Implicit ligand theory for relative binding free energies

Trung Hai Nguyen and David D. L. Minh^{a)}

Department of Chemistry, Illinois Institute of Technology, Chicago, Illinois 60616, USA

(Received 24 November 2017; accepted 20 February 2018; published online 14 March 2018)

Implicit ligand theory enables noncovalent binding free energies to be calculated based on an exponential average of the binding potential of mean force (BPMF)—the binding free energy between a flexible ligand and rigid receptor—over a precomputed ensemble of receptor configurations. In the original formalism, receptor configurations were drawn from or reweighted to the apo ensemble. Here we show that BPMFs averaged over a holo ensemble yield binding free energies relative to the reference ligand that specifies the ensemble. When using receptor snapshots from an alchemical simulation with a single ligand, the new statistical estimator outperforms the original. *Published by AIP Publishing.* <https://doi.org/10.1063/1.5017136>

I. INTRODUCTION

Noncovalent associations between receptors and small organic ligands play an important role in many important biological processes, including the recognition of substrates by enzymes, of hormones by receptors, and of pharmaceuticals by drug targets. Many computational methods have been developed to help characterize these interactions and to guide rational drug design.^{1–5}

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

$$\Delta G_{RL}^{\circ} = -k_B T \ln \left(\frac{C^{\circ} C_{RL}}{C_R C_L} \right), \quad (1)$$

where k_B is the Boltzmann's constant and T is the temperature in Kelvin. C_X is the equilibrium concentration of the species $X \in \{R, L, RL\}$ and C° is the standard concentration (usually $1\text{M} = 1/1660 \text{ \AA}^3$). It is often more convenient to calculate the relative binding free energy between a ligand of interest and a reference ligand, L_o ,

$$\Delta \Delta G_{RL}^{\circ} = \Delta G_{RL}^{\circ} - \Delta G_{RL_o}^{\circ}. \quad (2)$$

The definition of $\Delta G_{RL_o}^{\circ}$ is analogous to ΔG_{RL}° except that L_o replaces L . Methods to compute binding free energies span a wide spectrum of theoretical rigor and computational expense.²

Among methods to compute binding free energies, alchemical pathway methods⁶ are the most rigorous and computationally expensive. These methods, which employ physically unrealistic intermediate thermodynamic states, are used for both ΔG_{RL}° and $\Delta \Delta G_{RL}^{\circ}$ calculations. In the former, configurations of the complex are sampled from a series of states where the ligand is decoupled or physically separated from the receptor.^{7–10} In the latter, one ligand is converted into another. Accurate protein-ligand binding free energy calculations have been reported in many publications.^{11–18} Because

sampling of complex conformations needs to be repeated for every receptor-ligand pair, however, performing alchemical free energy calculations for a large library of ligands is very computationally demanding.

Implicit ligand theory (ILT)¹⁹ has the potential to facilitate fast and accurate binding free energy calculations between a large library of ligands and a single receptor.

ILT is based on a formal separation of receptor and ligand configurational sampling: binding free energies are an exponential average of the binding potential of mean force (BPMF)—the binding free energy between a *flexible* ligand and *rigid* receptor—over an ensemble of receptor conformations. Calculations based on ILT are as rigorous as any alchemical binding free energy calculation in implicit solvent. The approach has a key benefit that the receptor conformational ensemble can be thoroughly sampled once and recycled for binding free energy calculations with multiple ligands.

The name ILT originates from a mathematical analogy to implicit solvent models. In the same way that implicit solvent models formally invoke an integral over all solvent degrees of freedom to yield the solvation free energy, ILT invokes an integral over all ligand degrees of freedom to yield the BPMF. In practice, however, the computation of these integrals is different. While solvation free energies are usually based on a continuum dielectric model that ignore positions of solvent molecules, BPMFs have hitherto been based on Monte Carlo simulations²⁰ or fast Fourier transform operations²¹ that explicitly consider the position of the ligand.

Although BPMF calculations are nontrivial, they can be performed relatively quickly compared to binding free energy calculations with a flexible receptor. First, they only require sampling of ligands opposed to entire complexes. Second, receptor-ligand interaction energies can be quickly obtained by interpolating a 3D grid (this strategy is widely used in molecular docking^{22,23}). Once interaction energy grids have been computed and stored, calculation times become independent of the number of atoms in the receptor.

^{a)}Electronic mail: dminh@iit.edu

Although the approach is promising, a key barrier to the successful application of ILT is sampling receptor conformations relevant to the holo (ligand-bound) ensembles of the ligands of interest. The original formulation of ILT involves an exponential average of BPMFs over the apo (ligand-free) ensemble of the receptor. This strategy may be successful for binding processes that occur by conformational selection—that is, when binding-competent conformations are among those populated in the equilibrium *prior* to binding (Fig. 1). However, if binding occurs by induced fit, such that conformations relevant to binding are rarely accessed in the apo ensemble, then sampling from the apo ensemble is unlikely to be a successful strategy for ILT-based binding free energy calculations. In these cases, the simplest approach to obtaining binding-competent conformations of a receptor may be simulations of another holo ensemble. In the left panel of Fig. 1, for example, sampling the blue holo ensemble may be useful for binding free energy calculations of the red ligand.

In a recent study involving ILT-based binding free energy calculations between T4 lysozyme and 141 small molecules, Xie, Nguyen, and Minh²⁴ obtained binding-competent receptor configurations by performing alchemical pathway simulations with 6 different ligands. This approach could be described in the context of the left panel of Fig. 1 as sampling from *all* of the solid blue circles to calculate the binding free energy of the dashed red ligand. Sampling from the entire alchemical pathway was necessary in order to accurately reweight each receptor conformation to the apo ensemble.

In this paper, we extend ILT by deriving a statistical estimator for $\Delta\Delta G_{RL}^\circ$ that circumvents the requirement of reweighting receptor configurations to the apo ensemble. Instead, receptor configurations need to be drawn from or reweighted to a holo ensemble. The new formalism yields the binding free energy relative to the ligand used to sample the holo ensemble. We will refer to the original expression as the *apo* estimator and the new one as the *holo* estimator. The holo estimator should be especially useful if ligand binding induces a conformational change that is consistent between different ligands. In this case, receptor snapshots can be drawn from a single simulation of a complex with one ligand. On the other hand, if holo ensembles are distinct subregions within the apo ensemble (right panel in Fig. 1), another receptor sampling strategy is likely to yield more accurate binding free energies.

In addition to being based on BPMFs, our new framework for relative binding free energy calculations differs from previous methods in that it does not require ligand

transformations. Other $\Delta\Delta G_{RL}^\circ$ calculations require the free energy of transforming the reference ligand into the ligands of interest. This requirement limits the diversity of ligands for which $\Delta\Delta G_{RL}^\circ$ can be accurately computed. By contrast, the holo estimator only requires that ligands can bind to the same receptor ensemble.

After introducing the holo estimator, we apply it to binding free energy calculations between T4 lysozyme and 24 ligands based on previously calculated BPMFs.²⁴ Although the previous receptor sampling strategy was not necessarily ideal for the holo estimator, analyzing the same BPMFs allows us to directly compare the performance of the two estimators.

II. THEORY

The absolute standard binding free energy between a receptor R and ligand L can be expressed in terms of configurational integrals for the complex, Z_{RL} , receptor, Z_R , and ligand, Z_L , as⁶

$$\Delta G_{RL}^\circ = -k_B T \ln \left(\frac{Z_{RL}}{Z_R Z_L} \frac{C^\circ}{8\pi^2} \right), \quad (3)$$

$$Z_{RL} = \int I(\xi) J(\xi) e^{-\beta \mathcal{U}(r_{RL})} dr_{RL}, \quad (4)$$

$$Z_X = \int e^{-\beta \mathcal{U}(r_X)} dr_X, \quad X \in \{R, L\}, \quad (5)$$

where $\beta = (k_B T)^{-1}$. $\mathcal{U}(r_X)$ is the potential energy of species X with coordinates r_X in implicit solvent. Coordinates of the complex, r_{RL} , are partitioned into internal coordinates of the receptor and ligand, r_R and r_L , and external coordinates describing the relative translation and rotation of the species, ξ . $I(\xi)$ is an indicator function equal to one when the receptor and ligand are considered bound and zero otherwise. $J(\xi)$ is the Jacobian for the transformation from Cartesian coordinates into the coordinates used for ξ . $8\pi^2$ comes from an integral over the full range of Euler angles in the unbound complex.

The key insight of ILT is that standard binding free energies can be written in terms of an exponential average of the BPMF, $B(r_R)$,¹⁹

$$\Delta G_{RL}^\circ = -k_B T \ln \left[\frac{\int e^{-\beta B(r_R)} e^{-\beta \mathcal{U}(r_R)} dr_R}{\int e^{-\beta \mathcal{U}(r_R)} dr_R} \right] - k_B T \ln \left(\frac{\Omega C^\circ}{8\pi^2} \right), \quad (6)$$

$$B(r_R) = -k_B T \ln \left[\frac{\int I(\xi) J(\xi) e^{-\beta \Psi(r_{RL})} e^{-\beta \mathcal{U}(r_L)} dr_L d\xi}{\int I(\xi) J(\xi) e^{-\beta \mathcal{U}(r_L)} dr_L d\xi} \right], \quad (7)$$

$$\Psi(r_{RL}) = \mathcal{U}(r_{RL}) - \mathcal{U}(r_R) - \mathcal{U}(r_L), \quad (8)$$

FIG. 1. Schematic representation of apo and holo ensembles in configuration space. The circles denote regions in configuration space that are important to various thermodynamic ensembles. Solid blue and dashed red circles indicate two different ligands. In the left panel, a series of circles of the same color indicate one alchemical pathway from the apo to a holo ensemble. The left panel describes induced-fit binding processes in which apo and holo ensembles are distinct. The right panel describes conformational selection processes in which the holo ensembles are subsets of the apo ensemble.

where $\Omega = \int I(\xi)J(\xi)d\xi$, an integral of the indicator function over all external degrees of freedom. If no restraints are placed on rotation of the ligand, $\Omega = 8\pi^2 V_{site}$, where V_{site} is the volume of the binding site. Equation (6) is an expectation over the apo ensemble, where the probability density of r_R is proportional to $\exp[-\beta\mathcal{U}(r_R)]$. [It builds on the observation by Gallicchio, Lapelosa, and Levy²⁵ that Eq. (3) is an ensemble average of the interaction energy over the apo ensemble.] Hence, ΔG_{RL}° can be estimated using the apo estimator,

$$\Delta\hat{G}_{RL}^\circ = -k_B T \ln \left[\sum_{r_R} w_a(r_R) e^{-\beta\hat{B}(r_R)} \right] - k_B T \ln \left(\frac{\Omega C^\circ}{8\pi^2} \right), \quad (9)$$

where $\hat{B}(r_R)$ is an estimate of the BPMF for the receptor configuration r_R and $w_a(r_R)$ is the normalized statistical weight of r_R in the apo ensemble. If receptor snapshots are drawn from the apo ensemble, then $w_a(r_R) = 1/n$, where n is number of receptor samples.

An expression for relative binding free energies in terms of configurational integrals can be obtained by substituting Eq. (3) and the analogous expression for $\Delta G_{RL_o}^\circ$ into the definition of $\Delta\Delta G_{RL}^\circ$,

$$\Delta\Delta G_{RL}^\circ = -k_B T \ln \left(\frac{Z_{RL}}{Z_L} \frac{Z_{L_o}}{Z_{RL_o}} \right). \quad (10)$$

By expanding the configurational integrals, multiplying by $\int I(\xi_o)J(\xi_o)d\xi_o / \int I(\xi)J(\xi)d\xi = 1$, and factoring out $B(r_R)$, we obtain the main theoretical result of our present paper,

$$\begin{aligned} \Delta\Delta G_{RL}^\circ &= -k_B T \ln \left\{ \frac{\int I(\xi_o)J(\xi_o) e^{-\beta\mathcal{U}(r_{RL_o})} e^{-\beta[B(r_R)-\Psi(r_{RL_o})]} dr_{RL_o}}{\int I(\xi_o)J(\xi_o) e^{-\beta\mathcal{U}(r_{RL_o})} dr_{RL_o}} \right\}, \end{aligned} \quad (11)$$

where $\Psi(r_{RL_o}) = \mathcal{U}(r_{RL_o}) - \mathcal{U}(r_R) - \mathcal{U}(r_{L_o})$.

Equation (11) is an expectation over the holo ensemble with the reference ligand, where the probability density of r_{RL_o} is proportional to $\exp[-\beta\mathcal{U}(r_{RL_o})]$. Hence, $\Delta\Delta G_{RL}^\circ$ can be estimated using the holo estimator,

$$\begin{aligned} \Delta\Delta\hat{G}_{RL}^\circ &= -k_B T \ln \left\{ \sum_{r_{RL_o}} w_h(r_{RL_o}) e^{-\beta[\hat{B}(r_R)-\Psi(r_{RL_o})]} \right\} \\ &\quad - \Delta\Delta\hat{G}_{RL_o}^\circ, \end{aligned} \quad (12)$$

$$\Delta\Delta\hat{G}_{RL_o}^\circ = -k_B T \ln \left\{ \sum_{r_{RL_o}} w_h(r_{RL_o}) e^{-\beta[\hat{B}_o(r_R)-\Psi(r_{RL_o})]} \right\}, \quad (13)$$

where $w_h(r_{RL_o})$ is the statistical weight of r_R in the holo ensemble of the receptor complexed with the reference ligand. If receptor snapshots are drawn from the holo ensemble, then $w_h(r_{RL_o}) = 1/n$, where n is the number of receptor samples. $B_o(r_R)$ is the BPMF of the reference ligand.

$\Delta\Delta G_{RL_o}^\circ$, described in Eq. (13), is the “self” relative binding free energy—the relative binding free energy difference between the reference ligand and itself—which should be zero in the limit of infinite sampling. Because the estimated value can be nonzero for a finite number of receptor snapshots, correcting Eq. (12) by $\Delta\Delta\hat{G}_{RL_o}^\circ$ ensures that the free

energy difference between the reference ligand and itself is zero regardless of the number of receptor samples.

In the limit of infinite receptor sampling, Eqs. (9) and (12) should converge to the same values for binding free energies. Based on a finite number of BPMFs, however, they provide numerically distinct estimates. This can be seen by noting that factor $w_a(r_R)$ that appears in Eq. (9) is not equal to the factor $w_h(r_{RL_o})e^{\beta\Psi(r_{RL_o})}$ that appears in Eq. (12); their ratio,

$$\frac{w_h(r_{RL_o})e^{\beta\Psi(r_{RL_o})}}{w_a(r_R)} \propto e^{-\beta\mathcal{U}(r_{L_o})},$$

is not equal to one. Reweighting the holo to the apo ensemble requires not only the interaction energy but also the potential energy of the ligand by itself.

III. COMPUTATIONAL METHODS

The calculations in this study were built on the foundation laid by Xie, Nguyen, and Minh.²⁴ In the previous work, YANK,¹⁰ which implements an alchemical pathway method using a flexible receptor, was used to calculate absolute binding free energies between T4 lysozyme and 24 small organic molecules. The calculations yielded precise free energy estimates, with a standard deviation less than 0.5 kcal/mol for 22 systems and less than 1 kcal/mol for all 24 systems. YANK results for 18 systems were used as *benchmarks* for our ILT-based estimators. The other 6 ligands—methylpyrrole, benzene, *p*-xylene, phenol, *n*-hexylbenzene, and DL-camphor—were used as *reference* ligands.

In the present work, we used the Jensen-Shannon divergence (JSD)²⁶ to compare different ensembles sampled in the YANK calculations. The benchmark holo ensembles were compared to the reference holo ensembles and to the apo ensemble. Based on 15 000 snapshots from each ensemble, a principal components analysis was performed based on heavy atoms within 5 Å of valine 111, 2D probability densities were calculated using a Gaussian kernel density estimate with a bandwidth of 0.1. The probability densities were compared by the JSD, defined as

$$D_{JS}(p||q) = \frac{1}{2}D_{KL}(p||m) + \frac{1}{2}D_{KL}(q||m), \quad (14)$$

where p , q , and $m = \frac{1}{2}(p + q)$ are multivariate probability density functions. D_{KL} is Kullback-Leibler divergence, given by

$$D_{KL}(p||m) = \int p(x) \ln \frac{p(x)}{q(x)} dx, \quad (15)$$

where x is, in general, a multivariate variable.

For the ILT-based binding free energy calculations in the study of Xie, Nguyen, and Minh,²⁴ groups of 96 receptor snapshots were extracted from every alchemical state of YANK simulations with the reference ligands. (Subsequently, the term *group* will be used to refer snapshots drawn from a simulation with a single reference ligand.) BPMFs between 141 ligands and all 576 receptor snapshots were calculated using AIGDock²⁰ with the OBC2 generalized Born/surface area implicit solvent model.²⁷ In the AIGDock calculations,

the AMBER ff14 force field²⁸ was used for the receptor and the generalized AMBER force field²⁹ with Bondi radii³⁰ and AM1BCC partial charges^{31,32} for ligands. After obtaining the BPMFs, Eq. (9) was used to compute absolute binding free energies. The multistate Bennett Acceptance Ratio (MBAR)³³ was used to compute weights (which we will refer to as MBAR weights) for all receptor snapshots in the apo ensemble (see Ref. 24 for details). Using MBAR weights for $w_a(r_R)$, however, was found to not be the most numerically stable scheme. Better results were obtained with weighting scheme (c) in the study of Xie, Nguyen, and Minh²⁴—when each of the selected snapshots was assigned the cumulative MBAR weight of its alchemical state and each group was given equal weight.

In this work, Eq. (12) was used with the same BPMFs to calculate binding free energies for the 24 benchmark ligands relative to each of the 6 reference ligands. Each free energy estimate used up to 96 receptor snapshots from each group. Interaction energies $\Psi(r_{RL_o})$ were calculated using a python script based on OpenMM 6.3.1.^{34,35} The weights $w_h(r_{RL_o})$ were calculated using MBAR.³³ Unlike in the study of Xie, Nguyen, and Minh,²⁴ assigning each selected snapshot the

cumulative MBAR weight of the alchemical state from which it was sampled yielded poor results—an extremely large root mean square error (RMSE) and essentially no correlation with respect to YANK (see Fig. S1 of the [supplementary material](#)). This poor performance resulted from steric overlap between the ligand and receptor in some snapshots from mostly decoupled alchemical states leading to large $\Psi(r_{RL_o})$. Therefore, $w_h(r_{RL_o})$ in Eq. (12) was based on each snapshot's MBAR weight (this is actually a more statistically rigorous procedure).

Free energy calculations based on receptor snapshots from YANK calculations with different ligands were combined using simple rules. Starting with $\Delta\hat{G}_{RL}^\circ$ from the holo estimator, $\Delta\hat{G}_{RL}^\circ$ was calculated by adding the YANK free energy for the reference ligand. Given multiple $\Delta\hat{G}_{RL}^\circ$ from the holo estimator with multiple different ligands, the minimum value was taken as the consensus estimate. As a comparison, $\Delta\hat{G}_{RL}^\circ$ was also calculated using the best estimator from the study of Xie, Nguyen, and Minh²⁴—Eq. (9) with $w_a(r_R)$ based on the cumulative MBAR weight of alchemical states. The consensus estimate was taken to be the exponential average of the BPMF across all 6 groups.

FIG. 2. Binding free energies for the benchmark ligands estimated by YANK (x-axis) and AlGDock (y-axis) using the apo estimator (left column) or holo estimator (right column). Active ligands are shown as small dots and inactive ones as diamonds. Error bars denote the standard deviation from three independent YANK calculations (x-axis) or from bootstrapping BPMFs (y-axis), with the range of error bars representing a single standard deviation. A least-squares linear regression is shown as a dashed line. Each row corresponds to a group. Results for other ligands are shown in Fig. 3.

FIG. 3. Binding free energies for the benchmark ligands, continued from Fig. 2.

IV. RESULTS

A. For single groups, the holo estimator is generally more consistent with YANK than the apo estimator

Although both the apo and holo estimators should be consistent with each other and with YANK in the limit of asymptotic sampling, their performance can be compared based on results for a finite number of receptor snapshots.

For the benchmark ligands, performance was dependent on the group, but the holo estimator was generally more consistent with YANK (Figs. 2 and 3 and Table I).³⁶ Overall, the mean and median of Pearson's R is higher with the holo estimator (0.88 and 0.89, respectively) than the apo estimator (0.78 and 0.79, respectively). In 4 of 6 groups, the Pearson's R is significantly higher for the holo estimator. For the *p*-xylene and DL-camphor groups, the Pearson's R for the two estimators

TABLE I. Comparing the consistency of apo and holo estimators to YANK free energies. Statistics are based on the benchmark ligands. The RMSE is in kcal/mol.

Reference ligand	Apo estimator			Holo estimator		
	Linear regression	RMSE	Pearson's R	Linear regression	RMSE	Pearson's R
Methylpyrrole	$y = 0.74x + 1.73$	3.39(0.38)	0.81(0.08)	$y = 1.08x + 0.41$	1.20(0.31)	0.94(0.03)
Benzene	$y = 0.56x + 2.08$	4.48(0.45)	0.77(0.08)	$y = 0.89x - 0.89$	1.48(0.46)	0.89(0.06)
<i>p</i> -xylene	$y = 1.2x + 0.05$	2.14(0.43)	0.89(0.05)	$y = 0.75x - 1.02$	1.72(0.36)	0.83(0.07)
Phenol	$y = 0.59x - 0.69$	3.73(0.7)	0.59(0.18)	$y = 0.73x - 0.44$	1.94(0.33)	0.82(0.07)
DL-camphor	$y = 1.07x + 1.65$	1.96(0.19)	0.92(0.04)	$y = 0.88x + 3.23$	3.98(0.26)	0.91(0.04)
<i>n</i> -hexylbenzene	$y = 0.67x + 2.68$	4.76(0.69)	0.69(0.11)	$y = 0.98x + 7.41$	7.65(0.33)	0.89(0.05)

TABLE II. “Self” relative binding free energies for 6 reference ligands. The numbers in parentheses are standard deviations estimated by bootstrapping—drawing N samples from the previously sampled population with replacement—100 times.

Reference ligand	$\Delta\Delta\hat{G}_{RL_o}^{\circ}$ (kcal/mol)
Methylpyrrole	4.06(0.41)
Benzene	4.56(0.49)
<i>p</i> -xylene	4.61(0.98)
Phenol	1.96(0.85)
DL-camphor	4.15(1.41)
<i>n</i> -hexylbenzene	-1.81(2.69)

was comparable. On a related note, the least squares linear regression lines for the holo estimator have slopes closer to 1 except for the *p*-xylene and DL-camphor groups (Table I). Across the 6 groups, the mean and median RMSE is lower for the holo estimator (2.99 and 1.83 kcal/mol, respectively) than the apo estimator (3.41 and 3.56 kcal/mol, respectively). The RMSE of the holo estimator has a lower value in 4 of the 6 groups. In the exceptions, DL-camphor and *n*-hexylbenzene groups, much of the error comes from a constant offset, as evidenced by the large y-intercept in the linear regression.

The offset observed when using the holo estimator for the DL-camphor and *n*-hexylbenzene groups is due to the “self” relative binding free energy estimates. On one hand, the estimates have large variance. For the DL-camphor and *n*-hexylbenzene groups, which led to the highest RMSE, the standard deviation of the estimate is significantly larger than with the other ligands (Table II). It is also notable that for *n*-hexylbenzene, $\Delta\Delta\hat{G}_{RL_o}^{\circ}$ is negative, contributing to the large positive deviation observed in ILT-based $\Delta\Delta G_{RL}^{\circ}$ calculations with this reference ligand.

B. For single groups, the two estimators converge at a similar rate

The convergence of free energy calculations to their values for 96 snapshots is similar between the two estimators. The RMSE for both estimators follows a similar decreasing trend and starts to level off at essentially the same number of receptor snapshots (Fig. 4). Noticeable differences between the two estimators can be seen for the benzene group, where the holo estimator shows slightly faster convergence, and the DL-camphor group, where the opposite is true. Nevertheless, these differences may not be statistically significant because they are contained within the error bars. In 4 of the 6 groups in Fig. 5, the convergence of

FIG. 4. Convergence of binding free energies. RMSE with respect to the final results of binding free energies estimated using apo (blue solid lines) and holo (green dashed lines) estimators for single groups. Error bars denote the standard deviation from bootstrapping. Different panels correspond to different reference ligands.

FIG. 5. Convergence of binding free energies. Pearson's R with respect to the final results of binding free energies estimated using apo (blue solid lines) and holo (green dashed lines) estimators for single groups. Error bars denote the standard deviation from bootstrapping. Different panels correspond to different reference ligands.

Pearson's R is not very different between the two estimators. The holo estimator seems to converge slightly faster for the *n*-hexylbenzene group and slower for the DL-camphor group. Therefore, on average, the two estimators perform equally well in terms of convergence with respect to the number of receptor snapshots.

C. For multiple groups, the estimators are equally consistent with YANK but the holo estimator converges more slowly

In contrast to using single groups, consensus estimates for the two estimators are equally consistent with YANK (Fig. 6).

FIG. 6. Binding free energies for 24 ligands estimated by YANK (x-axis) and AIGDock (y-axis) using the apo estimator (left) or holo estimator (right). Results from 6 groups were combined using the exponential average across BPMFs from all groups for the apo estimator or the minimum of six estimates for the holo estimator. Active ligands are shown as dots and inactive ones as diamonds. Error bars denote the standard deviation from three independent YANK calculations (x-axis) or from bootstrapping BPMFs (y-axis), with the range of error bars representing a single standard deviation. The linear regression is shown as a dashed line.

FIG. 7. Convergence of binding free energies. RMSE (left) and Pearson's R (right) with respect to the final results of binding free energies estimated using apo (blue solid lines) and holo (green dashed lines) estimators for receptor snapshots combined from 6 YANK simulations. Error bars denote the standard deviation from bootstrapping.

The holo estimator gives slightly higher RMSE (1.74 ± 0.45 kcal/mol) than the apo estimator (1.64 ± 0.22 kcal/mol), but the difference is within their error bars. The Pearson's R is also the same, 0.90 ± 0.04 for the apo estimator and 0.89 ± 0.04 for the holo estimator.

When using a consensus estimate from multiple groups, the holo estimator converges to its final value for 576 snapshots more slowly than the apo estimator (Fig. 7). Both the RMSE and the Pearson's R for the holo estimator level off at about 150 receptor snapshots, whereas for the apo estimator, they level off after about only 20-50 receptor snapshots.

D. Configuration space overlap is associated with accurate ΔG_{RL}° estimates

Hitherto, we have used the same BPMFs for both the apo and holo estimators. While this approach has enabled us to directly compare the estimators, it does not specifically address the relationship between configuration space

overlap and the accuracy of binding free energy calculations; receptor snapshots are from the same configuration space. In this section, we explore this relationship by calculating ΔG_{RL}° based on receptor snapshots drawn from only the apo or holo ensembles.

In our simulations of T4 lysozyme complexes, most benchmark holo ensembles are closer to most reference holo ensembles than to the apo ensemble, but there are important exceptions. The JSD between most benchmark holo ensembles and reference holo ensembles with methylpyrrole, benzene, *p*-xylene, and phenol are lower than with the apo ensemble, leading to a lower median JSD (Fig. 8). However, the JSD between these ensembles and the benchmark ligands nitrobenzene, methanol, benzaldehyde oxime, and 1,2-diiodobenzene is very high. (The latter three of these benchmark ligands are inactive.) Due to these outliers, the mean JSD between the benchmark holo ensembles and the reference holo ensembles is higher than with the apo ensemble. For the other two reference holo ensembles, with DL-camphor and *n*-hexylbenzene,

FIG. 8. The Jensen-Shannon divergence between the benchmark holo ensembles (rows) and the apo and reference holo ensembles (columns).

FIG. 9. Jensen-Shannon divergence versus deviation of ILT estimates from YANK. Absolute binding free energy estimates using the apo estimator based on 36 BPMFs to receptor snapshots drawn from the apo state of six YANK simulations are shown as red circles. Relative binding free energy estimates using the holo estimator based on 6 BPMFs to receptor snapshots drawn from the holo state of the YANK simulation for the reference ligand are shown as follows: methylpyrrole (blue downward triangles), benzene (green upward triangles), *p*-xylene (cyan leftward triangles), phenol (magenta rightward triangles), DL-camphor (yellow squares), and *n*-hexylbenzene (black diamonds).

the JSD between these ensembles and the benchmark holo ensembles is very high. As previously mentioned, these two groups lead to the largest RMSE (Table II and Fig. 3).

Accurate ILT-based binding free energy estimates are associated with configuration space overlap between the reference ensemble and benchmark holo ensembles. When the JSD is low (below 0.1), all ILT-based estimates are close to the YANK estimates (within around 2 kcal/mol). On the other hand, when the JSD is higher, there is a much broader range of deviation from YANK (Fig. 9). Notably, for snapshots from the *n*-hexylbenzene and DL-camphor ensembles, the binding free energy estimated by ILT is significantly overestimated (weaker than YANK).

V. DISCUSSION

The most parsimonious explanation for the aforementioned JSD relationships is that binding of our ligands to T4 lysozyme occurs through conformational selection, similar to the right panel of Fig. 1. All of the benchmark holo ensembles appear to be subregions of the apo ensemble, as evidenced by the relatively low mean, median, and maximum JSD. Different holo ensembles, however, make up potentially distinct subregions of the apo ensemble. Most benchmark ligands, except for nitrobenzene, methanol, benzaldehyde oxime, and 1,2-diiodobenzene, favor the same subregion as the reference ligands methylpyrrole, benzene, *p*-xylene, and phenol, but different subregions from DL-camphor and *n*-hexylbenzene.

The comparative performance of apo and holo ΔG_{RL}° estimators may also be interpreted in terms of these configuration space relationships. The holo estimator outperforms the apo estimator for single groups because most benchmark holo ensembles are closer to reference holo ensembles than to apo ensembles. The similarity of ensembles means that reweighing factors have less variance and are more numerically stable. In

cases when holo ensembles do not overlap, the holo estimator is able to accurately calculate ΔG_{RL}° from groups because they contain snapshots from the apo ensemble. However, if only snapshots are exclusively from the holo ensemble and configuration space overlap is weak, then binding free energies are significantly overestimated.

Although the holo estimator exhibits superior performance for single groups, it comes with a caveat that the “self” relative binding free energy may lead to a constant offset. Equation (13), which we used to calculate $\Delta \Delta G_{RL_0}^\circ$, is an exponential average. Exponential averages converge slowly if the distribution of the exponent, in this case, $-\beta [\hat{B}_o(r_R) - \Psi(r_{RL_0})]$ are broadly distributed. In the specific case of the *n*-hexylbenzene group, the negative estimate implies that snapshots for which $\hat{B}_o(r_R) < \Psi(r_{RL_0})$ dominate the distribution. In turn, this implies that sampling of ligand configurations in the YANK simulation was limited, such that BPMF calculations accessed lower-energy configurations. In future calculations, an imprecise or negative $\Delta \Delta G_{RL_0}^\circ$ may indicate problems with accurate absolute ΔG_{RL}° estimates based on the given reference holo ensemble.

Another limitation of the holo estimator is in the analysis of multiple groups. In the apo estimator, methods for combining data from multiple groups²⁴ do not discard any information. On the other hand, our consensus holo estimator based on the minimum ΔG_{RL}° from multiple independent calculations essentially discards information from all but one group. This inefficient use of data is the likely reason that the consensus holo estimator converges to its final value more slowly.

Given these advantages and disadvantages of the holo estimator, the appropriate choice of estimator for future studies will likely depend on the receptor sampling scheme. If only receptor snapshots from an alchemical simulation with a single complex are available, the holo estimator is likely to be superior. The holo estimator is even more likely to perform better when the receptor is sampled from a single holo ensemble opposed to a series of alchemical states. On the other hand, the apo estimator converges faster when combining receptor snapshots from multiple alchemical simulations. It is also obviously more suited to situations when the receptor is sampled from the apo ensemble.

The present framework for relative binding free energy calculations has some similarities but is distinct from that developed by Oostenbrink and van Gunsteren.³⁷ Both approaches rely on precomputing a holo ensemble, which can be of an arbitrary (and possibly unphysical) reference state. Based on structural and chemical intuition, Oostenbrink and van Gunsteren³⁷ designed large and soft reference compounds to induce receptor conformations that are relevant for binding to a diverse set of ligands. This sampling strategy can also be applied for the current framework to generate samples for our new estimator, Eq. (12). However, our approach differs in that we do not compute the free energy of transforming a ligand into a reference compound. To calculate these transformation free energies, Oostenbrink and van Gunsteren³⁷ developed clever schemes to map the coordinates of the mostly planar aromatic ligands to the reference compounds. On the other hand, Eq. (12) is based on independent BPMF calculations

between each ligand and structures from the apo ensemble. Our approach is more computationally expensive but applicable to libraries of even more diverse and flexible ligands.

Very recently, Jandova *et al.*³⁸ adapted the one-step perturbation framework from Oostenbrink and van Gunsteren³⁷ to the free energy of amino acid mutagenesis. They used unphysical reference amino acids that incorporated soft-core particles. Free energies of transforming the reference amino acids into standard amino acids were computed based on mapping different rotational orientations of a precomputed library of amino acid side chain conformations onto the reference state. This strategy could potentially be applied to relative binding free energies for ligands with flexible regions branching from a common scaffold.

VI. CONCLUSIONS

We have derived a new binding free energy estimator for ILT. In contrast to the previous estimator, the new one is based on a holo ensemble and gives binding free energies relative to the reference ligand that defines the holo ensemble. For receptor snapshots generated by alchemical simulations of T4 lysozyme and 6 different ligands, we have demonstrated that when using a single group, the newly derived estimator is generally more accurate than the previous apo estimator and converges at a similar rate. When combining multiple groups, the two estimators are equally accurate in the limit of many receptor snapshots, but the holo estimator converged more slowly. Accuracy of the estimator appears to depend on configuration space overlap between the reference ensemble and the holo ensemble for the ligand of interest.

SUPPLEMENTARY MATERIAL

See [supplementary material](#) for binding free energies estimated using the holo estimator when each receptor snapshot is assigned the MBAR weight of the alchemical state it represents instead of the MBAR weight of the individual snapshot (Fig. S1). Figure S2 shows mean holo-ensemble MBAR weights for snapshots drawn from each alchemical state.

ACKNOWLEDGMENTS

We thank Bing Xie for providing data from Xie, Nguyen, and Minh.²⁴ This research was supported by the National Institutes of Health (Grant No. R15GM114781).

¹W. L. Jorgensen, *Science* **303**, 1813 (2004).

²M. K. Gilson and H.-X. Zhou, *Annu. Rev. Biophys. Biomol. Struct.* **36**, 21 (2007).

³J. Michel and J. W. Essex, *J. Comput.-Aided Mol. Des.* **24**, 639 (2010).

- ⁴J. D. Chodera, D. L. Mobley, M. R. Shirts, R. W. Dixon, K. Branson, and V. S. Pande, *Curr. Opin. Struct. Biol.* **21**, 150 (2011).
- ⁵D. L. Mobley and P. V. Klimovich, *J. Chem. Phys.* **137**, 230901 (2012).
- ⁶M. K. Gilson, J. A. Given, B. L. Bush, and J. A. McCammon, *Biophys. J.* **72**, 1047 (1997).
- ⁷Y. Deng and B. Roux, *J. Phys. Chem. B* **113**, 2234 (2009).
- ⁸G.-J. Bekker, N. Kamiya, M. Araki, I. Fukuda, Y. Okuno, and H. Nakamura, *J. Chem. Theory Comput.* **13**, 2389 (2017).
- ⁹G. Heinzelmann, N. M. Henriksen, and M. K. Gilson, *J. Chem. Theory Comput.* **13**, 3260 (2017).
- ¹⁰K. Wang, J. D. Chodera, Y. Yang, and M. R. Shirts, *J. Comput.-Aided Mol. Des.* **27**, 989 (2013).
- ¹¹J. Michel and J. W. Essex, *J. Med. Chem.* **51**, 6654 (2008).
- ¹²S. E. Boyce, D. L. Mobley, G. J. Rocklin, A. P. Graves, K. A. Dill, and B. K. Shoichet, *J. Mol. Biol.* **394**, 747 (2009).
- ¹³X. Ge and B. Roux, *J. Phys. Chem. B* **114**, 9525 (2010).
- ¹⁴L. Wang, B. J. Berne, and R. A. Friesner, *Proc. Natl. Acad. Sci. U. S. A.* **109**, 1937 (2012).
- ¹⁵S. Zhu, S. M. Travis, and A. H. Elcock, *J. Chem. Theory Comput.* **9**, 3151 (2013).
- ¹⁶L. Wang, Y. Wu, Y. Deng, B. Kim, L. Pierce, G. Krilov, D. Lupyan, S. Robinson, M. K. Dahlgren, J. Greenwood, D. L. Romero, C. Masse, J. L. Knight, T. Steinbrecher, T. Beumung, W. Damm, E. Harder, W. Sherman, M. Brewer, R. Wester, M. Murcko, L. Frye, R. Farid, T. Lin, D. L. Mobley, W. L. Jorgensen, B. J. Berne, R. A. Friesner, and R. Abel, *J. Am. Chem. Soc.* **137**, 2695 (2015).
- ¹⁷M. Aldeghi, A. Heifetz, M. J. Bodkin, S. Knapp, and P. C. Biggin, *Chem. Sci.* **7**, 207 (2016).
- ¹⁸S. Wan, A. P. Bhati, S. J. Zasada, I. Wall, D. Green, P. Bamborough, and P. V. Coveney, *J. Chem. Theory Comput.* **13**, 784 (2017).
- ¹⁹D. D. L. Minh, *J. Chem. Phys.* **137**, 104106 (2012).
- ²⁰D. D. L. Minh, e-print [arXiv:1507.03703v1](#) (2015).
- ²¹T. H. Nguyen, H.-X. Zhou, and D. D. L. Minh, *J. Comput. Chem.* **39**, 621 (2017).
- ²²N. Pattabiraman, M. Levitt, T. E. Ferrin, and R. Langridge, *J. Comput. Chem.* **6**, 432 (1985).
- ²³E. C. Meng, B. K. Shoichet, and I. D. Kuntz, *J. Comput. Chem.* **13**, 505 (1992).
- ²⁴B. Xie, T. H. Nguyen, and D. D. L. Minh, *J. Chem. Theory Comput.* **13**, 2930 (2017).
- ²⁵E. Gallicchio, M. Lapelosa, and R. M. Levy, *J. Chem. Theory Comput.* **6**, 2961 (2010).
- ²⁶J. Lin, *IEEE Trans. Inf. Theory* **37**, 145 (1991).
- ²⁷A. Onufriev, D. Bashford, and D. A. Case, *Proteins: Struct., Funct., Bioinf.* **55**, 383 (2004).
- ²⁸J. W. Ponder and D. A. Case, *Adv. Protein Chem.* **66**, 27 (2003).
- ²⁹J. Wang, W. Wang, P. A. Kollman, and D. A. Case, *J. Mol. Graphics Modell.* **25**, 247 (2006).
- ³⁰A. Bondi, *J. Phys. Chem.* **68**, 441 (1964).
- ³¹A. Jakalian, B. L. Bush, D. B. Jack, and C. I. Bayly, *J. Comput. Chem.* **21**, 132 (2000).
- ³²A. Jakalian, D. B. Jack, and C. I. Bayly, *J. Comput. Chem.* **23**, 1623 (2002).
- ³³M. R. Shirts and J. D. Chodera, *J. Chem. Phys.* **129**, 124105 (2008).
- ³⁴P. Eastman and V. S. Pande, *Comput. Sci. Eng.* **12**, 34 (2010).
- ³⁵P. Eastman, M. Friedrichs, and J. Chodera, *J. Chem. Theory Comput.* **9**, 461 (2013).
- ³⁶The strong performance of the holo estimator implies that the weights are numerically stable. As expected, snapshots drawn from alchemical states near the holo state contribute the largest weights to the holo ensemble (see Fig. S2 in the [supplementary material](#)).
- ³⁷C. Oostenbrink and W. F. van Gunsteren, *Proc. Natl. Acad. Sci. U. S. A.* **102**, 6750 (2005).
- ³⁸Z. Jandova, D. Fast, M. Setz, M. Pechlaner, and C. Oostenbrink, *J. Chem. Theory Comput.* **14**, 894 (2018).