Accurate calculation of absolute protein-ligand binding free energies

Wei Chen[1], Di Cui[1], Robert Abel[1], Richard A. Friesner[2], and Lingle Wang[1,*]

[1]Schrödinger, Inc., 1540 Broadway, 24th Floor, New York, NY 10036, United States
[2]Department of Chemistry, Columbia University, New York, NY 10027, United States

[*]Corresponding author. E-mail: lingle.wang@schrodinger.com.

Abstract

In the hit identification stage of drug discovery, a diverse chemical space needs to be explored to identify initial hits. Contrary to empirical scoring functions, absolute protein-ligand binding free energy perturbation (AB-FEP) provides a theoretically more rigorous and accurate description of protein ligand binding thermodynamics and could in principle greatly improve the hit rates in virtual screening. In this work, we describe an implementation of an accurate and reliable AB-FEP method in FEP+. We validated the AB-FEP method on eight congeneric compound series binding to eight protein receptors including both neutral and charged ligands. For ligands with net charges, the alchemical ion approach was adopted to avoid artifacts in electrostatic potential energy calculations. The calculated binding free energies were highly correlated with experimental results with the weighted average of R^2 of 0.55 for the entire dataset and an overall RMSE of 1.1 kcal/mol when protein reorganization effect upon ligand binding was accounted for. Through AB-FEP calculations using apo versus holo protein structures, we demonstrated that the protein conformational and protonation state changes between the apo and holo proteins are the main physical factors contributing to the protein reorganization free energy manifested by the overestimation of raw AB-FEP calculated binding free energies using the holo structures of the proteins. The highly accurate AB-FEP results demonstrated in this work position it as a useful tool to improve the hit rates in virtual screening, thus facilitate hit discovery.

Introduction

A primary objective of small molecule drug discovery is to design compounds that can tightly and selectively bind to a target protein. Accurate calculation of protein-ligand binding free energy is therefore of central importance in computational drug discovery. Benefiting from improved force fields and sampling algorithms and advanced hardware, rigorous free energy calculation by free energy perturbation (FEP) or related methods in explicit solvent simulations has dramatically improved the accuracy and begun to play an increasingly important role in modern computational drug discovery projects.[1,2,3] As an example, the FEP+ implementation of free energy calculations[4] has demonstrated a high level of accuracy in relative protein-ligand binding free energy calculations (RB-FEP), with an overall root-mean-square error (RMSE) of about 1.1 kal/mol over a broad range of protein targets and ligand series.[5,6] It also enables the accurate modeling of very complex perturbations including scaffold hopping,[7] macrocyclization,[8] net-charge changes,[9] fragment linking,[10] and linker enumeration. The high reliability and accuracy across a broad range of complex chemical modifications has also been validated in a large number of prospective studies in active drug discovery projects, positively impacting the projects through faster identification of novel potent chemical matters.[3] However, relative-binding free energy calculation through RB-FEP can only be applied on congeneric series of ligands with similar binding modes and scaffolds, limiting its application to the hit-to-lead and lead optimization stages of drug discovery where the structure of the binding complex of an initial reference ligand with the target receptor is known.

During the hit discovery stage of drug discovery, a diverse chemical space needs to be explored to identify initial hits. As the current best practices, empirical scoring functions are used in virtual screening to dock a large library of compounds. Due to the limited accuracy of empirical scoring functions, the hit rate in virtual screening is usually very low, about 1-2% on average, with only a few confirmed hits for most screenings and sometimes not a single hit for challenging targets.[11,12] Absolute protein-ligand binding free energy calculation through free energy perturbation (AB-FEP) provides a theoretically more rigorous description of protein-ligand binding thermodynamics, offering hope to dramatically improve the hit rates by rescoring the top compounds in virtual screening. However, due to the complexity for the implementation of AB-FEP methods, the difficulty to converge the simulations to a level useful in practical applications, and the large computational cost associated in these calculations, accurate and reliable calculations of protein-ligand binding free energies through AB-FEP in practical virtual screening for hit discovery have not been reported yet.

The first AB-FEP method has been proposed decades ago through the construction of a non-physical alchemical pathway.[13,14,15] The method involves the calculation of the free energy to transfer the ligand from the solution to the gas phase, and the free energy to transfer the ligand from the protein binding pocket to the gas phase. The difference in the above two free energies corresponds to the absolute binding free energy of the ligand. Initial applications of the method have been focused on model systems, such as fragments binding to T4 lysozyme[16,17,18] and FK506-binding protein[19] and host-guest systems,[20,21] resulting in a reasonable accuracy with RMSE between the calculated and experimental binding free energies of 2-3 kcal/mol, with the goal to showcase the feasibility of the method. Due to the large complexity and computational cost, over a very long period of time, the majority of AB-FEP literature on real protein-ligand systems only reported calculations of a small number of compounds binding to a handful of protein receptors, including FK506-binding proteins[22,23,24] and bromodomain-containing proteins.[25,26,27] The accuracies of the reported calculations varied, with the RMSE between calculation and experiment ranging from 1 to 3 kcal/mol.

Until very recently, with the great increase of computer power, AB-FEP calculations on a medium-to-large number of drug-like compounds for multiple protein targets were reported. In one such study, Li et al. performed AB-FEP simulations on 7 proteins and 101 congeneric ligands and reported surprisingly high accuracy with RMSEs of 0.6-1.5 kcal/mol and R^2 of 0.5-0.9 between calculated and experimental binding affinities, though the reported high accuracy was partially due to the removal of the systematic difference between calculated and experimental binding free energies (the RMSEs for the raw data are 0.9-5 kcal/mol).[28] In another study, Lin et al. performed AB-FEP simulations for 5 proteins and 105 congeneric ligands, and obtained RMSEs of 3-6 kcal/mol and R^2 of 0.6-0.8.[29] Another study from Khalak et al. employed a non-equilibrium method incorporating the apo states of the proteins in the AB-FEP calculations, and obtained RMSEs of 0.8-3 kcal/mol and R^2 of 0.02-0.76 on 7 proteins and 128 congeneric ligands.[30] In all these studies, AB-FEP was applied on congeneric ligands where RB-FEP worked better, with the goal

to validate the implementation. In addition, except for the work of Li et al.[28] with a few ligands carrying a net charge, all the other studies have focused exclusively on neutral ligands. Therefore, it is not clear how these methods would work in a practical virtual screening setting to score ligands with diverse structures and binding modes, particularly for ligands with net charges where the finite size effects[31] and strong electrostatic interactions between the protein and ligand are known to be prohibitively difficult to converge.

In this paper, we report an accurate and reliable AB-FEP method implemented in the FEP+ program.[4] We validated the implementation on all the eight protein systems and 199 ligands from our previous RB-FEP paper,[5] which was later used as the benchmark systems for free energy calculations by many groups. Four of the systems contain neutral ligands while the other four include ligands with net charges. To calculate the binding free energies of the charged ligands, the alchemical ion approach used for the charge-changing perturbations in RB-FEP[9] was adopted in AB-FEP as well. Using the holo conformation of the protein receptor, the raw AB-FEP calculated binding free energies were systematically more negative (favorable) than experiment. This is expected since the apo versus holo protein conformational and/or protonation/tautomeric state changes induced upon ligand binding was not sampled in the relatively short simulations. This is further verified through AB-FEP calculations using the apo conformation of the protein, where the calculated free energies are slightly more positive than experiment, providing evidence to elucidate a long standing puzzle in the literature regarding how AB-FEP calculations should be interpreted for real protein-ligand systems. After removing the systematic shift between the calculated and experimental binding free energies to account for the protein reorganization effect, the overall RMSE between calculation and experiment for the entire dataset is 1.1 kcal/mol with a weighted average R^2 of 0.55, comparable with the RB-FEP results on the same dataset (RMSE of 0.9 kcal/mol and R^2 of 0.56). Comparing with previous AB-FEP studies[28,29,30] which have reported results on some of the systems, the accuracy of our results are comparable or better.

Methods

Double decoupling method for AB-FEP

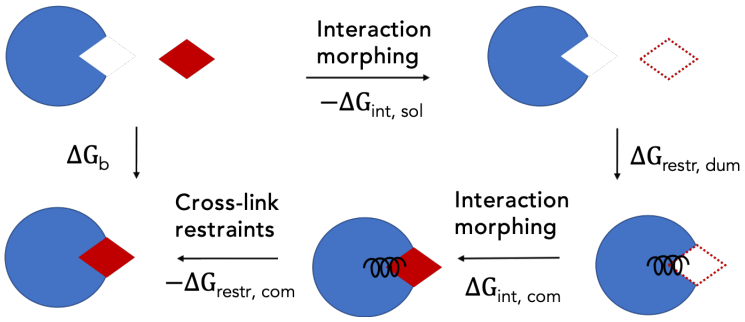

Figure 1: Thermodynamic cycle for AB-FEP. The binding of the ligand to the protein receptor is decomposed into a few alchemical steps similar to the originally proposed double decoupling scheme.[15] Starting from the physical ligand in water, the vdw and electrostatic interactions within the ligand and between the ligand and water are slowly turned off, i.e., $-\Delta G_{int,sol}$ in the cycle; then the relative position and orientation of the dummy ligand with respect to the protein binding pocket are restrained through a set of cross-link restraints ($\Delta G_{restr,dum}$); in the third step, the intra-ligand and ligand-environment vdw and electrostatic interactions for the restrained ligand are slowly turned on in the protein binding pocket ($\Delta G_{int,com}$) followed by relaxing the cross-link restrains when ligand interactions are fully turned on ($-\Delta G_{restr,com}$). The binding free energy is the sum of the free energies of these processes. In our protocol, the free energies for the two horizontal legs, $-\Delta G_{int,sol}$ and $\Delta G_{int,com}$-$\Delta G_{restr,com}$, are calculated through two independent simulations in solvent and protein complex, respectively, while the free energy of the right vertical leg, $\Delta G_{restr,dum}$, is calculated analytically.

Our AB-FEP protocol is similar to the originally proposed double decoupling scheme[15] with the thermodynamic cycle shown in Fig. 1. The binding of the ligand to the protein receptor is decomposed into a few alchemical steps. First, starting from the physical ligand in water, the vdw and electrostatic interactions within the ligand and between the ligand and water are slowly turned off until the ligand becomes dummy; second, the dummy ligand is attached to the protein binding pocket through a set of cross-link restraints similar to what was proposed by Boresch et al.;[15] in the third step, the intra-ligand and ligand-environment vdw and electrostatic interactions for the restrained ligand are slowly turned on in the protein binding pocket and the cross-link restrains are relaxed after that. The free energy to turn on/off the intra-ligand and ligand-water interactions ($-\Delta G_{int,sol}$) is calculated by one simulation in solvent, the free energy to turn on/off the intra-ligand and ligand-environment interactions and relaxing the restrain potentials in the protein binding pocket ($\Delta G_{int,com}$-$\Delta G_{restr,com}$) is calculated by another simulation in the protein binding pocket, and the free energy to restrain the relative position and orientation of the dummy ligand with respect to the protein binding pocket ($\Delta G_{restr,dum}$,) is calculated analytically with detailed derivations in the following section. The absolute binding free energy (ΔG_b) is the sum of these terms as follows:

$$\Delta G_b = \Delta G_{int,com} - \Delta G_{int,sol} + \Delta G_{restr,dum} - \Delta G_{restr,com} \tag{1}$$

Cross-link restraints for AB-FEP

The relative position and orientation of the dummy ligand with respect to the protein binding pocket is restrained by a set of cross-link restrains originally proposed by Boresch et al.[15] Three protein atoms (a, b, and c) and three ligand atoms (A, B, and C) are selected for setting up cross-link restraints (Fig. 2). One distance r_{aA}, two angles θ_{Aab} and θ_{aAB}, and three dihedral angles ϕ_{baAB}, ϕ_{Aabc} and ϕ_{aABC}, are restrained by harmonic potentials. The free energy difference for adding cross-link restraints can be calculated as

$$\Delta G_{restr,dum} = -k_B T ln \frac{Z_{CL}}{Z_P Z_L}, \tag{2}$$

where Z_P, Z_L, and Z_{CL} are the partition functions for the protein, the free dummy ligand, and the dummy ligand-protein complex with the above cross-link restrains, respectively (right vertical leg in Fig. 1). Since

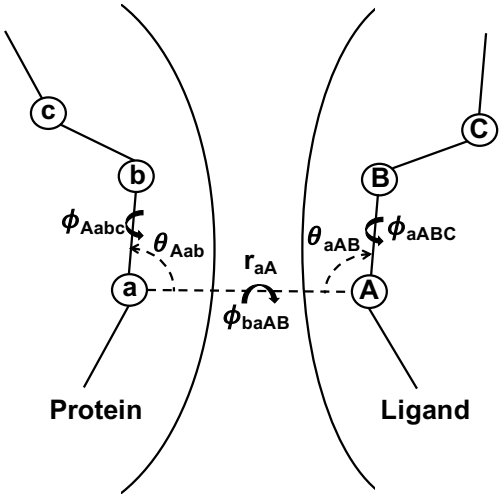

Figure 2: Cross-link restraints between the protein and ligand for AB-FEP. Following the work of Boresch et al.,[15] three protein atoms (a, b, and c) and three ligand heavy atoms (A, B, and C) are selected. A distance r_{aA}, two angles θ_{Aab} and θ_{aAB}, and three dihedral angles ϕ_{baAB}, ϕ_{Aabc} and ϕ_{aABC} are restrained by harmonic potentials.

the dummy ligand does not have interactions with the protein and water except for the cross-link restrains, an analytical formula for the above free energy can be easily derived:

$$\Delta G_{restr,dum} = -k_B T ln \frac{Z_r}{8\pi^2 V} = -k_B T ln \frac{Z^{dist}_{r,aA} Z^{ang}_{r,aAB} Z^{ang}_{r,Aab} Z^{dihed}_{r,baAB} Z^{dihed}_{r,Aabc} Z^{dihed}_{r,aABC}}{8\pi^2 V}, \tag{3}$$

where

$$\begin{aligned} Z^{dist}_r &= \int_0^\infty exp(-\beta K_r (r - r_0)^2) r^2 \, dr \\ &= \frac{r_0}{2\beta K_r} exp(-\beta K_r r_0^2) + \frac{\sqrt{\pi}}{4\beta K_r \sqrt{\beta K_r}} (1 + 2\beta K_r r_0^2)(1 + erf(\sqrt{\beta K_r} r_0), \end{aligned} \tag{4}$$

$$Z^{ang}_r = \int_0^\pi exp(-\beta K_\theta (\theta - \theta_0)^2) sin\theta d\theta \approx \sqrt{\frac{\pi}{\beta K_\theta}} exp(-\frac{1}{4\beta K_\theta}) sin\theta_0, \tag{5}$$

$$Z^{dihed}_r = \int_{\phi_0 - \pi}^{\phi_0 + \pi} exp(-\beta K_\phi (\phi - \phi_0)^2) d\phi = \sqrt{\frac{\pi}{\beta K_\phi}} erf(\pi \sqrt{\beta K_\phi}), \tag{6}$$

and V is 1660 $\mathring{A}^3$ for the standard state. In Eqns. 4-6, K_r, K_θ and K_ϕ are the force constants and r_0, θ_0 and ϕ_0 are equilibrium values for the distance, angle, and dihedral restraints, respectively.

Structure preparation

The input structures for AB-FEP calculations were obtained from the Protein Data Bank (PDB)[32] and prepared by the Protein Preparation Wizard[33] in Maestro[34] with the default settings. The same crystal structures as that used for RB-FEP benchmark[5] were also used for the AB-FEP calculations: 4DJW[35] for BACE1, 1H1Q[36] for CDK2, 2GMX[37] for JNK1, 4HW3[38] for MCL1, 3FLN for P38, 2QBS[39] for PTP1B, 2ZFF for Thrombin and 4GIH[40] for TYK2. For seven of the systems, apo protein structures were available in PDB, and the following structures were used for AB-FEP calculations using apo structures: 1SGZ[41] for BACE1, 1H27[42] for CDK2, 3O17 for JNK1, 6QB3[43] for MCL1, 1WFC[44] for P38, 2CM2[45] for PTP1B and 3D49 for Thrombin. The binding poses of compounds were taken from the previous RB-REP work.[5]

Simulation details

OPLS4 force field[6] was used for all simulations. In each AB-FEP simulation, the protein-ligand complex was solvated in an orthorhombic SPC[46] water box. The buffer width was 5 Å for a neutral ligand and 8 Å for a charged ligand (a charged ligand was defined as a ligand with a net charge or charged groups). For charged ligands, the protein ligand complexes were neutralized by counter-ions and additional salt ions of 150 mM were added to mimic the buffer solution of experimental assay. The systems were then relaxed by a series of short molecular dynamics (MD) relaxations including: (1)100 ps Brownian Dynamics at 10 K with solute heavy atoms restrained (force constant 50 kcal/mol/Å^2) to relieve minor steric clashes; (2) 12 ps NVT simulation at 10 K with solute heavy atoms restrained; (3) 20 ps Grand Canonical Monte Carlo (GCMC) μVT simulations[47] at 300 K with solute heavy atoms restrained to solvate the binding pocket; (4) 20 ps GCMC μVT simulation at 300 K with protein backbone heavy atoms restrained. After relaxation, a 1-ns GCMC μVT simulation at 300 K was performed with protein backbone heavy atoms restrained.

To identify the optimal set of atoms for the protein-ligand cross-link restraints, the interactions between the protein and ligand during the 1-ns MD relaxation were analyzed and the set of atoms with most frequent hydrogen bond or salt-bridge interactions in MD were selected to be restrained. In particular, the hydrogen bond and salt bridge interactions between the protein and ligand occurred in any frame of the MD trajectories were identified, and their frequencies were collected. The frequencies of atom-based interactions were then summed and assigned to a ligand non-terminal heavy atom-protein residue pair as follows: (1) any protein atom in an interaction is assigned to the corresponding residue it belongs to; (2) any ligand atom in an interaction is assigned to a non-terminal heavy atom, which is bonded to at least two other heavy atoms. Say for example a terminal oxygen atom in a carboxylate group of the ligand forming a hydrogen bond with the protein is assigned to the carbon atom of the carboxylate group. If multiple ligand atom-protein residue pairs have hydrogen bond or salt bridge interaction frequencies of at least 50%, the ligand atom closest to the centroid of the ligand was selected as one of the atoms for the cross-link restraints (terminal groups in the ligand like SO_2 that can easily flip orientation were excluded). After the selection of the anchoring atom in the ligand (atom A in Fig. 2), two other ligand heavy atoms that atom A is bonded to (B and C in Fig. 2), and the three backbone atoms (N, C_α and C) of the protein residue forming hydrogen bond and/or salt-bridge with atom A (atoms a, b and c in Fig. 2) were selected for the cross-link restrains. To avoid co-linear geometry leading to the singularity in the dihedral angle restraints, we limited the set of restrained atoms to those with the four angles (θ_{Aab}, θ_{aAB}, θ_{abc} and θ_{ABC}) between 45 and 135 degrees. If none of the ligand atom-protein residue pairs with interaction frequency of at least 50% satisfied the above criteria, the ligand heavy atom closest to the centroid of the ligand, two of its bonded heavy atoms (atoms A, B and C in Fig. 2) and three protein C_α atoms (atoms a, b and c in Fig. 2), which satisfy the above angle requirement, were selected to be restrained. The force constant was 1 kcal/mol/Å^2 for the distance restraint, and 40 kcal/mol/rad^2 for the angle and dihedral restraints.

After the MD relaxation and the selection of atoms for the cross-link restraints, a representative structure from the MD trajectory was used as the input for the following FEP simulations. To select the representative structure, the mean value of each rotatable bond in the ligand sampled during the MD trajectory was calculated, and the representative structure had the ligand torsions closest to the corresponding mean values. For charged ligands, counter-ions and additional salt ions of 150 mM were again added the same as that in the MD relaxation, and the same alchemical ion approach introduced for charged perturbation in RB-FEP[9] were also adopted for AB-FEP. A total of 68 and 108 λ windows were used for neutral and charged ligands, respectively. Each replica was run for 5 ns. For the solvent leg FEP, the ligand was extracted from the representative structure selected above and then solvated in a SPC water box with 10 Å buffer width. Again, counter-ions and additional salt ions of 150 mM were added and the alchemical ion approach was utilized for charged ligands. 60 λ windows were used for all ligands and each replica was run for 5 ns.

To benchmark the accuracy of AB-FEP versus RB-FEP on the congeneric series of ligands, we also performed RB-FEP calculations with OPLS4 on the eight congeneric series of ligands taken from Wang et al.,[5] and the default RB-FEP protocol as detailed in that paper was used for the RB-FEP calculations. The RB-FEP simulations lasted for 5 ns per replica.

To calculate the pK_a's of aspartic acids in the binding pocket of BACE1, protein-stability FEP[48] implemented in FEP+[4] was used. The free energy to mutate from the neutral to the charged ASP was calculated in the protein environment and for an isolated residue in solvent, and the difference corresponds to the shift of pKa due to protein environment, i.e., $\Delta\Delta$G = RTln10 (pK_a- 3.67), where 3.67 is the pK_a of an isolated Asp residue in solvent.[49] 24 λ windows were used and each replica was run for 20 ns.

Results and Discussion

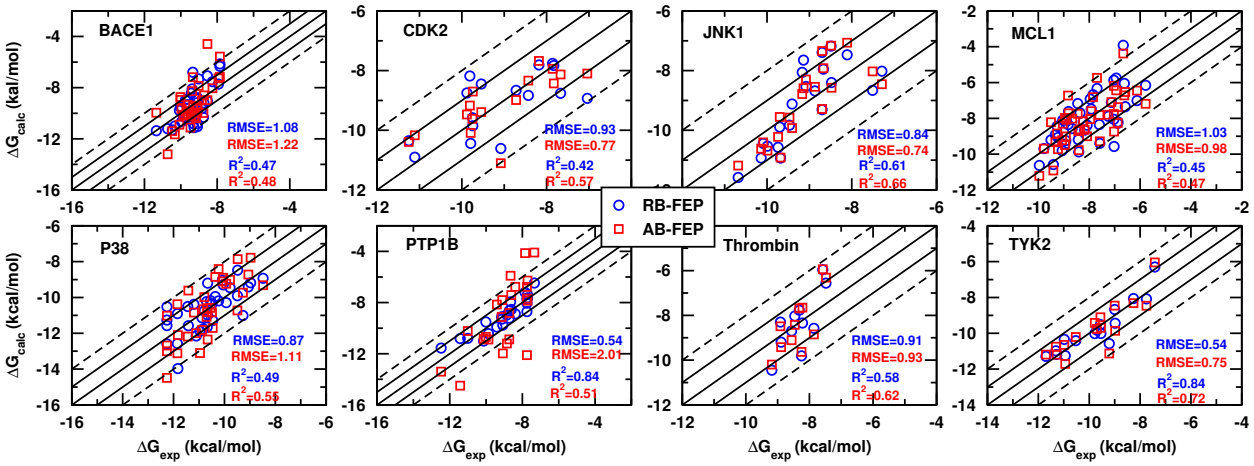

Figure 3: Comparison of calculated binding free energies versus experiment for both AB-FEP and RB-FEP. Blue circles and red squares are for RB-FEP and AB-FEP, respectively. The results of the first repeat of the AB-FEP simulation are shown, and the results from the other two repeats were reported in SI. The raw AB-FEP binding free energies were shifted to take into account the protein reorganization effect as detailed in the main text. Solid diagonal lines indicate the region within 1 kcal/mol of experimental values, while dashed diagonal lines indicate the region within 2 kcal/mol of experimental values. RMSE and R^2 are labeled.

We tested the AB-FEP method on the entire dataset from the previous RB-FEP benchmark[5] including a total number of 199 compounds binding to eight different protein systems. To gauge the convergence and consistency of the calculations, three replicas of independent AB-FEP simulations with different random seeds were performed for each ligand. The AB-FEP performance on these eight systems were presented in Fig. 3 (the results from the first replica were shown for simplicity, and results from the other two replicas are shown in Figs. S1 and S2) and summarized in Table I. Four series of ligands binding to BACE1, MCL1, PTP1B and Thrombin are charged while the other four are neutral. For comparison, the results from the previously established RB-FEP method[5] were also shown. The raw AB-FEP results for all these systems were more negative than experimental values, reflecting the miss of the protein reorganization effect for the structural difference between the apo protein and the holo complex due to limited sampling in the short AB-FEP simulations. For proper comparison between AB-FEP results and experiment, the protein reorganization contribution to the binding free energy for each system was estimated as the difference between the average of experimental binding free energies for all the ligands in that system and the average of the raw AB-FEP binding free energies for the same set of ligands. The estimated protein reorganization contribution for each system was added to the raw AB-FEP values for the final comparison with experimental data. The estimated protein reorganization contributions ($\Delta G_{\mathrm{prot-reorg}}$) are shown in Table I, and are ranged between 1.22 kcal/mol for CDK2 and 9.97 kcal/mol for PTP1B.

Taking protein reorganization contribution into account, the AB-FEP calculated binding free energies for all systems agree very well with experiment (Fig. 3), with RMSE between calculation and experiment ranged between 0.74 kcal/mol for TYK2 system and 1.78 kcal/mol for PTP1B, and R^2 ranged between 0.44 for CDK2 and 0.74 for TYK2 (Table I). The overall RMSE and weighted average R^2 from AB-FEP for the entire dataset is 1.13 kcal/mol and 0.55, respectively, slightly worse than the RB-FEP performance (overall RMSE and weighted average R^2 of 0.90 kcal/mol and 0.56, respectively), due to much larger configurational space to sample in AB-FEP. It should be noted that the protein reorganization contribution is a constant among all the ligands binding to the same protein conformation, thus does not affect the R^2 and rank ordering of the predictions. Our estimate of protein reorganization contribution minimizes the RMSE between prediction and experiment, but would not affect the prospective usage of AB-FEP for compound selection and prioritization.

Table I: Statistics of ABFEP and RB-FEP results for congeneric compound series

Protein	# of ligands	Ligand charge	RB-FEP		[a]AB-FEP with holo-protein			[b]AB-FEP with apo-protein		
			RMSE	R^2	[d]RMSE	R^2	[c]$\Delta G_{\mathrm{prot-reorg}}$	[d]RMSE	R^2	[c]$\Delta G_{\mathrm{prot-reorg}}$
BACE1	36	+1	1.08	0.47	1.21 ± 0.07	0.44 ± 0.04	2.91 ± 0.28	0.89 ± 0.06	0.26 ± 0.06	[e]-1.61 ± 0.27
CDK2	16	0	0.93	0.42	0.84 ± 0.06	0.50 ± 0.07	1.22 ± 0.20	1.05 ± 0.06	0.33 ± 0.10	-0.59 ± 0.10
JNK1	21	0	0.84	0.61	0.85 ± 0.13	0.63 ± 0.02	3.50 ± 0.12	0.64 ± 0.09	0.69 ± 0.09	2.72 ± 0.12
MCL1	42	-1	1.03	0.45	0.95 ± 0.06	0.53 ± 0.07	3.65 ± 0.15	1.03 ± 0.06	0.33 ± 0.04	2.48 ± 0.12
P38	34	0	0.87	0.49	1.09 ± 0.10	0.58 ± 0.06	5.27 ± 0.09	1.02 ± 0.10	0.46 ± 0.05	-1.51 ± 0.09
PTP1B	23	-2	0.54	0.84	1.78 ± 0.20	0.55 ± 0.03	9.97 ± 0.21	1.20 ± 0.06	0.50 ± 0.03	-1.12 ± 0.07
Thrombin	11	+1	0.91	0.58	1.01 ± 0.07	0.57 ± 0.07	1.82 ± 0.14	1.03 ± 0.19	0.48 ± 0.12	2.08 ± 0.08
TYK2	16	0	0.54	0.84	0.74 ± 0.02	0.74 ± 0.04	3.54 ± 0.17			
ALL	199		0.90	[f]0.56	1.13 ± 0.03	[f]0.55 ± 0.003		0.99 ± 0.03	[f]0.41 ± 0.02	

[a]AB-FEP simulations using the crystal structures of holo-proteins. [b]AB-FEP simulations using the crystal structures of apo-proteins. [c]$\Delta G_{\mathrm{prot-reorg}}$: protein reorganization contribution estimated as the difference between the average of experimental binding free energies and the average of the raw AB-FEP binding free energies for all ligands in a system. [d]RMSE was calculated after $\Delta G_{\mathrm{prot-reorg}}$ was added to the raw AB-FEP binding free energies. [e]$\Delta G_{\mathrm{prot-reorg}}$ for BACE1 using the apo protein structure was calculated after adding pK_{a} correction of 3.57 kcal/mol (see the main text for details) to raw AB-FEP binding free energies. [f]Average R^2 weighted by the numbers of ligands in the eight systems for AB-FEP with holo-proteins or the seven systems for AB-FEP with apo-proteins. The AB-FEP simulations were repeated for three times with different random seeds. The mean and standard deviation among the three repeats are reported in the table.

Among these eight systems, the protein reorganization contribution for PTP1B system of 9.97 kcal/mol is significantly larger than the other systems (in the range of 1.22 to 3.65 kcal/mol), and the overall RMSE for PTP1B system of 1.78 kcal/mol is also significantly worse than the others (in the range of 0.74 to 1.21 kcal/mol). To understand the possible reasons leading to the difference in the performance of AB-FEP on PTP1B versus the other systems, we searched for an apo crystal structure of PTP1B (PDB ID 2CM2),[45] and compared it with the holo crystal structure (PDB ID 2QBS[39]) used in the AB-FEP simulations (Fig. 4). The large conformational changes in the WPD-loop (Thr177-Pro188) of the binding pocket from an open conformation for the apo protein to the closed conformation for the holo protein[45] could potentially explain the large estimated protein reorganization contribution for this system. To validate this hypothesis, we repeated the AB-FEP simulations for the PTP1B ligands using the apo conformation of the receptor, and the resulting calculated binding free energies were more positive than experiment values by about 1 kcal/mol (Table I). This again is expected as the closed WPD-loop conformation is preferred in the presence of ligands but could not be sampled in the AB-FEP simulations using the apo structure of the protein with the WPD-loop in the open conformation. These results validated our hypothesis that the more negative binding free energies from the raw AB-FEP results using the holo conformation of the protein as compared to experiment is due to the protein reorganization contribution between the apo and holo conformations of the receptor not sampled in the AB-FEP simulations.

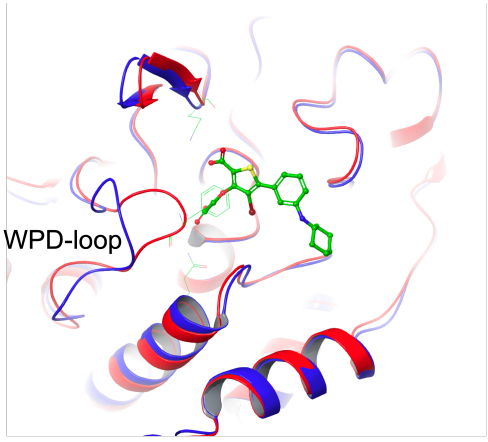

Figure 4: Comparison of holo- and apo-structures of PTP1B. The holo-crystal structure of PTP1B (red, PDB ID 2QBS[39]) is aligned to the apo-crystal structure (blue, PDB ID 2CM2[45]). The WPD-loop (Thr177-Pro188), which has the largest difference, is labeled.

Surprisingly, the RMSE between the AB-FEP calculated and experimental binding free energies (after accounting for the protein reorganization effect) for PTP1B ligands is greatly reduced to 1.2 kcal/mol

using the apo conformation of the receptor (so does the variance of the AB-FEP results among three independent repeats shown in Table I), suggesting that the simulations with the apo-structure converged much faster. We attribute the faster convergence of the calculated free energies using the apo structure to the much weaker hydrogen bond and salt bridge interactions between the two carboxylate groups conserved in all the ligands and surrounding charged residues in the binding pocket when the WPD-loop is open. In the holo structure, the two carboxyl groups in the ligands form multiple strong interactions with the protein including three salt bridges with Arg221 and Lys120, which were persistent in the AB-FEP simulations (Fig. 5). In constrast, the two salt bridge interactions with Arg221 were partially lost in the open conformation of the WPD-loop in the apo structure when exposed to solvent.

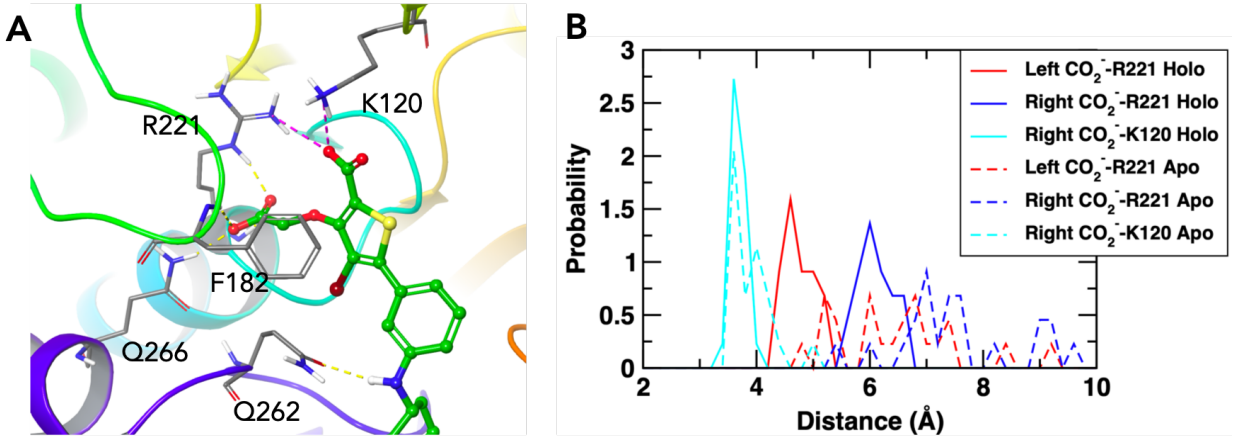

Figure 5: Salt bridge and hydrogen bond interactions between the carboxyl groups of the ligand and PTP1B. **A.** Structure of the binding pocket of PTP1B (PDB ID 2QBS[39]). Three salt bridges are formed between the carboxyl groups of the ligand and PTP1B: the first salt bridge is between the carboxyl group on the left and Arg221; the second is between the carboxyl group on the right and Arg221; and the third is between the carboxyl group on the right and Lys120. **B.** Probability distributions of distances for the three salt bridges between the ligand and protein in the AB-FEP simulation of one ligand. Red lines: distance between the left carboxyl group of the ligand and Arg221 (measured between C22 of the ligand and C_ζ of Arg221), blue lines: distance between the right carboxyl group of the ligand and Arg221 (measured between C6 of the ligand and C_ζ of Arg221), cyan lines: distance between the right carboxyl group of the ligand and Lys120 (measured between C6 of the ligand and N_ζ of Lys120). Solid lines are for the simulations with the holo structure while dashed lines for the simulations with the apo structure. The distributions were calculated based on the first repeat of the AB-FEP simulations of holo- and apo-PTP1B.

Besides PTP1B, we also investigated the possible physical factors contributing to the protein reorganization effects observed for other systems, although the magnitudes are much smaller than that for PTP1B (Table I). For seven out of the eight systems except for TYK2, we found apo crystal structures of the proteins from PDB. Repeating the AB-FEP simulations using the apo structures, the raw AB-FEP results of all systems except for Thrombin consistently got more positive as compared to that from the holo crystal structures (Table I), indicating that the conformational differences between the apo and holo structures are indeed the dominant factors contributing to the observed protein reorganization effects from the AB-FEP calculations. For Thrombin, the apo and holo structures are almost identical, and the AB-FEP results were not affected by the choice of input protein structures. For CDK2 and P38, the P-loop conformations are slightly different between the apo and holo structures, and similar to PTP1B the raw AB-FEP results using the apo protein structures were slightly more positive than the experimental results. Therefore, the P-loop conformational changes between the apo and holo structures are fully responsible to the observed protein reorganization effects for these two systems.

For BACE1, AB-FEP binding free energies using the apo protein structure are slightly more positive than that using the holo protein structure ($\Delta G_{\mathrm{prot-reorg}}$ is 2.91 and 1.96 kcal/mol for simulations with holo and apo structures, respectively), but do not fully explain the protein reorganization effects reflected in the AB-FEP simulations. Inspecting the binding pocket, we notice that two aspartic acids are located at the BACE1 binding pocket, forming salt bridges with the lgiand (Fig. 6). Because of the strong salt bridge

interactions, the two Asps should be deprotonated in the ligand bound form, which was the state used in the AB-FEP simulations. However, in the apo form without ligand, the two ASPs may prefer an alternative protonation state. To verify that, we performed protein FEP to calculate pK_a's of the two Asps. The resulting pK_a of Asp93 is 7.6 in the apo form, higher than experimental pH of 5, indicating the protonated form is preferred in the apo state. The corresponding penalty for the change of the protonation state from protonated form in the apo state to the deprotonated form of the ligand bound state, estimated to be RTln10(pK_a- pH) = 3.6 kcal/mol, fully explained the observed protein reorganization effects in the AB-FEP simulations.

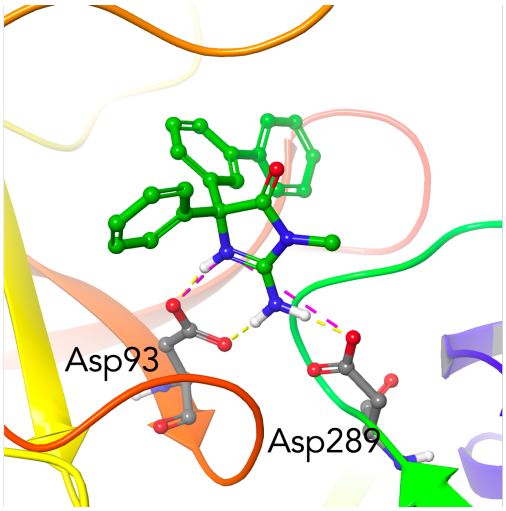

Figure 6: Aspartic acids in the binding pocket of BACE1. Two Asps are located at the binding pocket of BACE1 (PDB ID 4DJW[35]) and forming salt bridges with the ligand.

For JNK1, MCL1 and Thrombin, AB-FEP results using the apo protein structure are still slightly more negative than the experiments. Except for the conformational and protonation/tautomeric state changes between the apo and holo structures explored above, other factors could also contribute to the systematic overestimation of binding free energies from AB-FEP as compared to experiment. For example, some experimental assays measured the IC_{50} instead of K_i, and the binding free energies converted directly from the experimental IC_{50}s using $\Delta G=RTln(IC_{50})$ could systematically underestimate the real binding free energies. For the JNK1 system, the experimental IC_{50} and K_i were available for six compounds (Table S1),[37] and the differences in the free energies converted from IC_{50} versus that from K_i can be as large as 1.9 cal/mol, comparable to the magnitude of the observed shift between the raw AB-FEP results and IC_{50} based experimental free energies (Table I). Post-translational modifications and the different protein constructs and/or buffer solutions in experiment versus that in the AB-FEP simulations could also contribute to the systematic differences between AB-FEP results and experimental measurements.

Several groups have previously reported AB-FEP results on some of the systems we tested here.[28,29,30] The majority of these earlier studies were focused on the four protein systems with neutral ligands (CDK2, JNK1, P38 and TYK2), and only one paper reported results on Thrombin with charged ligands. Comparing with these earlier studies (Table II), after taking into account the protein reorganization effect for each method in the same way as we did above to remove the systematic difference between the raw AB-FEP results and experimental measurements, for the four systems with neutral ligands, our AB-FEP yielded RMSEs of 0.74-1.09 kcal/mol and R^2 of 0.5-0.74, while the work of Lin et al.[29] gave RMSEs of 0.86-1.11 kcal/mol and R^2 of 0.57-0.68, and the work from Khalak et al.[30] resulted in RMSEs of 0.76-1.17 kcal/mol and R^2 of 0.19-0.45, both slightly worse than our results. The protein reorganization effects from Khalak et al.[30] were very small for three systems, possibly because they incorporated the apo crystal structures into their AB-FEP calculations. However, surprisingly, their raw AB-FEP results for TYK2 (no apo crystal structure was available) were systematic more positive than experiments, which is opposite than expected based on possible protein conformational changes between apo and holo structures. Li et al.[28] reported AB-FEP results on the three systems with neutral ligands (CDK2, JNK1 and TYK2) and one with charged ligands (Thrombin). Their results for three systems (CDK2, JNK1 and Thrombin) were comparable to ours (RMSEs of 0.59-0.75 kcal/mol and R^2 of 0.47-0.79 versus RMSEs of 0.84-1.01 kcal/mol and R^2 of 0.5-0.63 from

ours), but were much worse than ours for TYK2 (RMSE of 1.32 kcal/mol and the R^2 of 0.52 as compared to RMSE of 0.74 kcal/mol and R^2 of 0.74 from ours). In addition, their AB-FEP results on TYK2 ligands are more positive than experiment, contradictory to the expected protein reorganization effect between the apo and holo conformations. Overall, our AB-FEP is competitive or slightly better in accuracy than previous benchmarks on the same systems and cover broader sets of ligands.

Table II: Comparison of AB-FEP results with previous studies

Protein	AB-FEP			Work of Lin et al.[29]			Work of Khalak et al.[30]			Work of Li et al.[28]		
	[b]RMSE	R^2	[a]$\Delta G_{\mathrm{prot-reorg}}$	[b]RMSE	R^2	[a]$\Delta G_{\mathrm{prot-reorg}}$	[b]RMSE	R^2	[a]$\Delta G_{\mathrm{prot-reorg}}$	[b] RMSE	R^2	[a]$\Delta G_{\mathrm{prot-reorg}}$
BACE1	1.21	0.44	2.91									
CDK2	0.84	0.50	1.22	1.11	0.62	3.50	1.13	0.19	-0.43	0.74	0.79	4.05
JNK1	0.85	0.63	3.50	1.03	0.68	4.61	0.76	0.45	0.32	0.75	0.49	2.97
MCL1	0.95	0.53	3.65									
P38	1.09	0.58	5.27	0.86	0.57	6.24	0.95	0.22	-0.03			
PTP1B	1.78	0.55	9.97									
Thrombin	1.01	0.57	1.82							0.59	0.47	5.02
TYK2	0.74	0.74	3.54	0.87	0.66	3.40	1.17	0.27	-2.59	1.32	0.52	-3.27

[a]$\Delta G_{\mathrm{prot-reorg}}$: protein reorganization contribution estimated as the difference between the average of experimental binding free energies and the average of the raw AB-FEP binding free energies for all ligands in a system. [b]RMSE was calculated after $\Delta G_{\mathrm{prot-reorg}}$ was added to the raw AB-FEP binding free energies.

Conclusions

We reported an implementation of the AB-FEP method in the FEP+ program.[4] The AB-FEP protocol utilizes the double decoupling scheme (Fig. 1), where the free energy to annihilate the ligand from the protein binding pocket into a restrained dummy ligand and the free energy to annihilate the ligand in solvent were calculated separately in two simulations. The translational and orientational entropy due to the restraints between the dummy ligand and the protein binding pocket (Fig. 2) was calculated by an analytical formula (Eqn. 3). For ligands with net charges, the alchemical ion approach introduced in the charge changing perturbations in RB-FEP[9] was adopted to resolve the finite-size effects and more replicas were used to converge the electrostatic interactions.

We validated our AB-FEP protocol on the entire dataset from our previous RB-FEP benchmark[5] including a total number of 199 ligands binding to eight protein targets. Different from previous AB-FEP publications focusing mainly on the neutral ligands, four of our validation systems have charged ligands. The overall RMSE between AB-FEP calculated and experimental binding free energies on the entire dataset is 1.13 kcal/mol and weighted average R^2 is 0.55, slightly worse than the well established RB-FEP method[5] with RMSE of 0.9 kcal/mol and R^2 of 0.56 (Fig. 3 and Table I). This is expected due to the much larger configurational space to sample in AB-FEP as compared to RB-FEP. Comparing with earlier AB-FEP publications,[29,30,28] which reported results on some of these systems, our AB-FEP results are comparable or slightly better (Table II).

It should be noted that the raw AB-FEP results are in general more negative than the experimental binding free energies for real protein-ligand systems. This is because the holo complexes with ligands bound most often have very different conformations and/or protonation states than the apo proteins, and the conformational and/or protonation state changes upon ligand binding, i.e., the protein reorganization contribution, can not be sampled in the short AB-FEP simulations. We verified this hypothesis by comparing AB-FEP calculations using the holo and apo protein structures and showing that AB-FEP calculated binding free energies are more positive using the apo protein structures than using the holo structures (Table I). In addition, for four of the systems, with the apo protein structures and after taking into account the protonation state changes between the apo and holo structures, AB-FEP calculated binding free energies are more positive than experiment, suggesting that the conformational and protonation state changes are the major factors contributing to the protein reorganization free energy. The hypothesis that the miss of protein reorganization free energy is the main reason for the systematic difference between AB-FEP calculation and experimental binding free energy was suggested in the past,[50] and our calculations are the first attempt to validate this hypothesis.

Although the AB-FEP results are usually more negative than experimental measurement, AB-FEP can still give correct rank ordering of ligand binding considering that the protein reorganization contribution is a constant for ligands binding to the same protein conformation. In fact, a unique advantage of AB-FEP as compared to RB-FEP is its ability to rank order compounds with diverse scaffolds and binding modes, with the potential application to separate correct binding poses from the decoy poses or selecting actives from decoys for hit discovery in virtual screening.

While the accuracy of AB-FEP calculations demonstrated in these validations is exciting, a few important notes should be considered in prospective applications. First, due to the miss of protein reorganization effect discussed above, the raw AB-FEP results usually over-estimate the binding free energies of the ligands. In the context of scoring ligands in virtual screening where no binding affinity information is available for any ligand, it should only be used to rank order the ligands. If experimental binding affinities are available for a handful of ligands, the difference in the AB-FEP results versus experimental data on these known ligands can be used to estimate the protein reorganization effect. As we demonstrated in the above validation, the absolute binding free energies of the ligands can be informed after accounting for protein reorganization effect. Second, due to the large configurational space to sample, AB-FEP calculations are computationally much more expensive than RB-FEP. For ranking congeneric series of ligands, RB-FEP is still recommended both for cost benefit and for the superior accuracy. Third, although the convergence of AB-FEP calculations is much improved with our optimized lambda schedules and sampling protocols, ligands with net charges involving strong salt-bridge interactions with the protein may still present challenges in some cases and further enhancements in the protocol is needed to make it more robust.

Notes

The authors declare the following competing financial interest(s): R.A.F. has a significant financial stake in, is a consultant for, and is on the Scientific Advisory Board of Schrödinger, Inc.

Acknowledgements

We thank Dr. Joe Kaus, Dr. Yuqing Deng and Dr. Yujie Wu for their helpful discussion.

References

[1] Abel, R.; Wang, L.; Harder, E. D.; Berne, B. J.; Friesner, R. A. Advancing drug discovery through enhanced free energy calculations. *Acc. Chem. Res.* **2017**, *50*, 1625–1632.

[2] Abel, R.; Wang, L.; Mobley, D. L.; Friesner, R. A. A critical review of validation, blind testing, and real-world use of alchemical protein-ligand binding free energy calculations. *Curr. Top. in Med. Chem.* **2017**, *17*, 2577–2585.

[3] Schindler, C. E. M. et al. Large-scale assessment of binding free energy calculations in active drug discovery projects. *J. Chem. Inf. Model.* **2020**, *60*, 5457–5474.

[4] *Schrödinger Release 2021-3: FEP+*; Schrödinger, LLC: New York, NY, 2021.

[5] Wang, L. et al. Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *J. Am. Chem. Soc.* **2015**, *137*, 2695–2703.

[6] Lu, C.; Wu, C.; Ghoreishi, D.; Chen, W.; Wang, L.; Damm, W.; Ross, G. A.; Dahlgren, M. K.; Russell, E.; Bargen, C. D. V.; Abel, R.; Friesner, R. A.; Harder, E. D. OPLS4: improving force field accuracy on challenging regimes of chemical space. *J. Chem. Theory Comput.* **2021**, *17*, 4291–4300.

[7] Wang, L.; Deng, Y.; Wu, Y.; Kim, B.; LeBard, D. N.; Wandschneider, D.; Beachy, M.; Friesner, R. A.; Abel, R. Accurate modeling of scaffold hopping transformations in drug discovery. *J. Chem. Theory Comput.* **2017**, *13*, 42–54.

[8] Yu, H. S.; Deng, Y.; Wu, Y.; Sindhikara, D.; Rask, A. R.; Kimura, T.; Abel, R.; Wang, L. Accurate and reliable prediction of the binding affinities of macrocycles to their protein targets. *J. Chem. Theory Comput.* **2017**, *13*, 6290–6300.

[9] Chen, W.; Deng, Y.; Russell, E.; Wu, Y.; Abel, R.; Wang, L. Accurate calculation of relative binding free energies between ligands with different net charges. *J. Chem. Theory Comput.* **2018**, *14*, 6346–6358.

[10] Yu, H. S.; Modugula, K.; Ichihara, O.; Kramschuster, K.; Keng, S.; Abel, R.; Wang, L. General theory of fragment linking in molecular design: Why fragment linking rarely succeeds and how to improve outcomes. *J. Chem. Theory Comput.* **2021**, *17*, 450–462.

[11] Murphy, R. B.; Repasky, M. P.; Greenwood, J. R.; Tubert-Brohman, I.; Jerome, S.; Annabhimoju, R.; Boyles, N. A.; Schmitz, C. D.; Abel, R.; Farid, R.; Friesner, R. A. WScore: a flexible and accurate treatment of explicit water molecules in ligand–receptor docking. *J. Med. Chem.* **2016**, *59*, 4364–4384.

[12] Sliwoski, G.; Kothiwale, S.; Meiler, J.; Edward W. Lowe, J. Computational methods in drug discovery. *Pharmacol. Rev.* **2014**, *66*, 334–395.

[13] Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; TiradoRives, J. Efficient computation of absolute free energies of binding by computer simulations: application to the methane dimer in water. *J. Chem. Phys.* **1988**, *89*, 3742.

[14] Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* **1997**, *72*, 1047–1069.

[15] Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. Absolute binding free energies: a quantitative approach for their calculation. *J. Phys. Chem. B* **2003**, *107*, 9535–9551.

[16] Deng, Y.; Roux, B. Calculation of standard binding free energies: aromatic molecules in the T4 lysozyme L99A mutant. *J. Chem. Theory Comput.* **2006**, *2*, 1255–1273.

[17] Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. Predicting absolute ligand binding free energies to a simple model site. *J. Mol. Biol.* **2007**, *371*, 1118–1134.

[18] Boyce, S. E.; Mobley, D. L.; Rocklin, G. J.; Graves, A. P.; Dill, K. A.; Shoichet, B. K. Predicting ligand binding affinity with alchemical free energy methods in a polar model binding site. *J. Mol. Biol.* **2009**, *394*, 747–763.

[19] Pan, A. C.; Xu, H.; Palpant, T.; Shaw, D. E. Quantitative characterization of the binding and unbinding of millimolar drug fragments with molecular dynamics simulations. *J. Chem. Theory Comput.* **2017**, *13*, 3372–3377.

[20] Lee, J.; Tofoleanu, F.; Pickard, F. C.; König, G.; Huang, J.; Damjanović, A.; Baek, M.; Seok, C.; Brooks, B. R. Absolute binding free energy calculations of CBClip host–guest systems in the SAMPL5 blind challenge. *J. Comput. Aided Mol. Des.* **2017**, *31*, 71–85.

[21] Laury, M. L.; Wang, Z.; Gordon, A. S.; Ponder, J. W. Absolute binding free energies for the SAMPL6 cucurbit[8]uril host–guest challenge via the AMOEBA polarizable force field. *J. Comput. Aided Mol. Des.* **2018**, *32*, 1087–1095.

[22] Jayachandran, G.; Shirts, M. R.; Park, S.; Pande, V. S. Parallelized-over-parts computation of absolute binding free energy with docking and molecular dynamics. *J. Chem. Phys.* **2006**, *125*, 084901.

[23] Wang, J.; Deng, Y.; Roux, B. Absolute binding free energy calculations using molecular dynamics simulations with restraining potentials. *Biophys. J.* **2006**, *91*, 2798–2814.

[24] Fujitani, H.; Tanida, Y.; Matsuura, A. Massively parallel computation of absolute binding free energy with well-equilibrated states. *Phys. Rev. E* **2009**, *79*, 021914.

[25] Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Accurate calculation of the absolute free energy of binding for drug molecules. *Chem. Sci.* **2016**, *7*, 207–218.

[26] Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Predictions of ligand selectivity from absolute binding free energy calculations. *J. Am. Chem. Soc.* **2017**, *139*, 946–957.

[27] Gapsys, V.; Yildirim, A.; Aldeghi, M.; Khalak, Y.; van der Spoel, D.; de Groot, B. L. Accurate absolute free energies for ligand–protein binding based on non-equilibrium approaches. *Commun. Chem.* **2021**, *4*, 1–13.

[28] Li, Z.; Huang, Y.; Wu, Y.; Chen, J.; Wu, D.; Zhan, C. G.; Luo, H. B. Absolute binding free energy calculation and design of a subnanomolar inhibitor of phosphodiesterase-10. *J. Med. Chem.* **2019**, *62*, 2099–2111.

[29] Lin, Z.; Zou, J.; Liu, S.; Peng, C.; Li, Z.; Wan, X.; Fang, D.; Yin, J.; Gobbo, G.; Chen, Y.; Ma, J.; Wen, S.; Zhang, P.; Yang, M. A cloud computing platform for scalable relative and absolute binding free energy predictions: new opportunities and challenges for drug discovery. *J. Chem. Inf. Model.* **2021**, *61*, 2720–2732.

[30] Khalak, Y.; Tresdern, G.; Aldeghi, M.; Baumann, H. M.; Mobley, D. L.; de Groot, B.; Gapsys, V. Alchemical absolute protein–ligand binding free energies for drug design. *Chem. Sci.* **2021**,

[31] Rocklin, G. J.; Mobley, D. L.; Dill, K. A.; Hünenberger, P. H. Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: an accurate correction scheme for electrostatic finite-size effects. *J. Chem. Phys.* **2013**, *139*, 184103.

[32] Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

[33] Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **2013**, *27*, 221–234.

[34] *Schrödinger Release 2021-3: Maestro*; Schrödinger, LLC: New York, NY, 2021.

[35] Cumming, J. N. et al. Structure based design of iminohydantoin BACE1 inhibitors: identification of an orally available, centrally active BACE1 inhibitor. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2444–2449.

[36] Davies, T. G. et al. Structure-based design of a potent purine-based cyclin-dependent kinase inhibitor. *Nat. Struct. Biol.* **2002**, *9*, 745–749.

[37] Szczepankiewicz, B. G. et al. Aminopyridine-based c-Jun N-terminal kinase inhibitors with cellular activity and minimal cross-kinase activity. *J. Med. Chem.* **2006**, *49*, 3563–3580.

[38] Friberg, A.; Vigil, D.; Zhao, B.; Daniels, R. N.; Burke, J. P.; Garcia-Barrantes, P. M.; Camper, D.; Chauder, B. A.; Lee, T.; Olejniczak, E. T.; Fesik, S. W. Discovery of potent myeloid cell leukemia 1 (Mcl-1) inhibitors using fragment-based methods and structure-based design. *J. Med. Chem.* **2012**, *56*, 15–30.

[39] Wilson, D. P. et al. Structure-based optimization of protein tyrosine phosphatase 1B inhibitors: from the active site to the second phosphotyrosine binding site. *J. Med. Chem.* **2007**, *50*, 4681–4698.

[40] Liang, J. et al. Lead identification of novel and selective TYK2 inhibitors. *Eur. J. Med. Chem.* **2013**, *67*, 175–187.

[41] Hong, L.; Tang, J. Flap position of free memapsin 2 (β-secretase), a model for flap opening in aspartic protease catalysis. *Biochemistry* **2004**, *43*, 4689–4695.

[42] Lowe, E. D.; Tews, I.; Cheng, K. Y.; Brown, N. R.; Gul, S.; Noble, M. E. M.; Gamblin, S. J.; Johnson, L. N. Specificity determinants of recruitment peptides bound to phospho-CDK2/cyclin A. *Biochemistry* **2002**, *41*, 15625–15634.

[43] Luptak, J.; Bista, M.; Fisher, D.; Flavell, L.; Gao, N.; Wickson, K.; Kazmirski, S.; Howard, T.; Rawlins, P.; Hargreaves, D.; IUCr, Antibody fragments structurally enable a drug-discovery campaign on the cancer target Mcl-1. *Acta Crystallogr. D Struct. Biol.* **2019**, *75*, 1003–1014.

[44] Wilson, K. P.; Fitzgibbon, M. J.; Caron, P. R.; Griffith, J. P.; Chen, W.; McCaffrey, P. G.; Chambers, S. P.; Su, M. S.-S. Crystal structure of p38 mitogen-activated protein kinase. *J. Biol. Chem.* **1996**, *271*, 27696–27700.

[45] Ala, P. J. et al. Structural basis for inhibition of protein-tyrosine phosphatase 1B by isothiazolidinone heterocyclic phosphonate mimetics. *J. Biol. Chem.* **2006**, *281*, 32784–32795.

[46] Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Springer, Dordrecht, 1981; pp 331–342.

[47] Ross, G. A.; Russell, E.; Deng, Y.; Lu, C.; Harder, E. D.; Abel, R.; Wang, L. Enhancing water sampling in free energy calculations with grand canonical Monte Carlo. *J. Chem. Theory Comput.* **2020**, *16*, 6061–6076.

[48] Steinbrecher, T.; Zhu, C.; Wang, L.; Abel, R.; Negron, C.; Pearlman, D.; Feyfant, E.; Duan, J.; Sherman, W. Predicting the effect of amino acid single-point mutations on protein stability–large-scale validation of MD-based relative free energy calculations. *J. Mol. Biol.* **2017**, *429*, 948–963.

[49] Thurlkill, R. L.; Grimsley, G. R.; Scholtz, J. M.; Pace, C. N. pK values of the ionizable groups of proteins. *Protein Sci.* **2006**, *15*, 1214–1218.

[50] Mobley, D. L.; Chodera, J. D.; Dill, K. A. Confine-and-release method: obtaining correct binding free energies in the presence of protein conformational change. *J. Chem. Theory Comput.* **2007**, *3*, 1231–1235.

Supporting Information

Table S1: Comparison between experimental binding free energies based on IC_{50} or K_i for JNK1 [37]

Ligand	IC_{50} (nM)	Ki	[a]RTln(IC_{50}) (kcal/mol)	[a]RTlnK_i (kcal/mol)	Difference (kcal/mol)
6a	750	190	-8.41	-9.23	-0.82
6o	45	2	-10.08	-11.94	-1.86
6s	38	3	-10.18	-11.7	-1.51
18b	14	1	-10.78	-12.35	-1.57
24	1900	550	-7.85	-8.59	-0.74
35	36	3	-10.22	-11.7	-1.48

[a]R is the perfect gas constant and T = 300 K.

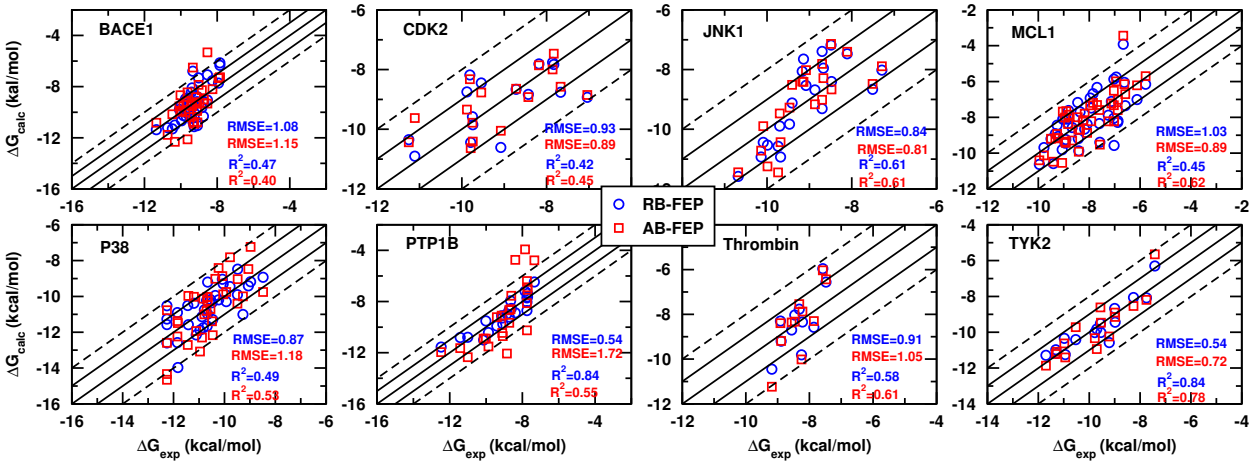

Figure S1: Comparison of calculated binding free energies with experiment for repeat 2 of AB-FEP. Blue circles and red squares are for RB-FEP and AB-FEP, respectively. The raw AB-FEP binding free energies were shifted to take into account the protein reorganization effect as detailed in the main text. Solid diagonal lines indicate the region within 1 kcal/mol of experimental values, while dashed diagonal lines indicate the region within 2 kcal/mol of experimental values. RMSE and R^2 are labeled.

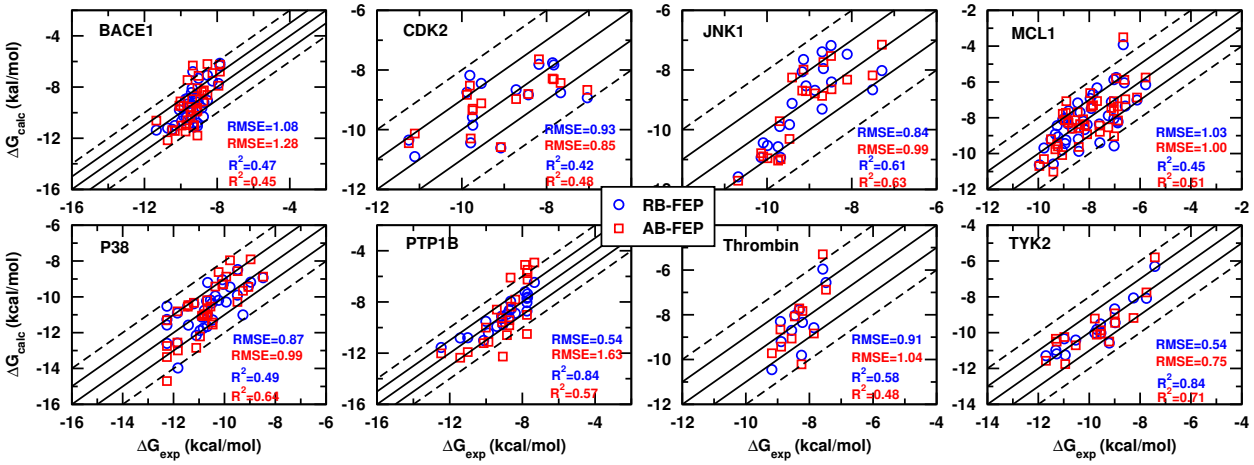

Figure S2: Comparison of calculated binding free energies with experiment for repeat 3 of AB-FEP. Blue circles and red squares are for RB-FEP and AB-FEP, respectively. The raw AB-FEP binding free energies were shifted to take into account the protein reorganization effect as detailed in the main text. Solid diagonal lines indicate the region within 1 kcal/mol of experimental values, while dashed diagonal lines indicate the region within 2 kcal/mol of experimental values. RMSE and R^2 are labeled.