Implicit ligand theory: Rigorous binding free energies and thermodynamic expectations from molecular docking

David D. L. Minh[a)]
Department of Chemistry, Duke University, Durham, North Carolina 27708, USA

A rigorous formalism for estimating noncovalent binding free energies and thermodynamic expectations from calculations in which receptor configurations are sampled independently from the ligand is derived. Due to this separation, receptor configurations only need to be sampled once, facilitating the use of binding free energy calculations in virtual screening. Demonstrative calculations on a host-guest system yield good agreement with previous free energy calculations and isothermal titration calorimetry measurements. Implicit ligand theory provides guidance on how to improve existing molecular docking algorithms and insight into the concepts of induced fit and conformational selection in noncovalent macromolecular recognition. © *2012 American Institute of Physics.* [http://dx.doi.org/10.1063/1.4751284]

I. INTRODUCTION

The goal of molecular docking is to predict the most stable configuration of a noncovalent complex between a ligand and receptor. Based on this configuration, the complex is assigned a score which may be used to approximately rank the binding affinity of one ligand to the receptor versus another. Molecular docking has many potential applications, and has been most prominently applied to the virtual screening[1,2] of chemical libraries to aid the development of pharmaceuticals.

Given the three-dimensional structure of a protein receptor, docking algorithms have proven reasonably adept at sampling stable conformations of small organic ligands in the complex. Unfortunately, current scoring functions perform poorly at predicting binding free energies.[3–5] Hence, docking is typically used to filter a large library of potential ligands to a smaller binder-enriched library that may be pursued experimentally or by more accurate and expensive computational methods.[6–10] Even in this capacity, however, scoring functions are inconsistent, frequently presenting false positives (ligands predicted to bind but actually have weak or no affinity) and false negatives (ligands predicted not to bind but actually have significant affinity). For example, docking programs often have difficulty distinguishing binding compounds from decoys in which the chemical connectivity has been randomized.[3,11] Improved scoring functions would increase the capability to discern binders from non-binders.

The improvement of scoring functions, however, has been hindered by the lack of a rigorous formalism for obtaining binding free energies from molecular docking. While molecular docking calculations are usually performed with a rigid receptor, existing formalisms for binding free energies require a flexible receptor. Here, I derive a formalism, *implicit ligand theory*, for estimating binding free energies and thermodynamic expectations based on docking ligands to rigid-receptor structures. I also describe practical aspects of

statistical estimation, present example calculations, and discuss how physics-based (opposed to empirical or knowledge-based) docking algorithms (see Ref. 12) may be modified to exploit it. Beyond molecular docking, implicit ligand theory provides insight into the concepts of induced fit and conformational selection in noncovalent macromolecular recognition.

II. THEORY

The standard binding free energy, the free energy of a noncovalent association between a receptor R and ligand L to form a complex RL, $R + L \rightleftharpoons RL$, is

$$\Delta G^\circ = -\beta^{-1} \ln \left(\frac{C^\circ C_{RL}}{C_R C_L} \right), \qquad (1)$$

where $\beta = (k_B T)^{-1}$ is the inverse of Boltzmann's constant, k_B, times the temperature in Kelvin, T, C° is the standard concentration (typically 1 M), and C_X is the equilibrium concentration of species $X \in \{R, L, RL\}$.[13]

Statistical thermodynamics relates the standard binding free energy to a ratio of configurational partition functions[14]

$$\Delta G^\circ = -\beta^{-1} \ln \left(\frac{Z_{RL,N} Z_N}{Z_{R,N} Z_{L,N}} \frac{C^\circ}{8\pi^2} \right), \qquad (2)$$

$$Z_{RL,N} = \int I_\xi e^{-\beta U(r_{RL}, r_S)} dr_{RL} dr_S, \qquad (3)$$

$$Z_{Y,N} = \int e^{-\beta U(r_Y, r_S)} dr_Y dr_S, \qquad (4)$$

$$Z_N = \int e^{-\beta U(r_S)} dr_S, \qquad (5)$$

in which symmetry numbers and a small pressure-volume term have been omitted from Eq. (2). $Z_{RL,N}$ and $Z_{Y,N}$ are configurational partition functions of the complex and of the species $Y \in \{R, L\}$, respectively, in N molecules of solvent.

---

a)Electronic mail: dm225@duke.edu.

The potential energy $U(r_X, r_S)$ depends on r_X, the *internal* coordinates of the receptor, ligand, or both in complex (the external degrees of freedom have been analytically integrated), and r_S, the coordinates of N molecules of solvent. The complex coordinates r_{RL} may be decomposed into the internal coordinates of the receptor, r_R, and of the ligand, r_L, and six degrees of freedom describing their relative translation and rotation, ξ_L. For simplicity, Jacobians for the transformation from Cartesian coordinates to a system with separated internal and external degrees of freedom are not shown in Eqs. (3) and (4). In $Z_{RL,N}$, the indicator function $I_\xi \equiv I(\xi_L)$ takes values between 0 and 1 and determines whether the receptor and ligand are complexed or not. For tight-binding complexes, the binding free energy is insensitive to the precise definition of I_ξ.[14]

A. Implicit solvent theory

The configurational integrals in Eq. (2) may be expressed in a formally equivalent but simpler form using implicit-solvent theory.[14] In implicit solvent theory, the *interaction energy* is defined as $\psi(r_X, r_S) = U(r_X, r_S) - U(r_X) - U(r_S)$, where $U(r_X)$ is the potential energy of species X by itself and $U(r_S)$ the potential energy of the solvent by itself. By integrating the configurational partition functions over r_S, we may define the ratios

$$Z_{RL} \equiv \frac{Z_{RL,N}}{Z_N} = \int I_\xi e^{-\beta[U(r_{RL}) + W(r_{RL})]} dr_{RL}, \quad (6)$$

$$Z_Y \equiv \frac{Z_{Y,N}}{Z_N} = \int e^{-\beta[U(r_Y) + W(r_Y)]} dr_Y, \quad (7)$$

where

$$W(r_X) = -\beta^{-1} \ln \left(\frac{\int e^{-\beta \psi(r_X, r_S)} e^{-\beta U(r_S)} dr_S}{\int e^{-\beta U(r_S)} dr_S} \right) \quad (8)$$

is a potential of mean force (PMF) that can be interpreted as the constant-pressure reversible work of transferring the species X from the gas phase into the solvent. In biomolecular modeling, $W(r_X)$ is frequently estimated as the sum of an electrostatic term from the Poisson-Boltzmann (PB) equation[15] (or the generalized Born approximation[16]), and a non-electrostatic term, which to a first approximation is proportional to the molecular surface area.

In terms of implicit solvent configurational integrals, the standard binding free energy is

$$\Delta G^\circ = -\beta^{-1} \ln \left(\frac{Z_{RL}}{Z_R Z_L} \frac{C^\circ}{8\pi^2} \right). \quad (9)$$

As most implicit solvent models fail to account for specific interactions, such as hydrogen bonding, that can have important structural and energetic consequences, binding free energy calculations in implicit solvent are generally expected to be less accurate than those in explicit solvent.[17] Nevertheless, binding free energy calculations in implicit solvent have yielded promising agreement with experimental results (e.g., Refs. 18–21).

B. Implicit ligand theory

The development of implicit ligand theory is very similar to that of implicit solvent theory. It involves defining the *effective potential* as $\mathcal{U}(r_X) = U(r_X) + W(r_X)$, the *effective interaction energy* as $\Psi(r_{RL}) = \mathcal{U}(r_{RL}) - \mathcal{U}(r_R) - \mathcal{U}(r_L)$, and

$$B(r_R) = -\beta^{-1} \ln \left(\frac{\int I_\xi e^{-\beta \Psi(r_{RL})} e^{-\beta \mathcal{U}(r_L)} dr_L d\xi_L}{\int I_\xi e^{-\beta \mathcal{U}(r_L)} dr_L d\xi_L} \right)$$

$$\equiv -\beta^{-1} \ln \langle e^{-\beta \Psi} \rangle_{L,I}^{r_L, \xi_L}, \quad (10)$$

which is a potential of mean force that will subsequently be referred to as the *binding PMF*. Throughout this paper, angled brackets $\langle \ldots \rangle_{X,\ldots}^r$ will be used to denote an ensemble average over the coordinates r listed in the superscript with respect to the density proportional to $q_{X,\ldots}$, where X describes the coordinates in the effective potential $\mathcal{U}(r_X)$, and $\ldots$ are labels. Here, $q_{L,I}(r_L, \xi_L) = I_\xi e^{-\beta \mathcal{U}(r_L)}$. Within angled brackets, I will use a shorthand notation in which functions implicitly depend on coordinates, e.g., $\Psi \equiv \Psi(r_{RL})$.

In terms of the binding PMF, Eq. (9) may be written as

$$\Delta G^\circ = -\beta^{-1} \ln \left(\frac{\int I_\xi e^{-\beta \mathcal{U}(r_{RL})} dr_{RL}}{\int e^{-\beta \mathcal{U}(r_R)} dr_R \int e^{-\beta \mathcal{U}(r_L)} dr_L} \frac{C^\circ}{8\pi^2} \right)$$

$$= -\beta^{-1} \ln \left(\frac{\int I_\xi e^{-\beta[\mathcal{U}(r_R) + \Psi(r_{RL}) + \mathcal{U}(r_L)]} dr_{RL}}{\int e^{-\beta \mathcal{U}(r_R)} dr_R \int e^{-\beta \mathcal{U}(r_L)} dr_L} \frac{C^\circ}{8\pi^2} \right)$$

$$= -\beta^{-1} \ln \left(\frac{\int e^{-\beta[B(r_R) + \mathcal{U}(r_R)]} dr_R}{\int e^{-\beta \mathcal{U}(r_R)} dr_R} \frac{\Omega C^\circ}{8\pi^2} \right)$$

$$\equiv -\beta^{-1} \ln \langle e^{-\beta B} \rangle_R^{r_R} + \Delta G_\xi, \quad (11)$$

where $\Omega = \int I_\xi d\xi_L$ (which may be analytically tractable) is the binding site volume, $\Delta G_\xi = -\beta^{-1} \ln \left(\frac{\Omega C^\circ}{8\pi^2} \right)$ is the free energy of confining the ligand external degrees of freedom to the binding site, and $q_R(r_R) = e^{-\beta \mathcal{U}(r_R)}$. Equations (10) and (11) are the central theoretical results of this paper.

Implicit ligand theory provides a rigorous framework for binding free energies that separates the sampling of receptor and ligand configurations. In Eq. (11), the receptor probability density is independent of any ligand configuration. Likewise, the probability density of ligand internal coordinates in Eq. (10) is independent from the receptor configuration. In practice, however, sampling from this ligand distribution may lead to slow convergence (this point will later be discussed in greater detail). The primary benefit of implicit ligand theory is that the computationally expensive step of sampling receptor configurations only needs to be performed once. Predicting binding free energies for a chemical library is then limited by the much faster process of sampling ligand conformations.

C. Thermodynamic expectations

In addition to estimating the binding free energy, implicit ligand theory may also be used to estimate expected values of observables in the bound ensemble. Observables may include, for example, the mean potential energy, interaction energy, or distance between a ligand and receptor atom. Towards this end, it is useful to define a rigid-receptor expectation of an

observable $O(r_{RL})$, weighted by the interaction energy,

$$\Theta(r_R) = \frac{\int I_\xi O(r_{RL}) e^{-\beta \Psi(r_{RL})} e^{-\beta \mathcal{U}(r_L)} dr_L d\xi_L}{\int I_\xi e^{-\beta \mathcal{U}(r_L)} dr_L d\xi_L}$$

$$\equiv \langle O e^{-\beta \Psi} \rangle_{L,I}^{r_L, \xi_L}. \quad (12)$$

If the observable is solely a function of the receptor configuration, then $\Theta(r_R)$ reduces to $O(r_R) e^{-\beta B(r_R)}$.

In terms of Eqs. (10) and (12), the expectation of $O(r_{RL})$ with respect to the density proportional to $q_{RL,I}(r_{RL}) = I_\xi e^{-\beta \mathcal{U}(r_{RL})}$ is

$$\langle O \rangle_{RL,I}^{r_{RL}} \equiv \frac{\int I_\xi O(r_{RL}) e^{-\beta \mathcal{U}(r_{RL})} dr_{RL}}{\int I_\xi e^{-\beta \mathcal{U}(r_{RL})} dr_{RL}}$$

$$= \frac{\int I_\xi O(r_{RL}) e^{-\beta [\mathcal{U}(r_R) + \Psi(r_{RL}) + \mathcal{U}(r_L)]} dr_{RL}}{\int I_\xi e^{-\beta [\mathcal{U}(r_R) + \Psi(r_{RL}) + \mathcal{U}(r_L)]} dr_{RL}}$$

$$= \frac{\int \Theta(r_R) e^{-\beta \mathcal{U}(r_R)} dr_R}{\int e^{-\beta [B(r_R) + \mathcal{U}(r_R)]} dr_R}$$

$$\equiv \frac{\langle \Theta \rangle_R^{r_R}}{\langle e^{-\beta B} \rangle_R^{r_R}} = \langle \Theta \rangle_R^{r_R} e^{\beta [\Delta G^\circ - \Delta G_\xi]}. \quad (13)$$

Equations (12) and (13) significantly generalize *implicit ligand sampling*,[22] a method to estimate the potential of mean force for the ligand center of mass. The results of Cohen *et al.*[22] may be obtained by choosing the observable as a Dirac delta function for the ligand center of mass, taking a natural logarithm, and multiplying by β^{-1}. Cohen *et al.*[22] applied implicit ligand sampling to study gas migration pathways in myoglobin, but the possibility of estimating other observables and binding free energies has not been previously recognized.

III. ESTIMATION

Applying implicit ligand theory to predicting binding free energies involves three steps:

1. Sampling receptor configurations.
2. Estimating the binding PMF, $B(r_R)$, for each receptor configuration.
3. Estimating ΔG° from $B(r_R)$ estimates.

In this section, I present several ways, roughly in order of increasing complexity, that these steps may be accomplished. A variant of one approach will be demonstrated later in the paper.

A. Receptor configurations

Receptor configurations can be drawn from $q_R(r_R)$, any (possibly unnormalized) distribution $q_{R,w}(r_R)$ on the same support as $q_R(r_R)$ and for which $w(r_R) = q_R(r_R)/q_{R,w}(r_R)$ may be calculated, or from multiple distributions satisfying these conditions. Regardless of the sampling method, however, convergence of free energy estimates requires representative sampling of both the bound and unbound receptor configuration space. A particularly straightforward protocol is to sample from the distribution proportional to $q_R(r_R) = e^{-\beta \mathcal{U}(r_R)}$; one conducts a molecular dynamics (MD) simulation in the implicit solvent used for $W(r_R)$, collecting snap-

shots at evenly spaced intervals that are longer than the statistical correlation time. This protocol may be satisfactory if receptor fluctuations are minimal and the ligand does not significantly perturb the receptor configurational ensemble.

For a receptor that undergoes larger structural fluctuations, sampling from multiple energetic minima may be facilitated by applying an external biasing potential (e.g., a harmonic bias) on one or more order parameters. If it is known that a ligand significantly perturbs the receptor configurational ensemble, it can be useful to introduce multiple *alchemical* intermediates into a simulation. Alchemical calculations may involve a coupling parameter λ, defined such that the two groups (e.g., the receptor and ligand) are non-interacting at $\lambda = 0$ and fully interacting with $\lambda = 1$. Simulations are conducted with λ at these end points and at multiple values in between. Sampling in each stage may be enhanced by Hamiltonian replica exchange (e.g., Jiang *et al.*,[23] Gallicchio *et al.*,[21] Gallicchio and Levy[24]), which entails stochastically swapping the coordinates of different simulations with a probability that preserves the Boltzmann distribution. Receptor configurations obtained through a flexible-receptor Hamiltonian replica exchange with a single ligand may subsequently be used for implicit ligand free energy calculations with other ligands in the chemical library.

As a caveat, implicit ligand theory does not provide a formal justification for docking to multiple experimentally determined structures (e.g., Ref. 25) or any other set of structures in which $w(r_R)$ is unknown (e.g., homology modeling or flexible docking). One potential way to use information about multiple structures is to conduct multiple MD simulations with external potentials biased towards one or more of the structures. To facilitate later analysis, the external potentials should be set up to promote overlap in the configuration space of different simulations.

B. Estimating a binding PMF

The binding PMF $B(r_R)$ may be expressed in terms of a ratio of partition functions,

$$B(r_R) = -\beta^{-1} \ln \left(\frac{\int I_\xi e^{-\beta \mathcal{U}(r_{RL})} dr_L d\xi_L}{\int I_\xi e^{-\beta [\mathcal{U}(r_L) + \mathcal{U}(r_R)]} dr_L d\xi_L} \right), \quad (14)$$

which clarifies that $B(r_R)$ is a special type of free energy difference in which the receptor configuration r_R is rigid. Thus, $B(r_R)$ may be calculated using any one of many available methods to estimate free energy differences,[26] including free energy perturbation (FEP),[27] thermodynamic integration (TI),[28] and the Bennett acceptance ratio (BAR).[29] While formally equivalent, free energy methods can have dramatically different convergence properties.

Based on the form of Eq. (10), the most straightforward estimation protocol is FEP. One can, for example, draw ligand configurations from the distribution proportional to $q_L(r_L, \xi_L) = e^{-\beta \mathcal{U}(r_L)}$ by conducting a MD simulation of the ligand in the appropriate implicit solvent and collecting snapshots at sufficiently long intervals. Because q_L is independent of ξ_L, the external degrees of freedom sampled from the simulation may be replaced by a new ξ_L sampled from the

distribution proportional to $q_{\xi,I} = I_\xi$. The expectation in Eq. (11) may then be estimated by the sample mean

$$\hat{B}(r_R) = -\beta^{-1} \ln \frac{1}{N} \sum_{n=1}^{N} e^{-\beta \Psi(r_{RL,n})}, \qquad (15)$$

where $r_{RL,n}$ is the nth of N samples of the complex. Throughout this paper, $\hat{A}$ will denote a statistical estimator—an equation used to calculate a quantity based on sampled data.

In exponential averages such as Eq. (15), a small subset of samples may contribute a large portion of the sum. The limiting case of an individual important sample inspires the severe *dominant state approximation*, in which a single value of $\Psi(r_{RL})$ is used to estimate $B(r_R)$. Exponential averages may also be estimated via a cumulant expansion,[27] here shown for Eq. (10) to the fourth order,

$$B(r_R) \approx \langle \Psi \rangle_{L,I}^{r_L,\xi_L} - \frac{\beta}{2!} \langle \xi \Psi^2 \rangle_{L,I}^{r_L,\xi_L} + \frac{\beta^2}{3!} \langle \xi \Psi^3 \rangle_{L,I}^{r_L,\xi_L}$$
$$- \frac{\beta^3}{4!} \big[\langle \xi \Psi^4 \rangle_{L,I}^{r_L,\xi_L} - 3 \big(\langle \xi \Psi^2 \rangle_{L,I}^{r_L,\xi_L} \big)^2 \big], \qquad (16)$$

where $\xi \Psi = \Psi(r_{RL}) - \langle \Psi \rangle_{L,I}^{r_L,\xi_L}$. Each expectation in the cumulant expansion may be estimated by the sample mean.

While formally correct, this approach to ligand sampling can converge slowly if most ligand configurations placed in the binding site have overlapping atoms and high values of $\Psi(r_{RL})$. One potential solution to this problem is to sample the external degrees of freedom from a distribution biased towards energetically favorable orientations by a confining potential $U_c(\xi_L)$. Multiplying and dividing Eq. (10) by $\Omega_c = \int I_\xi e^{-\beta U_c(\xi_L)} d\xi_L$ and the integrand in the numerator by $e^{-\beta U_c(\xi_L)}$ leads to

$$B(r_R) = -\beta^{-1} \ln \langle e^{-\beta[\Psi - U_c]} \rangle_{L,Ic}^{r_L,\xi_L} - \beta^{-1} \ln \left(\frac{\Omega_c}{\Omega} \right), \quad (17)$$

where $q_{L,Ic} = I_\xi e^{-\beta[\mathcal{U}(r_L) + U_c(\xi_L)]}$. Good choices for $U_c(\xi_L)$, which may be ascertained from existing molecular docking algorithms (as will be discussed later in the paper), will favor the sampling of poses with low $\Psi(r_{RL})$.

Alternatively, the binding PMF may be calculated using the inverse form of Eq. (10),

$$B(r_R) = \beta^{-1} \ln \left(\frac{\int I_\xi e^{\beta \Psi(r_{RL})} e^{-\beta \mathcal{U}(r_{RL})} dr_L d\xi_L}{\int I_\xi e^{-\beta \mathcal{U}(r_{RL})} dr_L d\xi_L} \right)$$
$$= \beta^{-1} \ln \langle e^{\beta \Psi} \rangle_{RL,I}^{r_L,\xi_L}. \qquad (18)$$

Ligand configurations from the distribution proportional to $q_{RL,I}(r_L, \xi_L) = I_\xi e^{-\beta \mathcal{U}(r_{RL})}$ may be sampled, for example, from an implicit-solvent MD simulation in which the receptor is held rigid and the ligand is allowed to move, and the expectation estimated using the sample mean estimator.

This straightforward procedure is also problematic because of the rarity of sampling configurations in which the ligand is separated from the receptor or in which they overlap. While these configurations are insignificant in the conformational ensemble in which receptor and ligand are fully interacting, they are relevant to the ensemble of noninteracting ligand and receptor, and the convergence of free energy differences requires phase space overlap between

adjacent thermodynamic states.[26] The phase space overlap problem may also be alleviated by calculating the free energy difference with a reference state in which the external degrees of freedom are confined,

$$B(r_R) = \beta^{-1} \ln \langle e^{\beta[\Psi - U_c]} \rangle_{RL,I}^{r_L,\xi_L} - \beta^{-1} \ln \frac{\Omega_c}{\Omega}. \qquad (19)$$

The binding PMF may be estimated from the same samples as with Eq. (18), and will be more accurate the more closely $e^{-\beta U_c(\xi_L)}$ resembles the distribution of ξ_L in the complex.

As discussed, phase space overlap problems are often resolved by introducing multiple alchemical stages into a calculation, and sampling may be enhanced by Hamiltonian replica exchange. With multiple stages, the total free energy difference between states with $\lambda = 0$ and $\lambda = 1$ is the sum of free energy differences between adjacent stages, each of which may be estimated by FEP,[27] TI,[28] or BAR.[29] Alternatively, the total free energy difference may be estimated by the multistate Bennett acceptance ratio (MBAR).[30]

C. Estimating the binding free energy

Once $\hat{B}(r_R)$ is evaluated for each receptor configuration, the binding free energy may be calculated by estimating an ensemble average. The appropriate method for estimating ΔG° depends on how the receptor configurations r_R are sampled. If they are drawn from the distribution $q_R(r_R)$, then the expectation in Eq. (11) may be estimated by the sample mean

$$\Delta \hat{G}^\circ = -\beta^{-1} \ln \frac{1}{N} \sum_{n=1}^{N} e^{-\beta \hat{B}(r_{R,n})} + \Delta G_\xi, \qquad (20)$$

in which $\hat{B}(r_{R,n})$ is the estimated binding PMF for the nth of N receptor configurations. Because the implicit-ligand expression for the binding free energy, Eq. (11), has the same form as Eq. (10), the dominant state approximation and cumulant expansion may also be applied.

If receptor configurations are drawn from a biased distribution, the importance sampling identity,

$$\langle O \rangle_T = \frac{\int O(r) q_T(r) dr}{\int q_T(r) dr}$$
$$= \frac{\int O(r) w(r) q_S(r) dr}{\int w(r) q_S(r) dr} = \frac{\langle w O \rangle_S}{\langle w \rangle_S}, \qquad (21)$$

may be applied. In this generic expression, $w(r) = q_T(r)/q_S(r)$ is a ratio of unnormalized densities $q_T(r)$ for the target distribution and $q_S(r)$ for the sampling distribution. Using the sample mean estimator and importance sampling identity for the expectation in Eq. (11) leads to

$$\Delta \hat{G}^\circ = -\beta^{-1} \ln \frac{\sum_{n=1}^{N} w(r_{R,n}) e^{-\beta \hat{B}(r_{R,n})}}{\sum_{n=1}^{N} w(r_{R,n})} + \Delta G_\xi. \quad (22)$$

If receptor configurations are drawn from multiple biased distributions, then the expectation may be estimated using MBAR.[30]

D. Thermodynamic expectations

Thermodynamic expectations may be estimated from the same data as the binding free energy. The appropriate estimator for $\Theta(r_R)$ will depend on how the ligand configurations were sampled. Once $\Theta(r_R)$ is estimated for every sampled receptor configuration, the appropriate estimator for the expectation in Eq. (13) similarly depends on how the receptor configurations were sampled. In the simplest case for $\Theta(r_R)$, if ligand configurations are sampled from $q_{\xi, I}$, then $\Theta(r_R)$ may be estimated by a sample mean. In other cases, $\Theta(r_R)$ and $\langle O \rangle_{RL, I}^{r_{RL}}$ may be estimated using importance sampling, MBAR,[30] or a combination thereof.

IV. DEMONSTRATION

As a demonstration, implicit ligand theory calculations were performed to estimate the standard binding free energy of various ligands to Cucurbit[7]uril (CB[7]) in water. The binding of CB[7] to a number of ferrocenes, adamantanes, and bicyclooctanes has been well characterized by both isothermal calorimetry and second-generation mining minima (M2)[18, 19] free energy calculations.[31, 32] Receptor configurations were sampled by molecular dynamics, binding PMFs estimated with a multi-stage alchemical calculation and MBAR,[30] and the binding free energy calculated using Eq. (20) or the dominant state approximation.

A. Methods

Molecular dynamics simulations at 300 K were performed with a slightly modified[33] compilation of NAMD[34] version 2.9. When appropriate, CB[7] was fixed using the fixedAtoms parameter. The "commercial" force field parameters and topologies from Moghaddam *et al.*[32] were used for both CB[7] and its ligands. To match the force field from Moghaddam *et al.*[32] as closely as possible, 1–4 electrostatics were scaled by 0.5 and the nonbonded cutoff was set to 999 Å, which effectively turns off cutoffs. Water was represented with the generalized Born surface area (GBSA) implicit-solvent model without ions and a surface tension of 0.006 kcal/mol/Å². The receptor dielectric was 1.0 and solvent dielectric was 78.5. A time step of 1 fs (using a 2 fs time step with fixed atoms led to unstable trajectories) was used with Langevin dynamics.

CB[7] was minimized for 2500 steps and thermalized by increasing the temperature by 10 K and reinitializing velocities every 100 steps from 0 to 300 K. Receptor snapshots were saved every 0.1 ns from a trajectory of 10 ns.

Binding PMFs for every ligand in Moghaddam *et al.*[32] with the minimized CB[7] structure (15 repetitions each) and 100 receptor simulation snapshots (1 repetition each) were estimated using Hamiltonian replica exchange, which can simultaneously dock a ligand and compute its binding free energy.[21, 24] The implementation is similar to that from Gallicchio and Levy,[24] except that the receptor configuration is fixed. A reservoir of ligand configurations[24] was generated by simulating the ligand for up to 10 ns and saving snapshots every 10 ps. Simulations of the complex in which λ controls the extent of interaction between the ligand and re-

ceptor were run with $\lambda \in \{0, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 1.0\}$. As implemented in NAMD, intermediate values of λ used a soft-core potential with a van der Waals shift coefficient of 5. Electrostatic interactions were turned on when $\lambda = 0.5$. Using the colvars module, a flat-bottom harmonic potential with a spring constant of 10 kcal mol^{-1} Å^{-1} and starting at 0.75 Å was used to restrain the center-of-mass distance between the ligand core (heavy atoms except for the R groups in Moghaddam *et al.*[32]) and the receptor heavy atoms. This potential keeps the ligand within the binding site when interactions are turned off. The binding site volume, $\Omega = \int I_\xi d\xi_L$, is approximated as $4/3\pi(0.75^3)(8\pi^2)$. Because NAMD does not allow the simultaneous use of alchemical decoupling and implicit solvent, simulations were conducted in vacuum.

The replica exchange simulation was initiated by taking a random ligand configuration, applying a random rotation, and randomly placing it within the binding site. This initial configuration was minimized and thermalized with the same protocol as with CB[7], except that it was done in vacuum. The thermalized structure was used to start each replica. Occasionally, the random placement of the ligand led to high forces that caused the simulations to crash; in this case, the simulation was restarted with a different random initial configuration.

After every 5 ps of simulation for every value of λ, 1000 replica exchanges were attempted between each pair of adjacent λ windows. After each set of replica exchange attempts, the ligand configuration for $\lambda = 0$ was replaced with a random ligand configuration from the reservoir, randomly rotated, and placed in the binding site. (This type of reservoir swap satisfies detailed balance.) The simulation was conducted for 25 cycles, saving snapshots every 0.5 ps, for a total of 2 ns of simulation for each binding PMF. The docking and equilibration period, defined as the time before the potential energy of the fully coupled state is within 20 k_BT of its energy for the final snapshot, was ignored in subsequent analysis.

Because alchemical coupling calculations were performed in vacuum, binding PMFs were estimated based on a decomposition of $B(r_R)$,

$$B(r_R) = B_{cpl} + B_{RL} - B_L - \Delta U(r_R),$$

$$B_{cpl} = -\beta^{-1} \ln \left(\frac{\int I_\xi e^{-\beta U(r_{RL})} dr_L d\xi_L}{\int I_\xi e^{-\beta [U(r_L) + U(r_R)]} dr_L d\xi_L} \right),$$

$$B_{RL} = -\beta^{-1} \ln \left(\frac{\int I_\xi e^{-\beta \Delta U(r_{RL})} e^{-\beta U(r_{RL})} dr_L d\xi_L}{\int I_\xi e^{-\beta U(r_{RL})} dr_L d\xi_L} \right), \quad (23)$$

$$B_L = -\beta^{-1} \ln \left(\frac{\int I_\xi e^{-\beta \Delta U(r_L)} e^{-\beta U(r_L)} dr_L d\xi_L}{\int I_\xi e^{-\beta U(r_L)} dr_L d\xi_L} \right).$$

B_{cpl} is the free energy of turning on the interactions between the ligand and the rigid receptor in vacuum. B_{RL}, B_L, and $\Delta U(r_R)$ are free energies of transferring the complex, ligand, and receptor, respectively, from vacuum to the target state (in implicit solvent). They are based on $\Delta U(r_X) = U_T(r_X) - U(r_X)$, the potential energy difference between r_X in the target state versus the state from which configurations were

sampled (in vacuum). B_{cpl} was estimated by applying MBAR[30] to snapshots from every 0.5 ps of simulation, and B_{RL} and B_L by single-step FEP. (Evaluating transfer free energies by MBAR would require the computationally expensive step of calculating target-state potential energies for every snapshot.)

This decomposition makes it straightforward to evaluate $B(r_R)$ for a variety of force fields using the same configurational samples. In this work, four are compared:

1. NAMD: the total potential energy from using GBSA in NAMD;[34]
2. M2: the total potential energy from using the GBSA model in the M2 program;[18,19]
3. PB: Poisson-Boltzmann electrostatic solvation free energies from the University of Houston Brownian Dynamics (UHBD) program[35] and bond, angle, dihedral, Coulomb, and van der Waals energies from the M2 program;[18,19]
4. PBSA: Poisson-Boltzmann electrostatic solvation free energies from UHBD[35] and bond, angle, dihedral, Coulomb, van der Waals, and nonpolar surface area energies from M2,[18,19] the combination used in Moghaddam et al.[32]

During this step, the NAMD, M2, and UHBD programs are used strictly for single-point energy evaluations, not for minimization or dynamics. Poisson-Boltzmann energies were calculated with a grid spacing of 0.18 Å with dimensions such that the maximum dimensions of the molecule are 0.7 (or less) of the final grid.[32] For comparison, binding PMFs were also calculated from the dominant state approximation with PBSA energies, using the lowest value of $\Psi(r_{RL})$ observed in the simulations with $\lambda = 0$ or $\lambda = 1$.

Because receptor configurations were sampled from a simulation in GBSA implicit solvent, binding free energies were estimated by using Eq. (22). Binding free energies were also estimated with the dominant state approximation: using the lowest observed value of $\hat{B}(r_R)$ to estimate $-\beta^{-1} \ln \langle e^{-\beta B} \rangle_R^{r_R}$.

To demonstrate the calculation of thermodynamic expectations and for comparison with results from Moghaddam et al.,[32] the mean values of six PBSA energies – van der Waals, Coulomb, electrostatic solvation, valence (bond + angle + dihedral), nonpolar solvation, and total – were estimated for the complex, the receptor, and the ligand. Mean PBSA energies for the ligand and receptor were estimated by applying the importance sampling identity to the ligand from the non-interacting system in vacuum and to the receptor from the GBSA simulation, respectively. Occasionally, energies in the ligand trajectory briefly spiked to very high values. In estimating the mean PBSA energies, these spikes were filtered out by removing data points in which the total PBSA energy is at least $100 \, k_B T$ larger than the PBSA energy of the final snapshot. As the spikes were likely caused by the finite molecular dynamics time step, they would probably be avoided by using a propagator that exactly preserves the Boltzmann distribution, e.g., hybrid Monte Carlo.[36]

Towards estimating the mean PBSA energies of the complex, rigid-receptor expectations were estimated by applying

MBAR[30] to snapshots from the non-interacting and fully interacting states,

$$\hat{\Theta}(r_R) = \frac{\sum_{n=1}^{N} w(r_{RL,n}) O(r_{RL,n})}{\sum_{n=1}^{N} w(r_{RL,n})},$$

$$w(r_{RL}) = \frac{e^{-\beta(U_{PBSA}(R_L) - U_0(R_L))}}{1 + \frac{N_1}{N_0} e^{-\beta(U_1(r_{RL}) - \hat{B}_{cpl} - U_0(r_{RL}))}},$$

(24)

where $U_0(r_{RL})$ and $U_1(r_{RL})$ are the potential energies of the non-interacting and fully interacting complexes, respectively, $U_{PBSA}(r_L)$ is the PBSA energy of only the ligand, and $r_{RL,n}$ is the nth of N snapshots of either the non-interacting (N_0 snapshots) or fully interacting complex (N_1 snapshots). $\hat{B}_{cpl}$ was estimated by using MBAR[30] with all replicas. While it would be possible to estimate the mean PBSA energies using all snapshots from all replicas, this was avoided because of the computational expense of Poisson-Boltzmann calculations, which can take over a minute per snapshot. After obtaining $\hat{\Theta}(r_R)$, the importance sampling identity, Eq. (21), was used to estimate the expectations in Eq. (13). To ensure consistency of the estimator—an estimate of a constant yields the same constant–$\hat{\Theta}(r_R)$ was calculated for $O = 1$, in which case $\langle \hat{\Theta}(r_R) \rangle_R^{r_R} = \langle e^{-\beta B} \rangle_R^{r_R}$. This estimate of $\langle e^{-\beta B} \rangle_R^{r_R}$ was used in the denominator of Eq. (13).

B. Results

Highlighting the importance of an accurate molecular mechanics model, binding PMF estimates are strongly dependent on the force field, as shown in Table I. For the large and highly charged bicyclooctane B11, switching the force field causes the binding PMF to change nearly 40 kcal/mol! With increasing magnitude of charge, larger Coulomb energies lead to larger values of B_{cpl} and larger electrostatic solvation free energies increase the magnitude of B_{RL}, B_L, and $\Delta U(r_R)$ (for estimates of B_{cpl}, B_{RL}, and B_L, see Table I of the supplementary material[37]). Thus, estimating the binding PMF with Eq. (23) entails the difficult task of computing a relatively small difference between large values. The importance of the force field has also been noted for M2 calculations.[18,19] An alternate implementation, e.g., conducting replica exchange within implicit solvent rather than vacuum, may not require the implicit-solvent model to be as accurate.

With 2 ns of total simulation for all replicas, the standard deviation of binding PMF estimates ranges from 0.12 to 1.63 kcal/mol (Table I), with most estimates on the lower range of imprecision. For all of the components of Eq. (23), the mean estimate does not appear to shift after about 0.75 ns, and additional sampling reduces the standard deviation of the estimate (see Fig. 1 and Fig. 1 in the supplementary material[37]). There is no unique component that limits the convergence of $\hat{B}(r_R)$; the slowest converging component varies from ligand to ligand. The binding PMF estimate $\hat{B}(r_R)$ and the minimal interaction energy $\min\{\Psi(r_{RL})\}$ converge at about the same rate, suggesting that the limiting factor for convergence is finding a configuration with the lowest interaction energy. This interpretation is corroborated by the fact that largest ligands

TABLE I. The mean and standard deviation of 15 independent estimates of the binding PMF, $B(r_R)$, (kcal/mol) for various ligands to the minimized structure of CB[7], based on applying Eq. (23) with different force fields (NAMD, M2, PB, and PBSA columns) or on using the minimum observed value of the interaction energy $\Psi(r_{RL})$ from PBSA energies during the $\lambda = 0$ and $\lambda = 1$ simulations (min$\{\Psi(r_{RL})\}$ column). The bottom rows show the correlation coefficient (R^2) and root mean square error (RMSE, Eq. (25)) with respect to isothermal titration calorimetry experiments (ITC) and mining minima calculations (Gilson) from Moghaddam et al.[32] that result from the dominant state approximation—calculating $\Delta \hat{G}^\circ$ by using a single binding PMF estimate $\hat{B}(r_R)$ as an estimate for $-\beta^{-1} \ln \langle e^{-\beta B} \rangle_R^{r_R}$ in Eq. (11).

Ligand	NAMD	M2	PB	PBSA	min$\{\Psi(r_{RL})\}$
AD1	− 14.1 (0.79)	− 22.0 (0.51)	− 23.0 (0.82)	− 25.5 (0.83)	− 31.3 (0.55)
AD2	− 32.5 (0.15)	− 29.0 (0.13)	− 26.8 (0.12)	− 29.4 (0.12)	− 36.9 (0.30)
AD3	− 31.0 (0.16)	− 30.7 (0.18)	− 28.9 (0.23)	− 31.6 (0.23)	− 40.3 (0.28)
AD4	− 44.0 (0.94)	− 36.7 (1.11)	− 24.0 (1.12)	− 26.9 (1.12)	− 36.1 (0.45)
AD5	− 32.2 (0.68)	− 29.0 (0.25)	− 26.0 (0.14)	− 28.5 (0.14)	− 36.2 (0.29)
B02	− 12.8 (0.41)	− 18.8 (0.38)	− 19.8 (0.53)	− 22.6 (0.53)	− 30.6 (0.54)
B05	− 40.4 (0.29)	− 30.6 (0.40)	− 19.5 (0.50)	− 22.3 (0.50)	− 34.3 (0.63)
B11	− 52.4 (1.50)	− 38.5 (1.81)	− 14.1 (1.72)	− 17.5 (1.63)	− 39.3 (1.90)
F01	− 1.8 (1.57)	− 5.5 (0.81)	− 10.9 (0.53)	− 13.6 (0.53)	− 24.7 (0.33)
F02	− 14.8 (0.96)	− 14.1 (0.67)	− 16.2 (0.42)	− 19.2 (0.42)	− 31.7 (0.38)
F03	− 16.3 (1.61)	− 13.1 (1.03)	− 16.4 (0.95)	− 19.5 (0.94)	− 31.1 (0.71)
F06	− 30.6 (0.18)	− 20.0 (0.21)	− 21.9 (0.18)	− 25.4 (0.18)	− 37.2 (0.24)
R^2_{ITC}	0.884	0.750	0.454	0.490	0.883
$RMSE_{ITC}$	12.8	7.9	4.9	4.7	12.2
R^2_{Gilson}	0.827	0.907	0.705	0.712	0.792
$RMSE_{Gilson}$	10.4	4.8	5.4	4.5	11.3

with the most rotatable bonds (see Moghaddam et al.[32] for structures) also have the most variance in $\hat{B}(r_R)$, as the flexibility increases the challenge of finding configurations with low $\Psi(r_{RL})$.

The accuracy of binding free energy estimates was assessed with the correlation coefficient and root mean square

FIG. 1. The mean and standard deviation of 15 independent estimates of $B(r_R)$, B_{cpl}, B_{RL}, B_L, and min$\{\Psi(r_{RL})\}$ (kcal/mol) based on PBSA energies as a function of total MD simulation time for the ligand B02. Analogous plots for the other ligands in this study are available as Fig. 1 in the supplementary material.[37]

error (RMSE)

$$RMSE(m1, m2) = \sqrt{\frac{1}{L} \sum_{l=1}^{L} (\Delta G^\circ_{l,m1} - \Delta G^\circ_{l,m2})^2} \quad (25)$$

between methods $m1$ and $m2$, where $\Delta G^\circ_{l,m}$ is the binding free energy estimate for ligand l of L ligands using method m (Tables I and II).

Binding free energy estimates based on the binding PMF for a minimized receptor structure suffices to provide high correlation with experiment ($R^2 = 0.884$ for NAMD) and M2 free energy calculations ($R^2 = 0.827$ for NAMD) (see Table I). Surprisingly, binding free energies from NAMD GBSA calculations are more highly correlated to these benchmarks than $\Delta \hat{G}^\circ$ from PBSA calculations. Ironically, the high correlation may be explained by inaccurately large binding PMF values resulting from highly charged ligands, as the molecules in this set with the strongest charges also tend to have stronger binding affinities. Although the correlation coefficient is high, the RMSE is also considerable, over 10 kcal/mol. Similar performance (R^2 and RMSE) is observed by using the dominant state approximation with PBSA calculations. In contrast, using Eq. (23) with PBSA leads to less correlated (lower R^2) but more accurate (lower RMSE) estimates of the binding free energy.

Even for this simple system, binding free energy estimates are substantially improved by using multiple receptor structures (Table II). With binding PMFs from PBSA energies for 100 receptor structures, there is both higher correlation and lower RMSE with respect to experiment ($R^2_{Exp} = 0.704$, $RMSE_{Exp} = 4.5$) and especially with respect to M2 free energy calculations ($R^2_{Gilson} = 0.925$, $RMSE_{Gilson} = 2.4$).
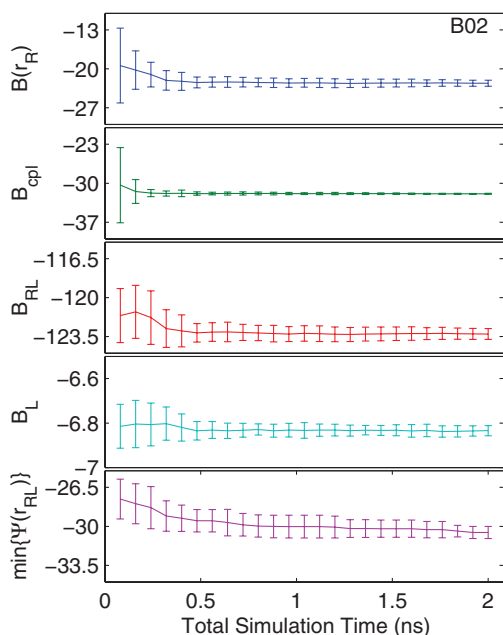
TABLE II. Estimates of the binding free energy ΔG° (kcal/mol) of various ligands to CB[7]. First, binding PMFs $B(r_R)$ are estimated based on Eq. (23) for 100 receptor snapshots from a simulation in GBSA implicit solvent. Then $\Delta \hat{G}^\circ$ is calculated using Eq. (22). The value in the parentheses is the standard deviation from bootstrapping: the binding free energy is estimated based on 1000 random selections of 100 binding PMFs. The experimental and Gilson columns are isothermal calorimetry measurements (ITC) and M2 calculations, respectively, taken from Moghaddam *et al.*[32] The bottom rows are the correlation coefficient (R^2) and root mean square error (RMSE, Eq. (25)) with respect to the ITC and Gilson columns.

Ligand	ITC	Gilson	NAMD	M2	PB	PBSA
AD1	− 14.1	− 18.2	− 9.4 (0.23)	− 16.3 (0.15)	− 17.6 (0.25)	− 20.1 (0.25)
AD2	− 19.4	− 25.9	− 27.9 (0.19)	− 24.3 (0.22)	− 22.9 (0.27)	− 25.4 (0.26)
AD3	− 20.4	− 25.6	− 35.7 (5.03)	− 28.6 (1.87)	− 23.5 (0.23)	− 26.2 (0.23)
AD4	− 21.5	− 29.7	− 40.5 (0.21)	− 33.7 (0.32)	− 24.3 (1.11)	− 27.1 (1.06)
AD5	− 19.1	− 24.1	− 29.5 (1.24)	− 24.0 (0.20)	− 22.0 (0.35)	− 24.4 (0.34)
B02	− 13.4	− 12.0	− 9.0 (0.38)	− 13.7 (0.16)	− 15.4 (0.26)	− 18.1 (0.25)
B05	− 19.5	− 23.1	− 38.0 (0.40)	− 27.7 (0.27)	− 18.6 (0.27)	− 21.4 (0.27)
B11	− 20.6	− 22.4	− 51.2 (0.34)	− 37.3 (0.24)	− 17.2 (0.53)	− 20.5 (0.51)
F01	− 12.9	− 10.2	0.3 (0.82)	− 0.6 (0.34)	− 4.9 (0.26)	− 7.6 (0.25)
F02	− 16.8	− 12.4	− 12.0 (0.70)	− 9.6 (0.75)	− 11.7 (0.70)	− 14.6 (0.71)
F03	− 17.2	− 12.2	− 10.2 (0.16)	− 7.3 (0.24)	− 10.2 (0.22)	− 13.2 (0.22)
F06	− 21.0	− 17.8	− 24.1 (0.34)	− 14.1 (0.46)	− 16.2 (0.51)	− 19.7 (0.52)
R^2_{ITC}		0.782	0.870	0.745	0.671	0.704
$RMSE_{ITC}$		4.6	14.0	9.0	4.4	4.5
R^2_{Gilson}			0.841	0.892	0.923	0.925
$RMSE_{Gilson}$			11.3	5.9	3.4	2.4

While there are some variations on the order of a few kcal/mol, mean potential energy changes upon complexation are also consistent with results from Moghaddam *et al.*[32] (Table III). Minor discrepancies between M2 and implicit ligand free energy and mean potential energy calculations may be explained by a combination of imperfect sampling in the current calculations and the approximations in M2. As the described calculations were performed in vacuum, the samples may not be from the same configurational space as those in implicit solvent. On the other hand, M2 assumes that the energy landscape of the ligand, receptor, and complex are a truncated harmonic wells with anharmonicity corrections.

Compared to the full procedure for estimating the binding PMF, applying the dominant state configuration leads to a reduction in the correlation with M2 results and an increase in the RMSE (Table IV). In contrast, applying the dominant state approximation to calculate $\Delta \hat{G}^\circ$ from $\hat{B}(r_R)$ leads to a near-constant reduction of about 3 kcal/mol in the estimated binding free energy. While the RMSE increases, the correlation with M2 results remains nearly identical. Given a set of $\hat{B}(r_R)$ results, however, there is essentially no reason to apply the dominant state approximation rather than Eq. (22).

There is considerable variation in the binding PMFs for the 100 receptor structures (Fig. 2 and Fig. 2 in the supplementary material[37]). For most of the ligands, the range of binding PMFs spans 10–20 kcal/mol. While the binding PMF of the minimized structure is often near the lower end of the binding PMF distribution, this is not always the case. In larger ligands, the binding PMF appears to be lower for other receptor structures. The fact that a single structure does not always lead to the lowest binding PMF shows a major limitation

TABLE III. Estimates of the mean potential energy changes (kcal/mol) upon the binding of various ligands to CB[7]. The columns refer to van der Waals (VDW), Coulomb (Coul), electrostatic solvation (PB), valence (Val, bond + angle + dihedral), nonpolar solvation (NP), and total energies. The value in the parentheses is the standard deviation from bootstrapping: the observable is estimated based on 1000 random selections of 100 values of $\hat{\Theta}$. In Table II of the supplementary material,[37] mean potential energies for the ligand, receptor, and complex are also shown.

Ligand	VDW	Coul	PB	Val	NP	Total
AD1	− 32.5 (0.471)	0.1 (1.509)	4.8 (1.547)	− 5.0 (2.785)	− 2.5 (0.011)	− 35.2 (2.547)
AD2	− 33.6 (0.931)	− 65.8 (1.032)	64.9 (0.783)	− 5.9 (1.910)	− 2.5 (0.017)	− 42.9 (1.693)
AD3	− 32.8 (0.718)	− 64.4 (0.693)	62.2 (0.855)	− 5.7 (2.128)	− 2.6 (0.009)	− 43.4 (2.388)
AD4	− 38.1 (1.400)	− 125.2 (3.283)	124.4 (1.003)	1.9 (4.475)	− 2.7 (0.070)	− 39.9 (2.817)
AD5	− 33.3 (1.374)	− 65.1 (1.549)	64.8 (1.415)	− 4.9 (1.782)	− 2.5 (0.025)	− 40.9 (1.834)
B02	− 33.3 (0.622)	− 5.8 (1.067)	9.7 (0.770)	1.4 (3.379)	− 2.7 (0.022)	− 30.6 (2.187)
B05	− 32.9 (0.896)	− 138.2 (1.231)	138.0 (1.151)	− 2.3 (1.236)	− 2.8 (0.013)	− 38.1 (1.673)
B11	− 39.9 (1.192)	− 199.3 (2.280)	212.0 (1.066)	− 5.6 (5.431)	− 3.4 (0.075)	− 36.2 (4.475)
F01	− 26.2 (0.497)	− 8.2 (1.824)	14.2 (1.119)	8.7 (4.842)	− 2.7 (0.017)	− 14.3 (4.079)
F02	− 26.9 (1.518)	− 65.7 (2.078)	65.9 (0.923)	− 0.9 (2.237)	− 3.0 (0.012)	− 30.6 (2.253)
F03	− 28.7 (0.832)	− 58.0 (0.987)	64.2 (0.649)	− 0.4 (3.552)	− 3.0 (0.015)	− 26.1 (3.428)
F06	− 35.1 (1.154)	− 116.1 (0.810)	120.9 (0.651)	− 8.8 (4.862)	− 3.5 (0.013)	− 42.6 (4.132)

TABLE IV. Estimates of the binding free energy $\Delta G°$ (kcal/mol) using the PBSA model. First, the binding PMF $B(r_R)$ is estimated with the dominant state approximation (min $\{\Psi(r_R)\}$) or based on Eq. (23), using Hamiltonian replica exchange for B_{cpl} (HREX). Then, $\Delta\hat{G}°$ is from the dominant state approximation (min$\{\hat{B}(r_R)\}$) or based on Eq. (22) (EXP). The bottom rows show the correlation coefficient (R^2) and root mean square error (RMSE, Eq. (25)) with respect to isothermal titration calorimetry experiments (ITC) and mining minima calculations (Gilson) from Moghaddam *et al.*,[32] and the fourth column.

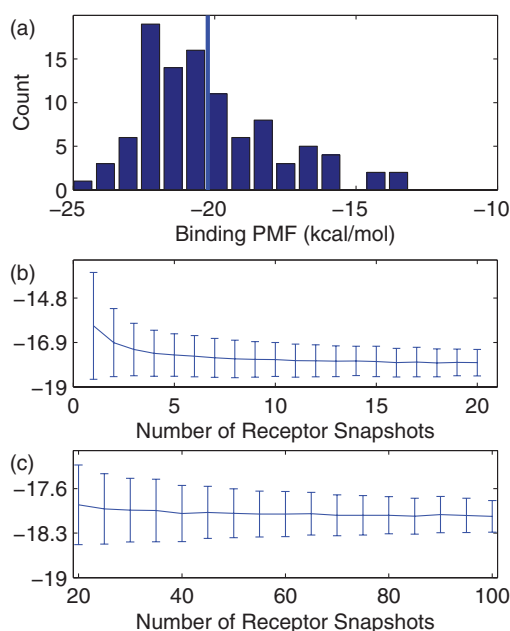| Ligand $\hat{B}(r_R)$ $\Delta\hat{G}°$	min$\{\Psi(r_R)\}$ min$\{\hat{B}(r_R)\}$	min$\{\Psi(r_R)\}$ EXP	HREX min$\{\hat{B}(r_R)\}$	HREX EXP
AD1	−28.6	−27.2	−22.0	−20.1
AD2	−36.4	−34.6	−27.6	−25.4
AD3	−38.1	−36.8	−27.6	−26.2
AD4	−43.1	−40.4	−29.8	−27.1
AD5	−35.8	−33.6	−26.8	−24.4
B02	−29.8	−27.9	−21.0	−18.1
B05	−37.9	−35.6	−23.7	−21.4
B11	−48.5	−45.7	−23.1	−20.5
F01	−22.7	−21.3	−10.2	−7.6
F02	−30.9	−28.8	−17.0	−14.6
F03	−28.7	−27.0	−14.5	−13.2
F06	−35.6	−33.8	−21.3	−19.7
R^2_{ITC}	0.849	0.855	0.684	0.704
RMSE$_{ITC}$	17.3	15.3	5.8	4.5
R^2_{Gilson}	0.787	0.795	0.926	0.925
RMSE$_{Gilson}$	15.8	13.9	3.5	2.4
R^2_{Exp}	0.723	0.736	0.996	
RMSE$_{Exp}$	15.5	13.6	2.3	

FIG. 2. (a) Histogram of binding PMF estimates $\hat{B}(r_R)$ (kcal/mol) of B02 to 100 snapshots of CB[7] using PBSA energies. The vertical line shows the mean binding PMF for the minimized receptor structure. (b) and (c) Estimates of the binding free energy $\Delta G°$ of B02 to CB[7] (kcal/mol), using PBSA energies, as a function of the number of receptor snapshots. The line and error bars denote the mean and standard deviation from bootstrapping: the binding free energy is estimated 100 times using random selections of N out of 100 binding PMFs. Analogous plots for the other ligands in this study are available as Figs. 2 and 3 in the supplementary material.[37]

of using a single receptor structure to estimate binding free energies.

In spite of the variability of binding PMFs, for the ligands in the test set, the average value of $\Delta G°$ appears to stabilize after a relatively small number (about 15) of receptor snapshots (Fig. 2 and Fig. 3 in the supplementary material[37]). Using a greater number of snapshots slightly reduces the variance of binding free energy estimates. After the certain point, however, further reduction in the variance of $\Delta\hat{G}°$ is limited by the variance in binding PMF estimates.

V. DISCUSSION

While the good agreement between implicit ligand and M2 calculations provides a proof of principle, the convergence and accuracy of implicit ligand calculations will differ with other classes of receptor-ligand pairs. With protein-ligand pairs, for example, representative sampling of receptors and finding low-energy poses of the ligand will likely require much more MD simulation time. On the other hand, many protein-ligand systems are not as strongly charged and may be less sensitive to the electrostatic solvation free energy. Due to these variabilities, assessments for the feasibility of implicit ligand calculations in different classes of systems will prove valuable. Tests for convergence and accuracy may be similar to those performed for the CB[7] system.

Numerous opportunities remain for further methodological improvement and optimization of implicit ligand free en-

ergy calculations. The accuracy of implicit ligand calculations (and M2 calculations) may be limited by the quality of the force field. The decomposition of the binding PMF in Eq. (23) provides a facile means to integrate alternate and potentially more expensive potential energies, e.g., quantum mechanical calculations or more sophisticated nonpolar solvation free energies. Modeling may also be improved by the inclusion of a few explicit water molecules (see supplementary material.[37]) Another potential avenue for improvement is the fine-tuning of the replica exchange protocol (e.g., using implicit solvent or optimizing the number of stages and values of λ for a particular system) or implementing alternative methods to estimate the binding PMF and binding free energy.

Even without modifying the replica exchange protocol, computations may be accelerated by optimizing existing MD simulation packages for implicit ligand theory. Few modern MD simulation programs take full advantage of rigid degrees of freedom by skipping the calculation of pairwise interactions between rigid atoms. Even fewer implicit-solvent models are designed with rigid receptors in mind;[38] implicit ligand theory may inspire the development of such models.

Implicit ligand theory also provides guidance on how to understand and improve existing molecular docking algorithms. The definition of $\Psi(r_{RL})$ provides a straightforward functional form that can be used to account for solvation free energies and ligand internal energies (strain), which have been noted to be important factors in binding free energies,[39] but are frequently ignored in the interaction energy functions used by docking packages. Implicit ligand theory also delineates

how to improve the ranking of different ligands. Molecular docking packages currently rank receptor-ligand binding free energies based on a single low-energy configuration. As such, they apply the crudest form of implicit ligand theory, the dominant state approximation, to estimate both the binding PMF and the binding free energy. With important modifications to existing algorithms and the application of more complex estimators, the accuracy of scoring functions should be enhanced.

One important potential change to molecular docking is the inclusion of multiple receptor configurations. While most modern docking packages account for the orientation and flexibility of the ligand, the large number of coordinates makes the treatment of receptor flexibility challenging. A number of groups have improved docking performance by treating receptor flexibility by using multiple structures from crystallography[25,40,41] or MD simulations[42] (the *relaxed complex method*[7,43,44]). Molecular dynamics simulations have also revealed binding sites not discovered by crystallography.[45,46] In the case of Human Immunodeficiency Virus integrase, insight into a new binding site even inspired the development of a new drug.[47]

Despite of this success, it has hitherto remained unclear how to combine information from docking to different receptor snapshots. While averaging strategies have been empirically compared,[48] the default strategy has been to rank the ligand using the minimal energy from docking to all the snapshots. With implicit ligand theory, it is clear that the binding free energy may also be estimated by using an exponential average or cumulant expansion of the binding PMF (which may still come from the dominant state approximation) for different snapshots.

The computational expense of the relaxed complex approach may be reduced by clustering snapshots and selecting a representative snapshot from each cluster.[44] Assuming that the binding PMF is constant within the cluster, estimated averages may be weighed by the cluster size. Using a clustering algorithm based on QR factorization to select 33 representative structures Amaro *et al.*[44] were able to accurately reproduce a histogram of docking scores to over 400 structures. Further research will be necessary to develop and validate algorithms that reliably cluster receptor configurations in which the binding PMF is nearly constant.

Compared to the inclusion of multiple receptor structures, a more difficult task is the estimation of $B(r_R)$ using molecular docking, as this involves a paradigm shift from searching for a minimum to sampling from a distribution. Docking algorithms may be broadly classified into two categories: matching and docking simulation.[49] Matching algorithms such as DOCK[50] attempt to match a ligand into a model of the binding site. DOCK models both the ligand and the binding site as a set of spheres and uses algorithms from graph theory to align the ligand spheres into the binding site spheres. In docking simulation methods such as AutoDOCK[49,51–53] and MCDOCK,[54] the ligand starts outside the binding site and its configuration and orientation are progressively modified to search for the lowest energy configuration of the complex.

Matching algorithms can estimate $B(r_R)$ by a postprocessing algorithm. That is, after low-energy complexes are found, they may be used to bias receptor-independent random sampling of the ligand orientation by a confining potential $U_c(\xi_L)$, for use in Eq. (17). With a harmonic potential for $U_c(\xi_L)$, the ligand orientation will come from a Gaussian distribution. An alternative postprocessing algorithm is to use the lowest energy structure from a matching algorithm as a starting point for a rigid-receptor MD simulation. This is not prohibitively expensive; Graves *et al.*[8] even used MD simulations with flexibility near the binding site as a postprocessing step for molecular docking. Samples from this simulation would be used to estimate $B(r_R)$ based on Eq. (18) or (19).

Docking simulation methods, on the other hand, will need to be modified to sample from a known distribution rather than to search for the minimum energy. This change may not require a complete revamp. Docking simulation algorithms are often based on Monte Carlo approaches, which preserve a desired distribution or may be readily modified to do so. For example, MCDOCK[54] and early generations of AutoDOCK[51,52] use simulated annealing, a procedure for which it is possible to calculate the importance sampling weight.[55]

In addition to providing a path to rigorous binding free energies from molecular docking, implicit ligand theory also quantifies existing notions[56] about whether molecular recognition proceeds by induced fit or conformational selection.[57,58] As all receptor configurations have finite Boltzmann probability, the issue is a matter of degree. Suppose that a receptor binds to two different ligands with the same binding free energy, one by conformational selection and the other by induced fit. If, to a good approximation, the complex is dominated by a single structure with receptor configuration r_R^* such that $B(r_R) = \infty$ for all other receptor configurations, then Eq. (11) simplifies to $\Delta G^\circ = \mathcal{U}(r_R^*) + B(r_R^*) + \beta^{-1} \ln Z_R + \Delta G_\xi$. For the ligand that binds by conformational selection, $p(r_R^*) = e^{-\beta \mathcal{U}(r_R^*)}/Z_R$ has a reasonably high probability. In the induced fit complex, $\mathcal{U}(r_R^*)$ is much less favorable and $B(r_R^*)$ must compensate accordingly to achieve the same ΔG°.

Source code and data used in this paper are available at https://simtk.org/home/implicit_ligand .

[1] B. K. Shoichet, Nature (London) **432**, 862 (2004).
[2] G. Klebe, Drug Discovery Today **11**, 580 (2006).
[3] R. Kim and J. Skolnick, J. Comput. Chem. **29**, 1316 (2008).

[4]N. Moitessier, P. Englebienne, D. Lee, J. Lawandi, and C. R. Corbeil, Br. J. Pharmacol. **153**, S7 (2009).

[5]D. Plewczynski, M. Laźniewski, R. Augustyniak, and K. Ginalski, J. Comput. Chem. **32**, 742 (2010).

[6]J. Wang, P. Morin, W. Wang, and P. Kollman, J. Am. Chem. Soc. **123**, 5221 (2001).

[7]J. Lin, A. Perryman, J. Schames, and J. McCammon, Biopolymers **68**, 47 (2003).

[8]A. P. Graves, D. M. Shivakumar, S. E. Boyce, M. P. Jacobson, D. A. Case, and B. K. Shoichet, J. Mol. Biol. **377**, 914 (2008).

[9]D. C. Thompson, C. Humblet, and D. Joseph-McCarthy, J. Chem. Inf. Model. **48**, 1081 (2008).

[10]T. Hou, J. Wang, Y. Li, and W. Wang, J. Comput. Chem. **32**, 866 (2010).

[11]M. W. Chang, C. Ayeni, S. Breuer, and B. E. Torbett, PLoS ONE **5**, e11955 (2010).

[12]N. Huang, C. Kalyanaraman, K. Bernacki, and M. P. Jacobson, Phys. Chem. Chem. Phys. **8**, 5166 (2006).

[13]Activities have been assumed to be unity, a reasonable approximation in the limit of low concentrations.

[14]M. K. Gilson, J. Given, B. Bush, and J. McCammon, Biophys. J. **72**, 1047 (1997).

[15]J. Wang, C. Tan, Y.-H. Tan, Q. Lu, and R. Luo, Commun. Comput. Phys. **3**, 1010 (2008), available at : http://www.global-sci.com/issue/abstract/readabs.php?vol=3&page=1010.

[16]M. Feig and C. Brooks, Curr. Opin. Struct. Biol. **14**, 217 (2004).

[17]J. Michel and J. W. Essex, J. Med. Chem. **51**, 6654 (2008).

[18]C.-E. Chang and M. K. Gilson, J. Comput. Chem. **24**, 1987 (2003).

[19]C.-E. Chang, M. J. Potter, and M. K. Gilson, J. Phys. Chem. B **107**, 1048 (2003).

[20]M. Lee, Biophys. J. **90**, 864 (2005).

[21]E. Gallicchio, M. Lapelosa, and R. M. Levy, J. Chem. Theory Comput. **6**, 2961 (2010).

[22]J. Cohen, A. Arkhipov, R. Braun, and K. Schulten, Biophys. J. **91**, 1844 (2006).

[23]W. Jiang, M. Hodoscek, and B. Roux, J. Chem. Theory Comput. **5**, 2583 (2009).

[24]E. Gallicchio and R. M. Levy, J. Comput.-Aided Mol. Des. **26**(5), 505 (2012).

[25]S. Rao, P. C. Sanschagrin, J. R. Greenwood, M. P. Repasky, W. Sherman, and R. Farid, J. Comput.-Aided Mol. Des. **22**, 621 (2008).

[26]*Free Energy Calculations*, edited by C. Chipot and A. Pohorille (Springer, Berlin, 2007), Vol. 86.

[27]R. Zwanzig, J. Chem. Phys. **22**, 1420 (1954).

[28]J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).

[29]C. H. Bennett, J. Comput. Phys. **22**, 245 (1976).

[30]M. R. Shirts and J. D. Chodera, J. Chem. Phys. **129**, 124105 (2008).

[31]S. Moghaddam, Y. Inoue, and M. K. Gilson, J. Am. Chem. Soc. **131**, 4012 (2009).

[32]S. Moghaddam, C. Yang, M. Rekharsky, Y. H. Ko, K. Kim, Y. Inoue, and M. K. Gilson, J. Am. Chem. Soc. **133**, 3570 (2011).

[33]Using the linear combination of pairwise overlap[59] algorithm, NAMD 2.9 calculates a negative surface area for CB[7]. NAMD directly uses

Appendix B of Weiser *et al.*,[59] in which the P_1 parameter for the N *sp3* atom type with 1 bonded neighbor is 7.8602×10^{-2}, which is smaller than other P_1 values and the corresponding P_2 parameter. By definition, P_1 should be larger than P_2. To bring this parameter in line with other P_1 and to make it larger than P_2, this parameter was multiplied by 10. The modified code yields a positive surface area for CB[7].

[34]J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, J. Computat. Chem. **26**, 1781 (2005).

[35]M. E. Davis, J. D. Madura, B. A. Luty, and J. McCammon, Comput. Phys. Commun. **62**, 187 (1991).

[36]S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987).

[37]See supplementary material at http://dx.doi.org/10.1063/1.4751284 for tables and figures.

[38]O. Guvench, J. Weiser, P. Shenkin, I. Kolossváry, and W. Still, J. Comput. Chem. **23**, 214 (2001).

[39]D. L. Mobley and K. A. Dill, Structure **17**, 489 (2009).

[40]I. R. Craig, J. W. Essex, and K. Spiegel, J. Chem. Inf. Model. **50**, 511 (2010).

[41]G. Bottegoni, W. Rocchia, M. Rueda, R. Abagyan, and A. Cavalli, PLoS ONE **6**, e18845 (2011).

[42]S. E. Nichols, R. Baron, A. Ivetac, and J. A. McCammon, J. Chem. Inf. Model. **51**, 1439 (2011).

[43]J. Lin, A. Perryman, J. Schames, and J. McCammon, J. Am. Chem. Soc. **124**, 5632 (2002).

[44]R. E. Amaro, R. Baron, and J. A. McCammon, J. Comput.-Aided Mol. Des. **22**, 693 (2008).

[45]J. Schames, R. Henchman, J. Siegel, C. Sotriffer, H. Ni, and J. McCammon, J. Med. Chem. **47**, 1879 (2004).

[46]R. E. Amaro, D. D. L. Minh, L. S. Cheng, W. M. Lindstrom, A. J. Olson, J.-H. Lin, W. W. Li, and J. A. McCammon, J. Am. Chem. Soc. **129**, 7764 (2007).

[47]J. D. Durrant and J. A. McCammon, BMC Biol. **9**, 71 (2011).

[48]J. L. Paulsen and A. C. Anderson, J. Chem. Inf. Model. **49**, 2813 (2009).

[49]G. Morris, D. Goodsell, R. Halliday, R. Huey, W. Hart, R. Belew, and A. Olson, J. Comput. Chem. **19**, 1639 (1998).

[50]I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge, and T. E. Ferrin, J. Mol. Biol. **161**, 269 (1982).

[51]D. Goodsell and A. Olson, Proteins **8**, 195 (1990).

[52]G. Morris, D. Goodsell, R. Huey, and A. Olson, J. Comput.-Aided Mol. Des. **10**, 293 (1996).

[53]R. Huey, G. M. Morris, A. J. Olson, and D. S. Goodsell, J. Comput. Chem. **28**, 1145 (2007).

[54]M. Liu and S. Wang, J. Comput.-Aided Mol. Des. **13**, 435 (1999).

[55]R. Neal, Stat. Comput. **11**, 125 (2001).

[56]H. Carlson, Curr. Opin. Chem. Biol. **6**, 447 (2002).

[57]Y. Xu, J. P. Colletier, H. Jiang, I. Silman, J. L. Sussman, and M. Weik, Protein Sci. **17**, 601 (2008).

[58]D. Bucher, B. J. Grant, and J. A. McCammon, Biochemistry **50**, 10530 (2011).

[59]J. Weiser, P. S. Shenkin, and W. C. Still, J. Comput. Chem. **20**, 217 (1999).