for Connecting Classical MD Simulations to QM or QM/MM Free Energies Using Non-Boltzmann Bennett Reweighting Schemes. J. Chem. Theory Comput 2014, 10, 1406–1419. [PubMed: 24803863]

(252). Olsson MA; Ryde U Comparison of QM/MM Methods To Obtain Ligand-Binding Free Energies. J. Chem. Theory Comput 2017, 13, 2245–2253. [PubMed: 28355487]

(253). Shaw KE; Woods CJ; Mulholland AJ Compatibility of Quantum Chemical Methods and Empirical (MM) Water Models in Quantum Mechanics/Molecular Mechanics Liquid Water Simulations. J. Phys. Chem. Lett 2010, 1, 219–223.

(254). Heimdal J; Ryde U Convergence of QM/MM Free-Energy Perturbations Based on Molecular-Mechanics or Semiempirical Simulations. Phys. Chem. Chem. Phys 2012, 14, 12592–12604. [PubMed: 22797613]

(255). Li P; Jia X; Pan X; Shao Y; Mei Y Accelerated Computation of Free Energy Profile at ab Initio Quantum Mechanical/Molecular Mechanics Accuracy via a Semi-Empirical Reference Potential. I. Weighted Thermodynamics Perturbation. J. Chem. Theory Comput 2018, 14, 5583–5596. [PubMed: 30336015]

(256). Wang M; Mei Y; Ryde U Host-Guest Relative Binding Affinities at Density-Functional Theory Level from Semiempirical Molecular Dynamics Simulations. J. Chem. Theory Comput 2019, 15, 2659–2671. [PubMed: 30811192]

(257). Štrajbl M; Hong G; Warshel A Ab Initio QM/MM Simulation with Proper Sampling: “First Principle” Calculations of the Free Energy of the Autodissociation of Water in Aqueous Solution. J. Phys. Chem. B 2002, 106, 13333–13343.

(258). Plotnikov NV; Warshel A Exploring, Refining, and Validating the Paradynamics QM/MM Sampling. J. Phys. Chem. B 2012, 116, 10342–10356. [PubMed: 22853800]

(259). Bentzien J; Muller RP; Florián J; Warshel A Hybrid ab Initio Quantum Mechanics/Molecular Mechanics Calculations of Free Energy Surfaces for Enzymatic Reactions: The Nucleophilic Attack in Subtilisin. J. Phys. Chem. B 1998, 102, 2293–2301.

(260). Kroonblawd MP; Pietrucci F; Marco Saitta A; Goldman N Generating Converged Accurate Free Energy Surfaces for Chemical Reactions with a Force-Matched Semiempirical Model. J. Chem. Theory Comput 2018, 14, 2207–2218. [PubMed: 29543444]

(261). Pinnick ER; Calderon CE; Rusnak AJ; Wang F Achieving Fast Convergence of ab Initio Free Energy Perturbation Calculations with the Adaptive Force-Matching Method. Theor. Chem. Acc 2012, 131, 1146.

(262). Ercolessi F; Adams JB Interatomic Potentials from First-Principles Calculations: The Force-Matching Method. EPL 1994, 26, 583.

(263). Maurer P; Laio A; Hugosson HW; Colombo MC; Rothlisberger U Automated Parametrization of Biomolecular Force Fields from Quantum Mechanics/Molecular Mechanics (QM/MM) Simulations through Force Matching. J. Chem. Theory Comput 2007, 3, 628–639. [PubMed: 26637041]

(264). Izvekov S; Parrinello M; Burnham CJ; Voth GA Effective Force Fields for Condensed Phase Systems from ab Initio Molecular Dynamics Simulation: A New Method for Force-Matching. J. Chem. Phys 2004, 120, 10896. [PubMed: 15268120]

(265). Zhou Y; Pu J Reaction Path Force Matching: A New Strategy of Fitting Specific Reaction Parameters for Semiempirical Methods in Combined QM/ MM Simulations. J. Chem. Theory Comput 2014, 10, 3038. [PubMed: 26588275]

(266). Akin-Ojo O; Song Y; Wang F Developing ab Initio Quality Force Fields from Condensed Phase Quantum-Mechanics/Molecular-Mechanics Calculations through the Adaptive Force Matching Method. J. Chem. Phys 2008, 129, 64108.

(267). Akin-Ojo O; Wang F The Quest for the Best Nonpolarizable Water Model from the Adaptive Force Matching Method. J. Comput. Chem 2011, 32, 453–462. [PubMed: 20730778]

(268). Sampson C; Fox T; Tautermann CS; Woods C; Skylaris C-KA ”Stepping Stone” Approach for Obtaining Quantum Free Energies of Hydration. J. Phys. Chem. B 2015, 119, 7030–7040. [PubMed: 25985723]

(269). Olsson MA; Söderhjelm P; Ryde U Converging Ligand-Binding Free Energies Obtained with Free-Energy Perturbations at the Quantum Mechanical Level. J. Comput. Chem 2016, 37, 1589–1600. [PubMed: 27117350]

(270). Dybeck EC; König G; Brooks BR; Shirts MR Comparison of Methods To Reweight from Classical Molecular Simulations to QM/MM Potentials. J. Chem. Theory Comput 2016, 12, 1466–1480. [PubMed: 26928941]

(271). Genheden S; Ryde U; Söderhjelm P Binding Affinities by Alchemical Perturbation Using QM/MM with a Large QM System and Polarizable MM Model. J. Comput. Chem 2015, 36, 2114–2124. [PubMed: 26280564]

(272). Wesolowski T; Warshel A Ab Initio Free Energy Perturbation Calculations of Solvation Free Energy Using the Frozen Density Functional Approach. J. Phys. Chem 1994, 98, 5183–5187.

(273). Olsson MHM; Hong G; Warshel A Frozen Density Functional Free Energy Simulations of Redox Proteins: Computational Studies of the Reduction Potential of Plastocyanin and Rusticyanin. J. Am. Chem. Soc 2003, 125, 5025–5039. [PubMed: 12708852]

(274). Mori T; Hamers RJ; Pedersen JA; Cui Q Integrated Hamiltonian Sampling: A Simple and Versatile Method for Free Energy Simulations and Conformational Sampling. J. Phys. Chem. B 2014, 118, 8210–8220. [PubMed: 24641518]

(275). Min D; Zheng L; Harris W; Chen M; Lv C; Yang W Practically Efficient QM/MM Alchemical Free Energy Simulations: The Orthogonal Space Random Walk Strategy. J. Chem. Theory Comput 2010, 6, 2253–2266. [PubMed: 26613484]

(276). Plotnikov NV; Kamerlin SCL; Warshel A Paradynamics: An Effective and Reliable Model for Ab Initio QM/MM Free-Energy Calculations and Related Tasks. J. Phys. Chem. B 2011, 115, 7950–7962. [PubMed: 21618985]

(277). Liu P; Kim B; Friesner RA; Berne BJ Replica Exchange with Solute Tempering: A Method for Sampling Biological Systems in Explicit Water. Proc. Natl. Acad. Sci. USA 2005, 102, 13749–13754. [PubMed: 16172406]

(278). Wang L; Friesner RA; Berne B Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2). J. Phys. Chem. B 2011, 115, 9431–9438. [PubMed: 21714551]

(279). Piana S; Robustelli P; Tan D; Chen S; Shaw DE Development of a Force Field for the Simulation of Single-Chain Proteins and Protein-Protein Complexes. J. Chem. Theory Comput 2020, 16, 2494–2507. [PubMed: 31914313]

(280). Izadi S; Anandakrishnan R; Onufriev AV Building Water Models: A Different Approach. J. Phys. Chem. Lett 2014, 5, 3863–3871. [PubMed: 25400877]

(281). Kolář MH; Hobza P Computer Modeling of Halogen Bonds and Other σ-Hole Interactions. Chem. Rev 2016, 116, 5155–5187. [PubMed: 26840433]

(282). Spinn A; Handle PH; Kraml J; Hofer TS; Liedl KR Charge Anisotropy of Nitrogen: Where Chemical Intuition Fails. J. Chem. Theory Comput 2020, 4443–4453.

(283). Buckingham RA The Classical Equation of State of Gaseous Helium, Neon and Argon. P. Roy. Soc. A-Math. Phy 1938, 168, 264–283.

(284). Wade AD; Wang L-P; Huggins DJ Assimilating Radial Distribution Functions To Build Water Models with Improved Structural Properties. J. Chem. Inf. Model 2018, 58, 1766–1778. [PubMed: 30113842]

(285). Migliorati V; Serva A; Terenzio FM; DâAngelo P Development of Lennard-Jones and Buckingham Potentials for Lanthanoid Ions in Water. Inorg. Chem 2017, 56, 6214–6224. [PubMed: 28493693]

(286). Jiao D; Golubkov PA; Darden TA; Ren P Calculation of Protein-Ligand Binding Free Energy by Using a Polarizable Potential. Proc. Natl. Acad. Sci. USA 2008, 105, 6290–6295. [PubMed: 18427113]

(287). Ren P; Wu C; Ponder JW Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules. J. Chem. Theory Comput 2011, 7, DOI: 10.1021/ct200304d.

(288). Lemkul JA; Huang J; Roux B; MacKerell AD Jr. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. Chem. Rev 2016, 116, 4983–5013. [PubMed: 26815602]

(289). Wang J; Cieplak P; Li J; Cai Q; Hsieh M-J; Luo R; Duan Y Development of Polarizable Models for Molecular Mechanical Calculations. 4. van der Waals Parametrization. J. Phys. Chem. B 2012, 116, 7088–7101. [PubMed: 22612331]

(290). Shi Y; Xia Z; Zhang J; Best R; Wu C; Ponder JW; Ren P Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins. J. Chem. Theory Comput 2013, 9, DOI: 10.1021/ct4003702.

(291). Yang Q; Burchett W; Steeno GS; Liu S; Yang M; Mobley DL; Hou X Optimal Designs for Pairwise Calculation: An Application to Free Energy Perturbation in Minimizing Prediction Variability. J. Comp. Chem 2019, DOI: 10.1002/jcc.26095.

(292). Xu H Optimal Measurement Network of Pairwise Differences. J. Chem. Inf. Model 2019, 59, DOI: 10.1021/acs.jcim.9b00528.

(293). Ding X; Vilseck JZ; Brooks III CL Fast Solver for Large Scale Multistate Bennett Acceptance Ratio Equations. J. Chem. Theory Comput 2019, 15, 799–802. [PubMed: 30689377]

(294). Towns J; Cockerill T; Dahan M; Foster I; Gaither K; Grimshaw A; Hazlewood V; Lathrop S; Lifka D; Peterson GD; Roskies R; Scott JR; Wilkins-Diehr N XSEDE: Accelerating Scientific Discovery. Comput. Sci. Eng 2014, 16, 62–74.

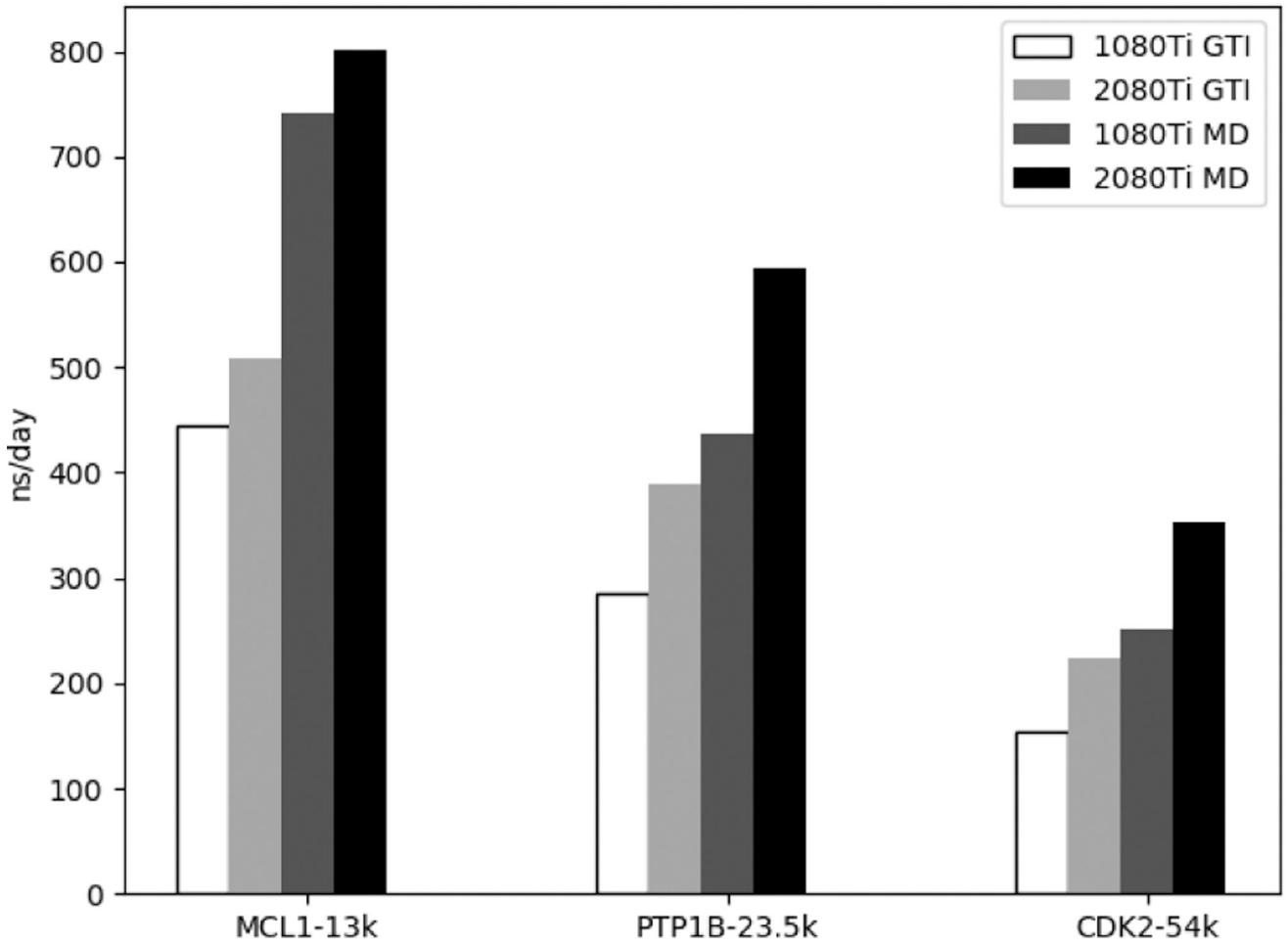



**Figure 1:**

Performance of AMBER20 for standard MD and thermodynamic integration (TI) on GeForce 1080Ti and 2080Ti graphics cards compiled on CUDA 9.1 using a Monte Carlo barostat, Langevin thermostat, and 4 fs time step.

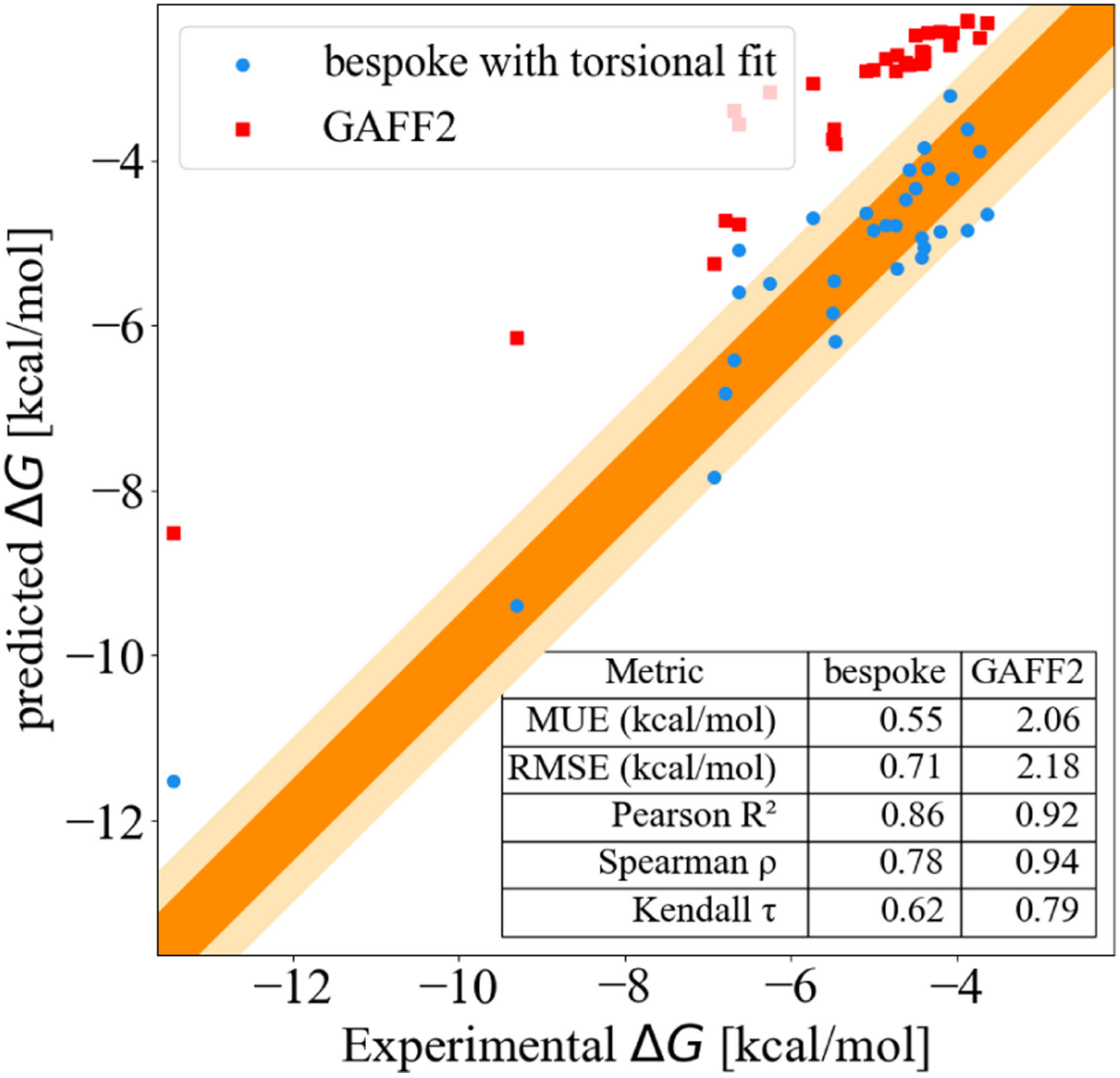



**Figure 2:**

Hydration free energies of molecules containing alcohol functional groups. The bespoke force field uses the same parameters for bond stretch, bond angle, and van der Waals interactions as GAFF2. It derives the partial charges by fitting to the electrostatic potential computed using restricted Hartree-Fock with the 6–31G* basis set.[142] The torsional parameters are then optimized to fit the potential energy surface computed by B3LYP/6–31G** for conformations generated at different torsional values of the rotatable bonds. Refitting the torsional parameters improves the agreement between the predicted and experimental hydration free energies.

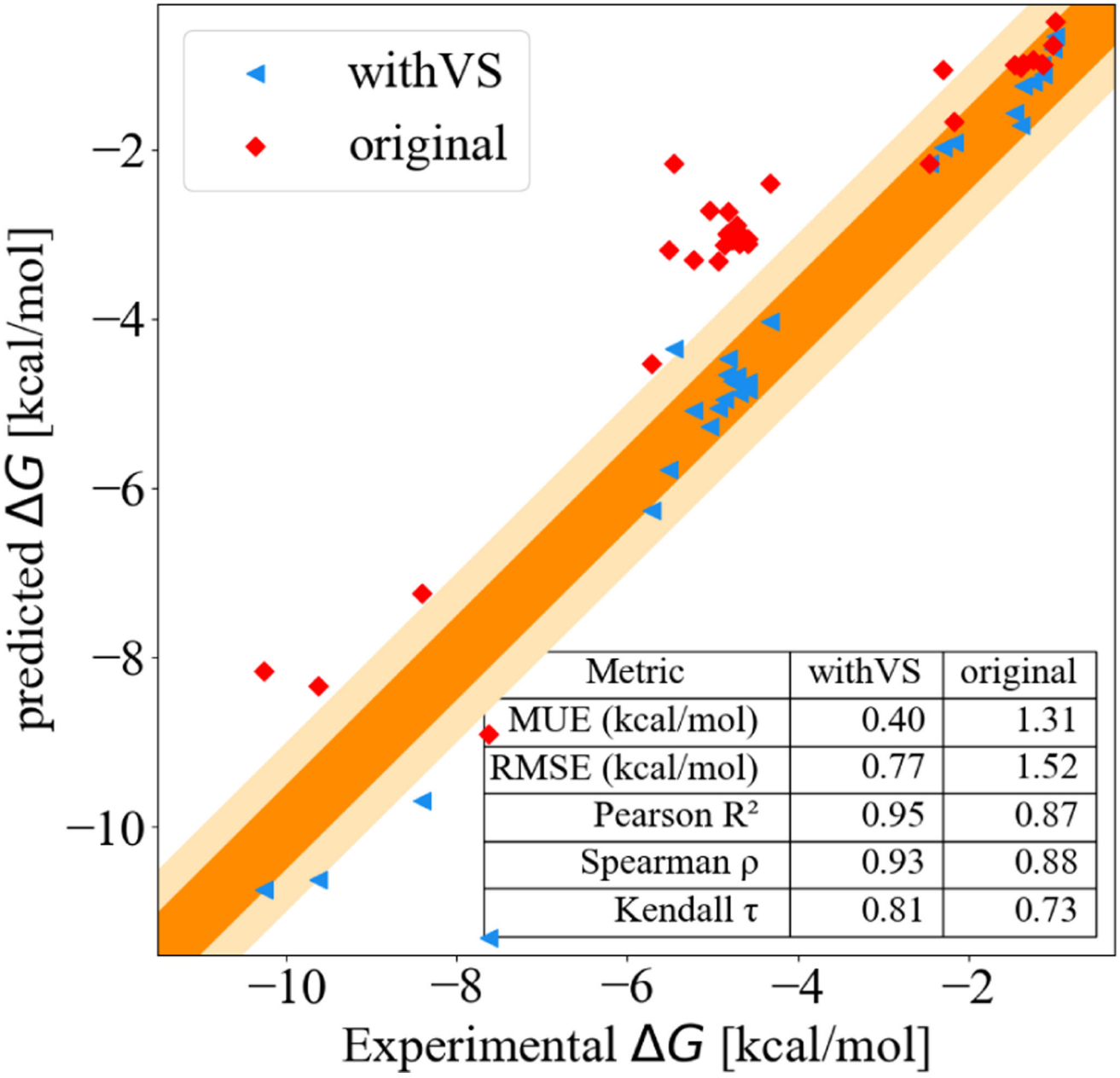



**Figure 3:**
Hydration free energies of molecules containing aromatic halogens and aromatic nitrogens. Including virtual sites on the halogen and nitrogen atoms improves fitting to the electrostatic potential calculated by quantum chemistry and the agreement between the predicted and experimental hydration free energies.

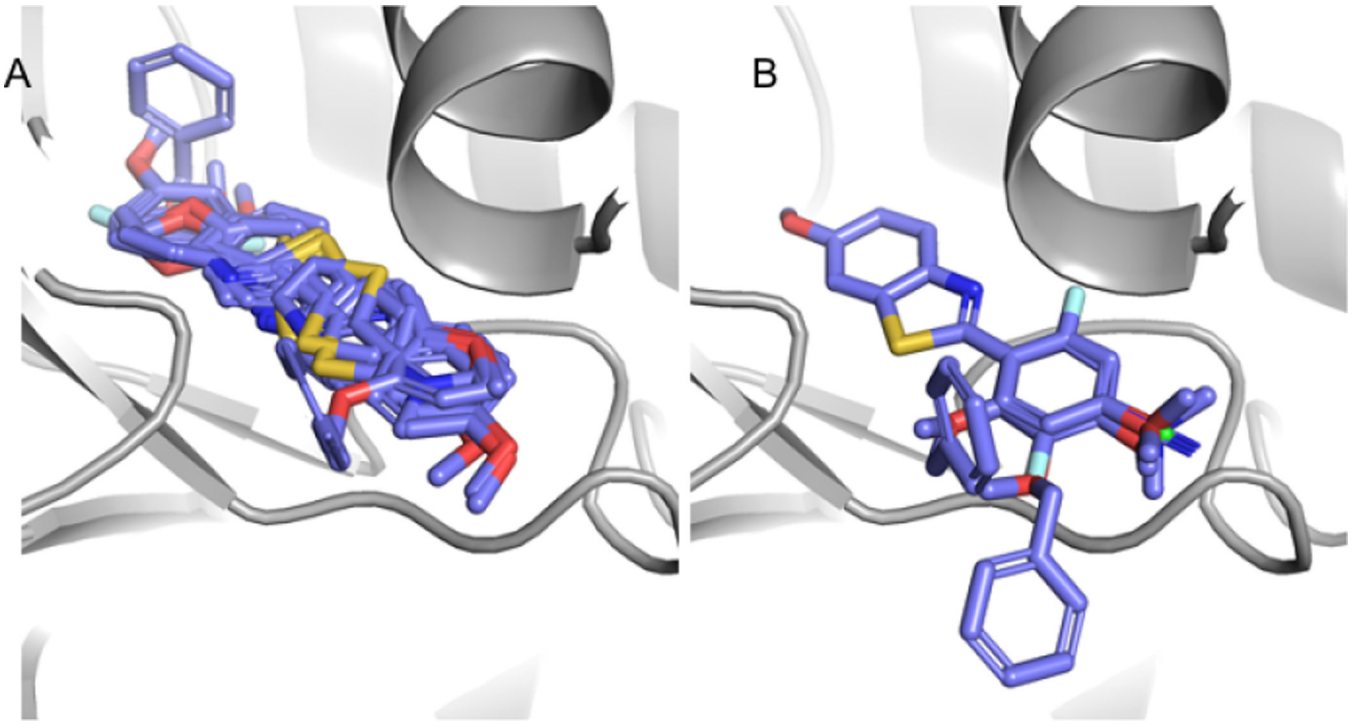



**Figure 4:**
An example of ligand poses (purple carbons) docked A) not using core-restraints and B) using core-restraints. Employing core-constraints ensures that the binding mode is conserved between all of ligands in a congeneric series.

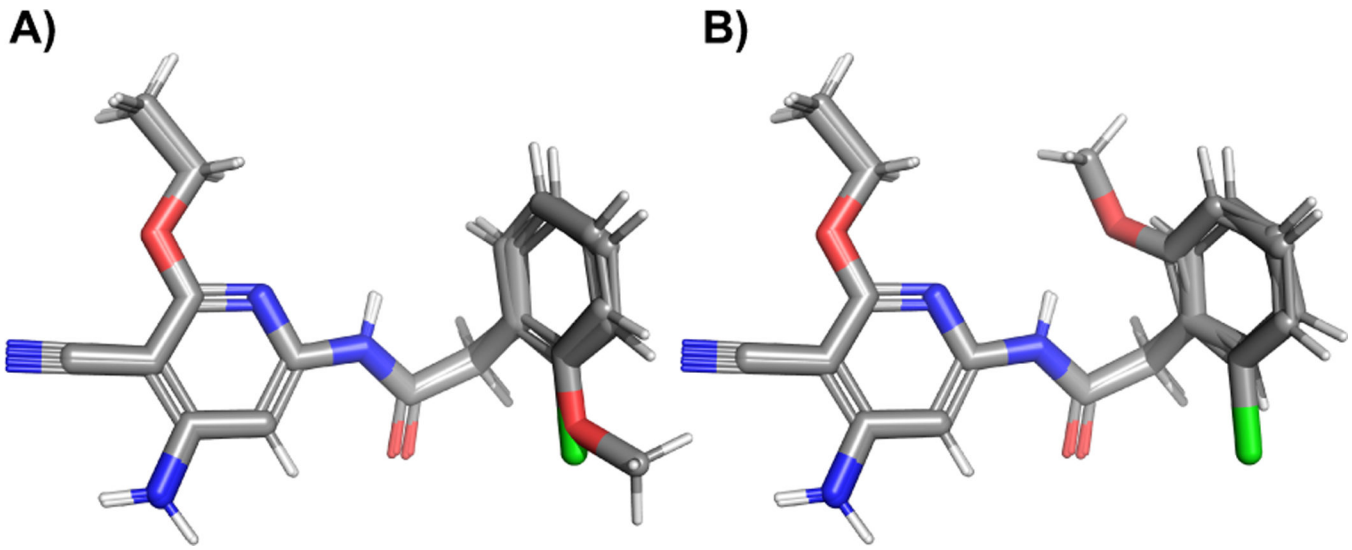



**Figure 5:**
Potential mappings using A) 2D or B) 3D information related to differing ortho substitutions on a terminal phenyl ring.

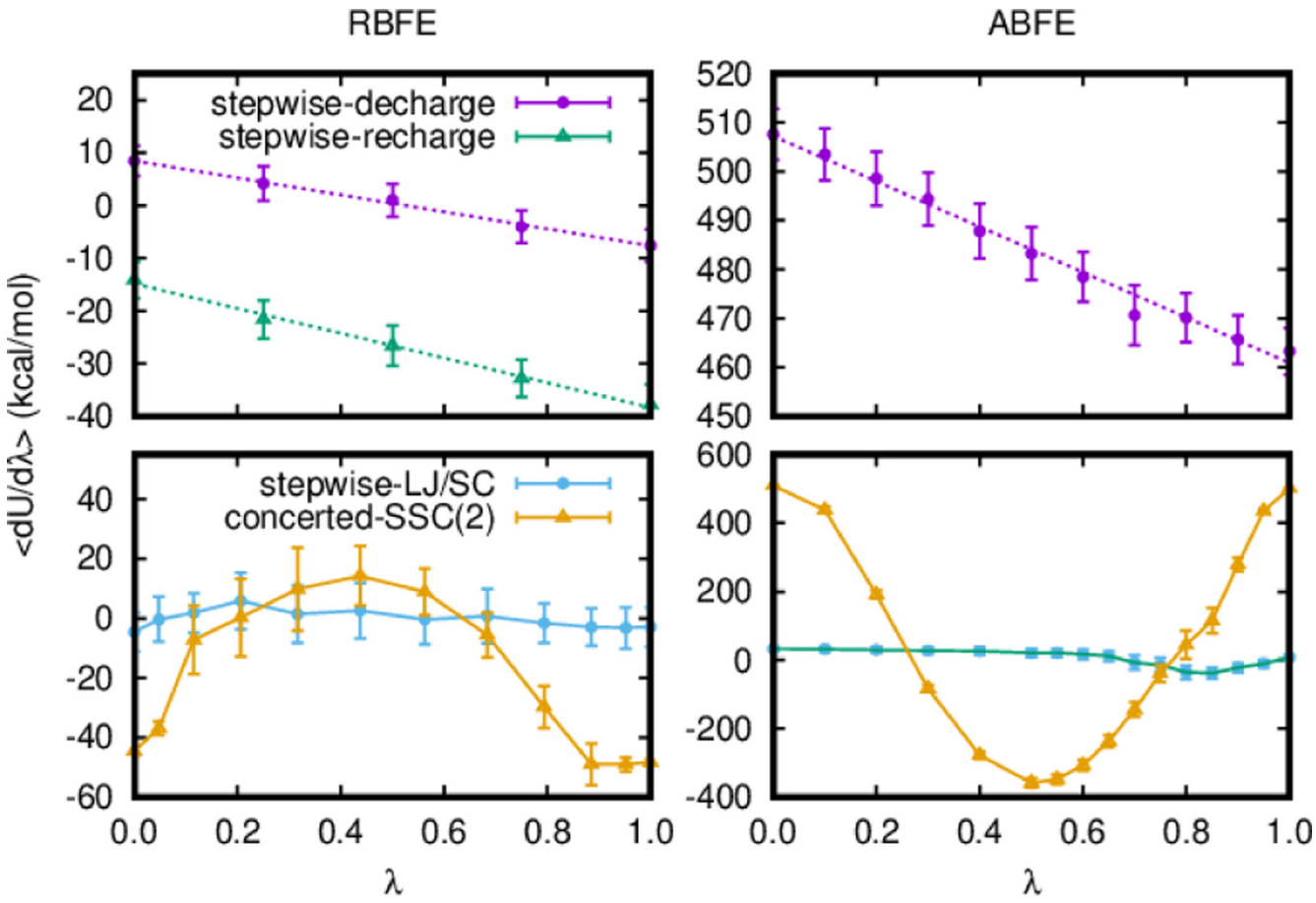



**Figure 6:**

Representative TI integrands for different $\lambda$ schedules used in binding free energy calculations. The shape of the curve is highly dependant on the use of a single step versus multistep protocol. For simple charge changing transformations the curves may have a near linear character (fit dashed lines). The number of atoms being transformed (*e.g.* RBFE versus ABFE) also has a strong effect (left and right columns, respectively). The specific transformations are the Tyk2 ejm-47 (ABFE) and p38 2v→3fhm (RBFE) perturbations from the Wang, *et al.* data set. The specific coupling protocols are as outlined in Section 2.3. Note that decharge and recharge use opposite conventions for direction ($\lambda$=1 is fully coupled for decharge and $\lambda$ = 0 is fully coupled for recharge).

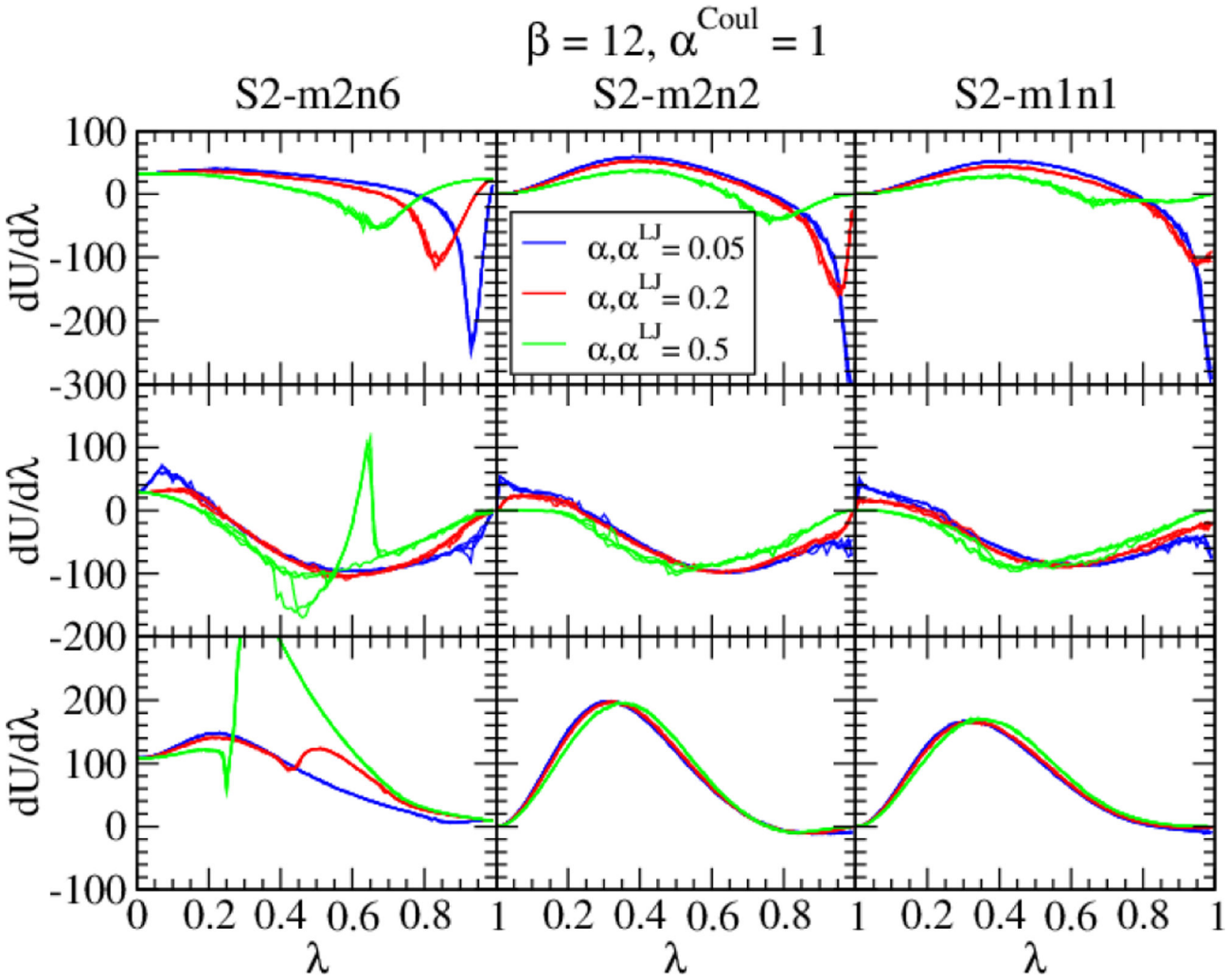



**Figure 7:**

Comparison of $\langle dU/d\lambda \rangle$ curves from (*leftmost column*:) the original SSC(2) scheme with (m=2, n=6), from (*middle column*:) the modified SSC(2) scheme with (m=2, n=2), and from (*rightmost column*:) the modified SSC(2) scheme with (m=1, n=1) (defined in Eq.(20)). The $\beta$ value is 12 Å$^2$ and $\alpha^{\text{Coul}}$ is 1. The molecular systems are *upper row:* the absolute hydration free energy of diphenyltoluene; *middle row*: the relative hydration free energy between the Factor Xa ligand L51h and L51c *lower row*: the absolute hydration free energy of a single $Na^+$ ion.

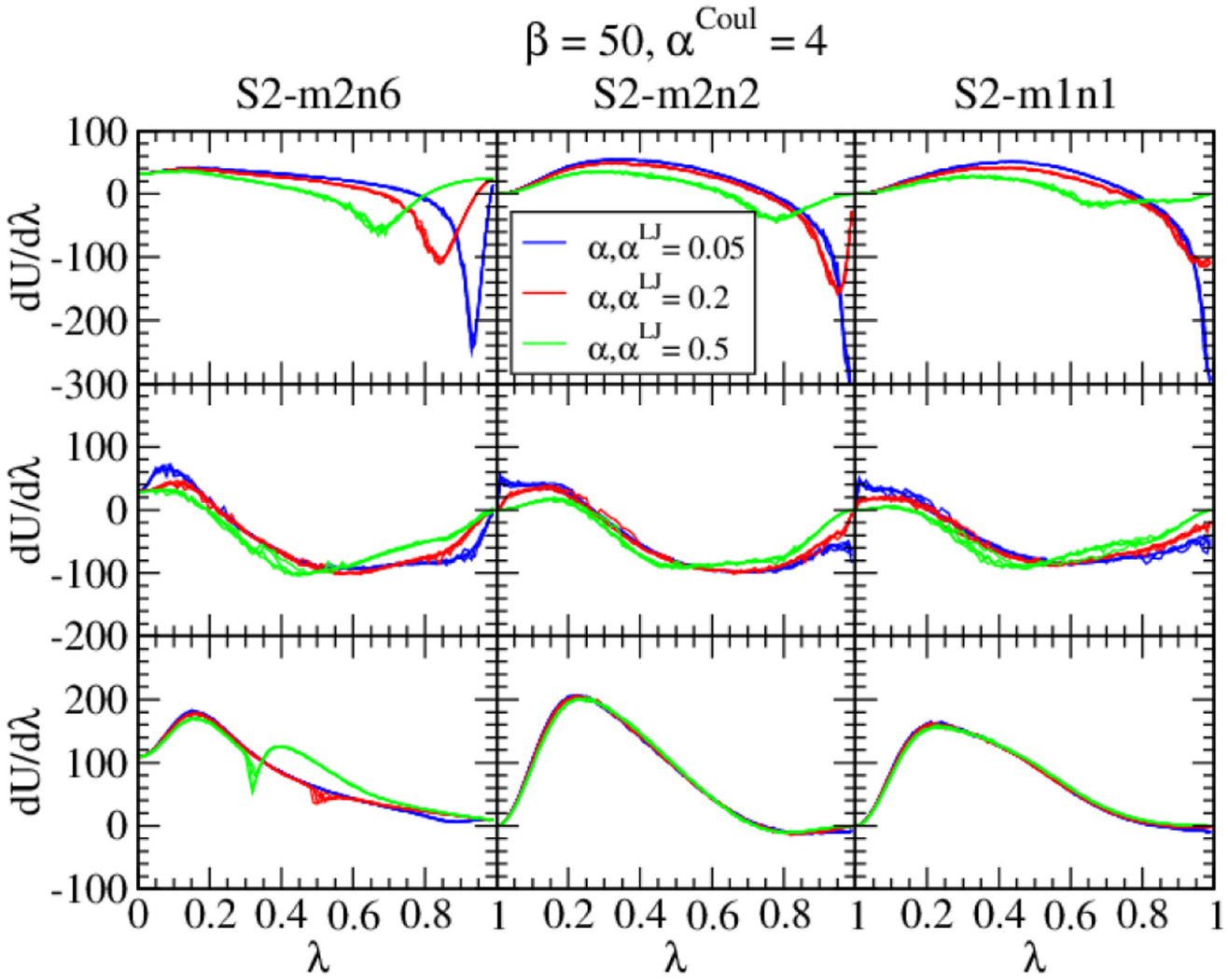



**Figure 8:**

Comparison of $\langle dU/d\lambda \rangle$ curves from (*leftmost column*:) the original SSC(2) scheme with (m=2, n=6), from (*middle column*:) the modified SSC(2) scheme with (m=2, n=2), and from (*rightmost column*:) the modified SSC(2) scheme with (m=1, n=1) (defined in Eq.(20)). The $\beta$ value is 50 Å$^2$ and $\alpha^{Coul}$ is 4. The molecular systems are *upper row*: the absolute hydration free energy of diphenyltoluene; *middle row*: the relative hydration free energy between the Factor Xa ligand L51h and L51c *lower row:* the absolute hydration free energy of a single $Na^+$ ion.

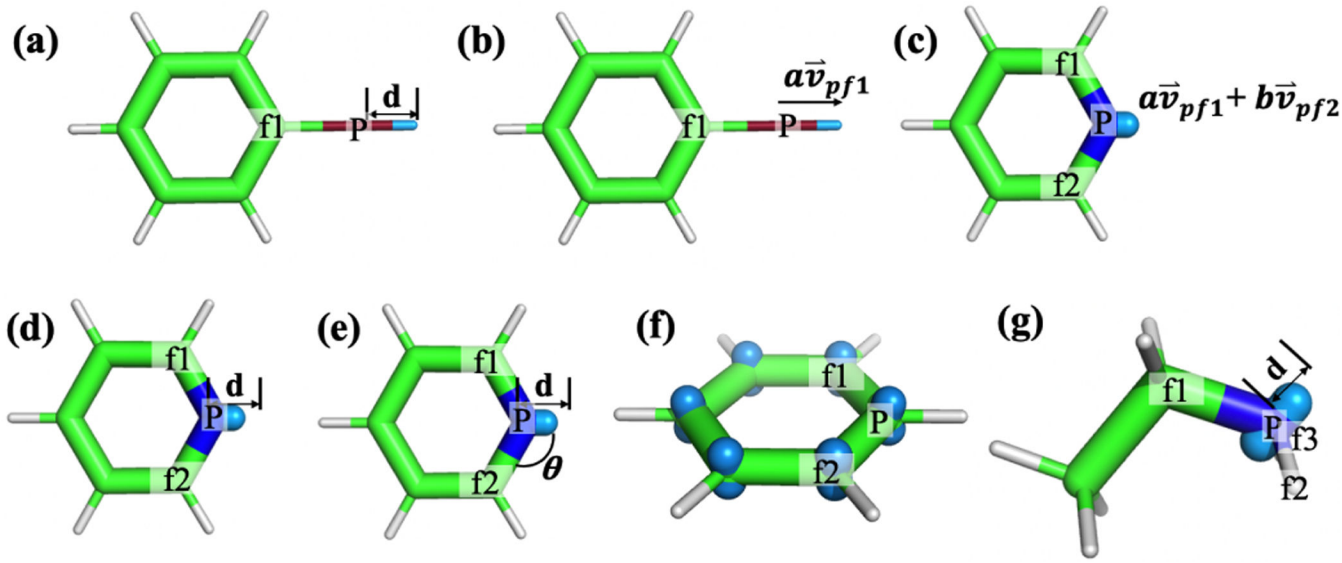



**Figure 9:**

Seven types of virtual sites will be made available in a future release of AMBER. (a) aromatic halogens: fixed-distance VS from 2-atom frame; (b) aromatic halogens: flexible-distance VS from 2-atom frame; (c) aromatic nitrogens: flexible-distance VS from 3-atom frame; (d) aromatic nitrogens: fixed-distance VS from 3-atom frame; (e) aromatic nitrogens: fixed-distance-with-angle VS from 3-atom frame; (f) aromatic carbons: out-of-plane VS from 3-atom frame; (g) amines: in(out-of)-pyramid VS from 4-atom-frame. The virtual sites are shown as cyan beads; P,fn (n=1,2,3) are parent and frame atoms to define virtual sites. The relative positions of the virtual sites are specified by the illustrated geometric parameters.

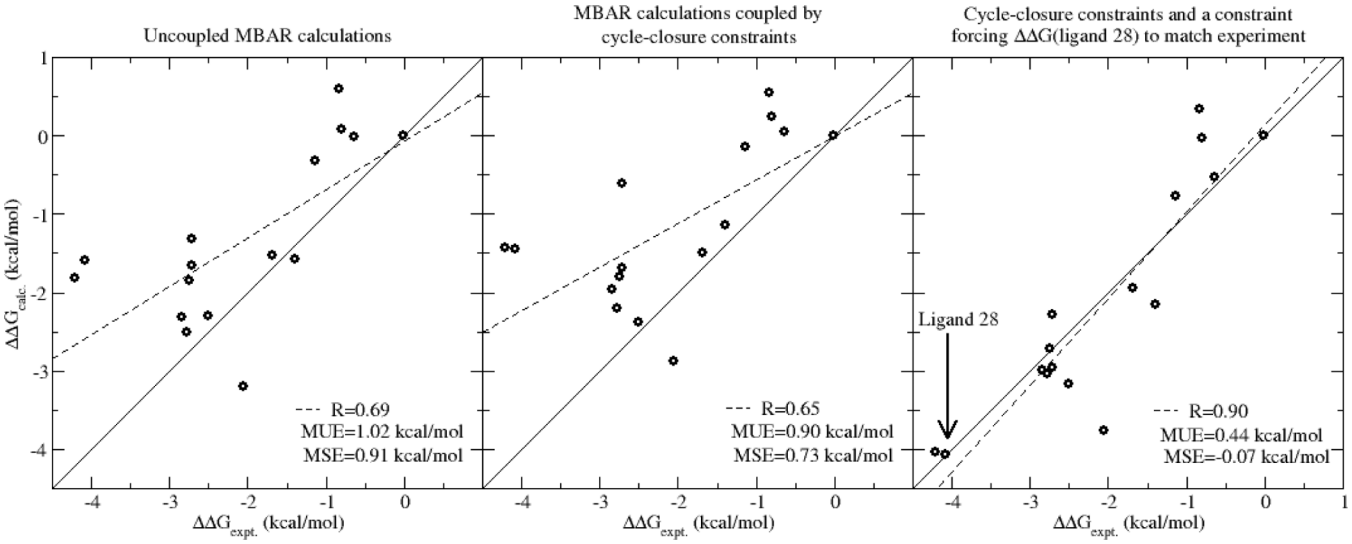



**Figure 10:**

RBFE results for CDK2 (16 ligands). The left pane computes each edge of the RBFE network from independent MBAR optimizations. The center pane simultaneously optimizes all edges in the network, coupling the results through 22 cycle closure constraints. The right pane further includes a constraint that forces the RBFE of “ligand 28” to match experiment.

**Table 1:**

Mean unsigned error (MUE) in kcal/mol for 8 protein targets. Results are presented for the work here and other recent relative binding free energy (RBFE) publications on the same data set.[121] The AMBER20 results are an average of 10 independent runs using the smoothstep softcore SSC(2)[58] and the GAFF2 force field. Results for each individual perturbation, statistical errors across the 10 runs, correlation coefficients, and null model results can be found in Supporting Information.

| Method | BACE | CDK2 | JNK1 | MCL1 | P38 | PTP1B | Thrombin | TYK2 | Mean |
|---|---|---|---|---|---|---|---|---|---|
| AMBER20 (this work) | 0.88 | 0.93 | 0.73 | 1.30 | 0.79 | 0.79 | 0.39 | 0.93 | 0.84 |
| FEP+ (Wang et al.[121]) | 0.84 | 0.91 | 0.78 | 1.16 | 0.80 | 0.89 | 0.76 | 0.75 | 0.86 |
| Cresset (Kuhn et al.[116]) | 0.95 | 0.95 | 0.78 | 1.36 | 1.18 | 1.04 | 0.23 | 0.71 | 0.90 |
| PMX (Gapsys et al.[34]) | 0.84 | 0.68 | 0.80 | 1.23 | 0.77 | 0.90 | 0.77 | 1.01 | 0.88 |