JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

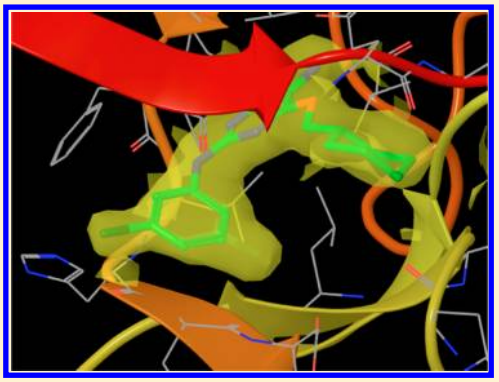
Modeling Local Structural Rearrangements Using FEP/REST: Application to Relative Binding Affinity Predictions of CDK2 Inhibitors

Lingle Wang,* Yuqing Deng, Jennifer L. Knight, Yujie Wu, Byungchan Kim, Woody Sherman, John C. Shelley, Teng Lin, and Robert Abel*

Schrodinger, 120 West 45st Street, New York, New York 10036, United States

Supporting Information

ABSTRACT: Accurate and reliable calculation of protein−ligand binding affinities remains a hotbed of computer-aided drug design research. Despite the potentially large impact FEP (free energy perturbation) may have in drug design projects, practical applications of FEP in industrial contexts have been limited. In this work, we use a recently developed method, FEP/REST (free energy perturbation/replica exchange with solute tempering), to calculate the relative binding affinities for a set of congeneric ligands binding to the CDK2 receptor. We compare the FEP/REST results with traditional FEP/MD (molecular dynamics) results and MM/GBSA (molecular mechanics/Generalized Born Surface Area model) results and examine why FEP/REST performed notably better than these other methods, as well as why certain ligand mutations lead to large increases of the binding affinity while others do not. We also introduce a mathematical framework for assessing the consistency and reliability of the calculations using cycle closures in FEP mutation paths.



INTRODUCTION

Biological processes often depend on protein−ligand binding events, and thus accurate calculation of the associated energetics is a central goal of computational structure-based drug design.[1−3] Many different factors contribute to the protein−ligand binding free energy, including the direct interaction between the protein and the ligand, the desolvation of the protein and the ligand, the protein and ligand strain energies associated with adopting conformations associated with binding, and the change in the configurational entropies of the protein and the ligand upon binding.[1−4] A variety of different approaches have been developed to calculate protein−ligand binding free energies with different tradeoffs between speed and accuracy, including the fast end point methods such as empirical scoring functions in virtual screening and molecular mechanics/generalized Born surface area (MM/GBSA) or Poisson−Boltzmann (MM/PBSA) models, the WM/MM (WaterMap/Molecular Mechanics) method that more explicitly considers the contribution from the desolvation of protein than GB methods, and free energy perturbation (FEP) methods that should offer a complete energetic description of binding within the accuracy limits of the underlying force field and complete sampling.[1,2,5−11]

Among the various methods to calculate protein−ligand binding affinities, free energy perturbation calculations performed by way of explicitly solvated molecular dynamics simulations (FEP/MD) are expected to provide a thermodynamically complete description of the binding event and yield maximally accurate predictions within the limits of the force field used and complete sampling of the phase space. As such, there is a widespread belief that such calculations should have a large impact on drug design projects.[12,13] However, practical applications of FEP in industrial contexts have been limited and have not been especially successful to date. In addition to the large computational resources required for FEP calculations, adequate sampling of all of the relevant conformations and converging the free energy calculation to the desired accuracy, in our estimate 0.8−1.0 kcal/mol mean unsigned error, has been very difficult, especially on time scales needed to impact live projects.[5,13,14]

In this work, we use FEP/REST (Free Energy Perturbation/Replica Exchange with Solute Tempering),[14−17] a recently developed enhanced sampling technique for free energy perturbation, to calculate the relative binding affinities of a subset from a series of congeneric ligands binding to the cyclin dependent kinase CDK2-cyclin A receptor, from a recent medicinal chemistry study.[18] CDK2 is a member of the CDK family, performing various functions in the regulation of the cell proliferation and the RNA polymerase II transcription cycles. CDK2 has also been identified as an important drug target for tumor-selective therapeutic strategies.[19,20] Our FEP/REST calculated relative binding affinities are highly correlated with the experimental data, while those calculated with normal FEP/

Received: October 22, 2012
Published: January 16, 2013

MD have less agreement. The main reason for the better performance of FEP/REST is due to the more thorough sampling of the different conformations of the ligand in the active site of the protein. In particular, we identified that for two of the ligands, there are multiple binding modes and that sampling those is important for the correct prediction of the binding affinities. While FEP/REST can sample all these important conformations in a relatively short simulation time, the ligands were trapped in the initial conformations using FEP/MD. MM/GBSA calculations using the FEP/MD and FEP/REST trajectories for this series of ligands suggested that the additional binding modes of some of the ligands in the series made comparatively favorable entropic contributions to binding that may be difficult to capture with end-state methods, such as MM/GBSA.

For a free energy calculation technique to become a true engineering platform in everyday drug design projects, a reasonable estimate of the associated errors in the calculation and the reliability of the predictions is also very important.[21] In this paper, we build up cycle closures in the FEP mutation paths as has been done by several other groups[11,22,23] and develop an approach to evaluate the consistency and reliability of the predictions by determining how much the sum of the calculated free energy changes for each closed thermodynamic cycle deviates from the theoretical value of 0. In Appendix A, we construct a mathematical framework that yields optimal predictions of the free energy from the multiple free energy estimates obtained by way of the cycle closure calculations, as well as the expected error bounds associated with those predictions and a mechanism to flag mutations that have systematic errors.

In what follows, we first describe the enhanced sampling method and the simulation protocol in detail and then compare the FEP/REST results with FEP/MD and MM/GBSA results. In particular, we examine why FEP/REST performed better than FEP/MD and why certain mutations lead to large increases of the binding affinity while others do not. In the appendices, we detail the mathematical framework for assessing the consistency and reliability of the calculations using cycle closures in FEP mutation paths and how the calculated total relative free energy can be decomposed into meaningful components.

METHODS

FEP/REST Sampling Algorithm. Converging explicit solvent model FEP calculations to the accuracy needed to provide utility in an industrial setting has proven to be difficult, especially when the protein residues surrounding the binding pocket move significantly when different ligands bind or the two ligands adopt different binding modes or manifest one or more degenerate binding modes.[5,13,21] In such cases, sampling all relevant conformations is essential for accurately estimating the free energy difference, yet the transition times to interchange between the different conformations are often too large to be accessible on the time scale of most FEP/MD simulation lengths. Here, we use a recently developed method, FEP/REST introduced by Wang et al.,[14] which incorporates the enhanced sampling method, REST (Replica Exchange with Solute Tempering), into FEP calculations through an efficient λ-hopping protocol, to sample relevant local structural rearrangements in relative protein−ligand binding affinity calculations within easily accessible simulation times.

FEP involves calculating the free energy difference between two systems by alchemically transforming between the systems in a series of discrete steps represented by λ values, where λ varies from 0 for the initial state to 1 for the final state. FEP/REST modifies the potential energy for a localized region surrounding the binding pocket (which we call the "hot" region) which may include protein residues as well as the ligand. For the intermediate λ windows, the potential energy for the localized region is scaled by a factor less than 1. In this way, energy barriers are lowered, enabling the efficient sampling of the different conformations in these intermediate λ windows, and the different conformations are propagated to the end states through the Hamiltonian replica exchange method.[17,24,25] We call the region where the potential energy is scaled the "hot" region, since the local contributions to the potential energy have been scaled to smaller values such that the effective temperature of that region is higher, although this analogy does not carry over to the particles actually having higher kinetic energies in the simulation. The details about the FEP/REST algorithm are described in the original FEP/REST work by Wang et al.[14]

As discussed by Wang et al.,[14] an appropriate selection of the "hot" region and an effective temperature profile reflects a tradeoff between the precision of the free energy results and the efficiency of the enhanced sampling. On the one hand, as one makes the "hot" region larger and the effective temperature higher, a more diverse range of conformations will be sampled. However, on the other hand, as one increases the effective temperature and size of the "hot" region, one also makes the free energy differences between neighboring λ windows larger, and the precision of the free energy results will correspondingly be reduced. Therefore, the "hot" region should be made as small and as cool as possible, while still allowing the complex to sample the relevant structural rearrangements within the allowed sampling times. Thus, generating a transferable set of parameters for FEP/REST, including the size of the "hot" region, the effective temperature profile, and the alchemical λ schedules, presents a technical challenge for the application of large-scale FEP/REST calculations.

In this paper, a transferable set of FEP/REST parameters is developed. The "hot" region for REST enhanced sampling is defined here to include the ligand functional group mutated in the SAR studies. If the ligand functional group is attached to an aromatic ring, the aromatic ring is also included in the "hot" region, allowing the ring to flip through REST enhanced sampling. The effective temperature profile for the REST "hot" region is calculated according to eq 9 of ref 26, setting the expected acceptance ratio of replica exchange to be 0.3, an optimal value reflecting the tradeoffs between the number of replicas needed and exchange efficiency. A total of 12 λ windows are used for both the normal FEP and FEP/REST calculations. The electrostatic interactions that are unique to the initial state were turned off before the Lennard-Jones (LJ) interactions, and the LJ interactions that are unique for the final state were turned on before the electrostatic interactions. The bonded interactions are smoothly turned from the initial state to the final state. The core of the LJ interactions was softened to avoid the singularities and instabilities in the simulations.[27] In our current simulations, the benzene ring on which substituents were varied in the SAR studies[18] are included in the "hot" region, and the highest effective temperature with a total of 12 λ windows achieved is roughly 900 K.

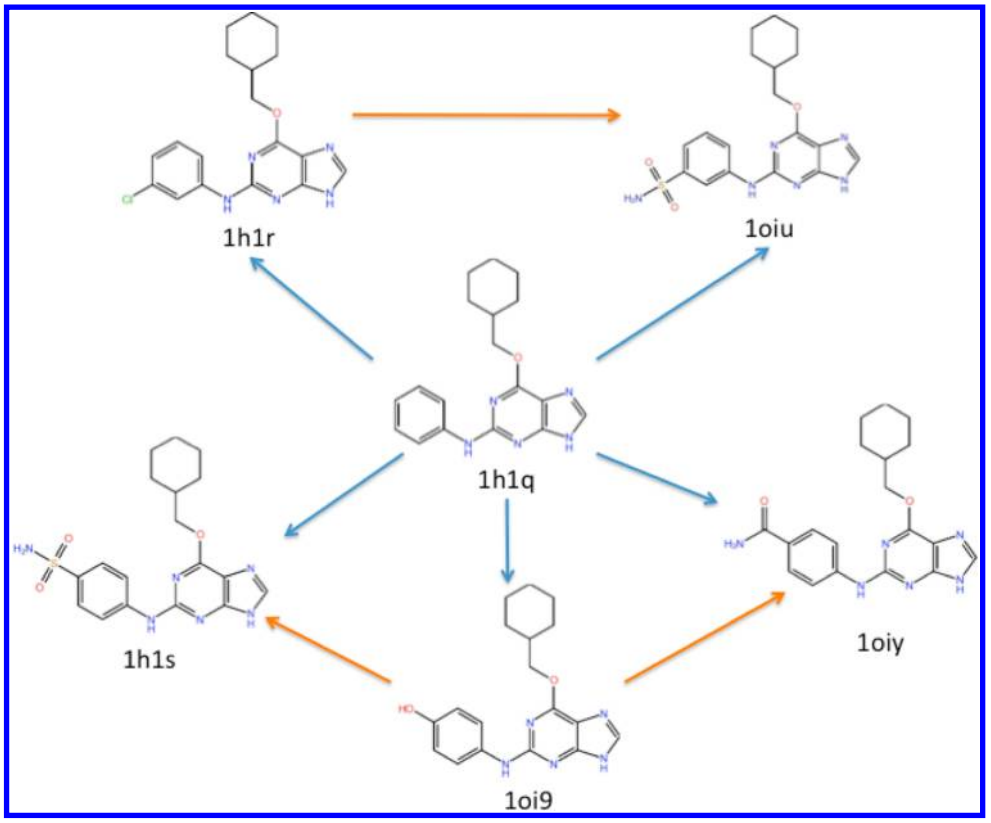


Figure 1. The structures of the ligands and the mutation paths used for the FEP/REST calculations. The three red arrows represent the mutations which complete three distinct cycles for assessing the consistency and reliability of the calculations. Ligand 1h1q was used for the five blue mutations since the other compounds can all be generated by simply adding different substituents on the benzene ring of ligand 1h1q.

Details of the Simulation Protocol. The ligand functional group mutation module implemented in Desmond[28,29] was used to set up the calculations.[30] The OPLS2005 force field[31,32] was used for the proteins and the ligands along with the SPC[33] water model. The starting structures for the simulations were taken from the PDB structures with IDs 1h1q, 1h1r, 1h1s, 1oiu, 1oi9, and 1oiy.[18] Whenever multiple conformations were present in the PDB, we selected the first conformation for use in the simulations. The protein was prepared using the Protein Preparation Wizard[34,35] during which the protonation states were assigned assuming a pH of 7.0.[36] Cyclin A was retained with CDK2 and treated as part of the receptor for all calculations. Na^+ or Cl^- ions were added to maintain electric neutrality. The systems were relaxed and equilibrated using the default Desmond relaxation protocol, which involves a series of minimizations and short molecular dynamics simulations. A total of 12 λ windows were used for both the normal FEP and FEP/REST calculations. The production stage lasted 2 ns for the complex simulations and 5 ns for the solvent simulations. Replica exchanges between neighboring λ windows were attempted every 1.2 ps. The production stages were sampled using both FEP/MD and FEP/REST. The Bennett acceptance ratio method (BAR)[37] was used to calculate the free energy. Errors were estimated for each free energy calculation using both bootstrapping[21,38] and the BAR analytical error prediction,[21,38,39] and the larger of the two errors was reported.

Assessing the Consistency and Reliability of the Predictions Using Cycle Closure. One can rank order the relative binding affinities for the whole set of ligands in relative protein−ligand binding affinity FEP calculations, most simply, by performing all FEP simulations from a common reference ligand. In addition, if one obtains a fully converged result for each calculation, the final predictions are, theoretically, independent of the paths chosen to do the mutations. For example, for three ligands A, B, and C, there are two strategies to obtain the relative binding affinity between ligand A and ligand C: (1) sample the mutation path directly from A to C or (2) sample the mutation in two steps, from A to B and B to C, and then sum the two obtained free energies. Theoretically, the final converged free energy estimate should be the same from the above two methods. In practice, however, due to the errors in each calculation—including unbiased errors due to the random fluctuations in the sampling of the phase space which may be characterized by the Bennett error bound, and biased error due to a *systematically* incomplete sampling of the phase space which would not be characterized by the Bennett error bound—both the variances and the average values from the above two mutation paths are usually somewhat different.[40] Given these limitations, how might one assess the reliability of the results obtained by either path?

To gain additional information about the consistency and the reliability of the predictions, we include cycle closures for all the mutation paths. In the example given above, if the results from the two mutation paths agree to within a specified threshold, it indicates that the predictions are consistent, and likely reliable within the limits of the employed force field. However, if the results from the two mutation paths differ by more than some threshold value, it may indicate that the results may not be converged and the predictions may not be reliable. The threshold value depends on the number of members in each cycle and the error we expect for each mutation in the cycle. We propose a method to rigorously assess the reliability of the

Table 1. The FEP/MD and FEP/REST Predicted Relative Binding Free Energies for the Five Ligands Compared to the Reference Ligand 1h1q[a]

ligand	ΔΔG(EXP)	ΔΔG(FEP/MD)	Δ(|FEP/MD-EXP|)	ΔΔG(FEP/REST)	Δ(|FEP/REST-EXP|)
1h1q→1h1r	0.51	−1.29 ± 0.09	**1.80**	0.14 ± 0.11	0.37
1h1q→1h1s	−3.07	−1.64 ± 0.27	**1.43**	−2.72 ± 0.28	0.35
1h1q→1oi9	−1.56	−1.83 ± 0.14	0.27	−1.70 ± 0.16	0.14
1h1q→1oiu	−0.91	−1.68 ± 0.25	0.77	−1.71 ± 0.26	0.80
1h1q→1oiy	−1.61	−1.69 ± 0.22	0.08	−1.92 ± 0.23	0.31

[a]The errors for the calculated free energies using BAR analytical error estimation are also included in the table. The FEP/MD predicted results have large deviations from the experimental data for ligands 1h1r and 1h1s (**bold** in the table), while the FEP/REST predicted results all agree well with the experimental data. Free energies are reported in units of kcal/mol.

predictions and how to adaptively assign an appropriate threshold value for each cycle in Appendix A. In addition, once we include cycle closures in the mutation path, there are multiple independent ways to calculate the free energy difference between two states. In this appendix, we also build up a model and derive a formula to obtain optimal predictions from the multiple free energy estimates obtained along the cycle. The hysteresis of each cycle closure itself also directly provides information about the convergence error associated with the free energy calculations. Last, we derive a formula in Appendix A to directly estimate the error associated with the free energy calculation using the hysteresis of each cycle closure.

MM/GBSA Calculations. *Single-Conformation MM/GBSA.* MM/GBSA calculations were performed for each ligand using the 1h1q receptor, the OPLS2005 force field,[31,32] and the VSGB2.0 implicit solvent model.[41] The initial ligand conformations were obtained by superimposing the binding pockets of each of the crystallographic structures. The side chains of two binding pocket residues (K88 and K89) along with 10° rotations and 0.1 Å translations of each ligand around its initial conformation were exhaustively sampled within a 1 Å hypersphere using the Protein R-Group Ligand (PGL) sampling method.[42] The energies of the receptor and ligand were then evaluated individually in solvent from the optimized bound conformation. The overall binding affinity for a given ligand was estimated by the difference in the energy calculated for the complex and the receptor and ligand modeled in the solvent environment.

Thermal MM/GBSA. MM/GBSA calculations were also performed by evaluating the energy of each of the 168 saved snapshots (with the explicit water removed) for each of the FEP/REST trajectories. The binding affinity for a given ligand was estimated as the difference between the average energy of the complex and that of the free receptor and free ligand. The average energy of the free receptor was estimated by analyzing the protein component of the snapshots from the FEP/REST trajectories of the complex. The free ligand in the MM/GBSA calculations was modeled either as the ligand component in the FEP/REST snapshot or in the FEP solvent simulation. These calculations were performed with and without minimization of the snapshots and the snapshot components.

RESULTS AND DISCUSSIONS

Performance on Ligands with Crystal Structures. FEP/MD and FEP/REST simulations were carried out to predict the relative binding affinities for the six ligands displayed in Figure 1 binding to the CDK2-cyclin A receptor. These are the only six ligands with crystal structures available for each holo complex, with PDB IDs 1h1q, 1h1r, 1h1s, 1oi9, 1oiu, and 1oiy,[18] and experimental binding affinity data for these ligands came from the same publication using the same method. Thus, the calculations reported here are essentially purely a scoring exercise, since the binding modes of ligands are known from the crystallography. In these crystal structures, the CDK2 receptor conformation is essentially the same regardless of which ligand is bound. Hence, we did not expect the sampling of protein conformational changes to significantly limit simulation convergence. In the calculations, we chose to use the protein conformation and ligand from 1h1q as the starting structure for all five perturbations (drawn with blue arrows in Figure 1), since all other ligands in the set are generated by adding different substituents on the benzene ring of ligand 1h1q.

Comparison between FEP/MD and FEP/REST Results. The FEP/MD and the FEP/REST results for the mutations among the ligands with crystal structures are given in Table 1. The correlation of the predictions with experimental binding affinity data is displayed in Figure 2. Despite the ∼3.5 kcal/mol range of experimental binding affinities, FEP/MD predicts that the five compounds (excluding the reference compound 1h1q) have almost the same binding affinities, resulting in an R^2 value

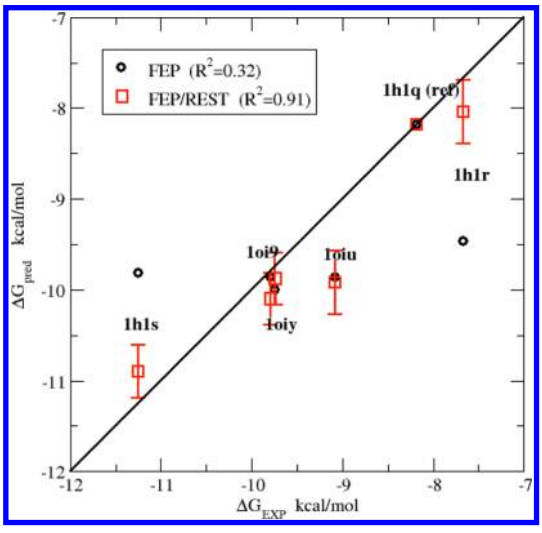


Figure 2. The FEP/MD and FEP/REST predicted results versus the experimental binding affinity data for the set of CDK2 inhibitors. The FEP/MD predicted binding affinities do not vary significantly, making it impossible to reliably rank ligands 1h1r, 1h1s, 1oiu, 1oi9, and 1oiy using these values. By contrast, the FEP/REST predicted results have a very high correlation with experimental data (R^2 = 0.91), and the deviation is less than 0.8 kcal/mol for all mutation pairs. The error bars for the FEP/REST predictions reported in this figure are calculated on the basis of the method detailed in Appendix A using cycle closures.

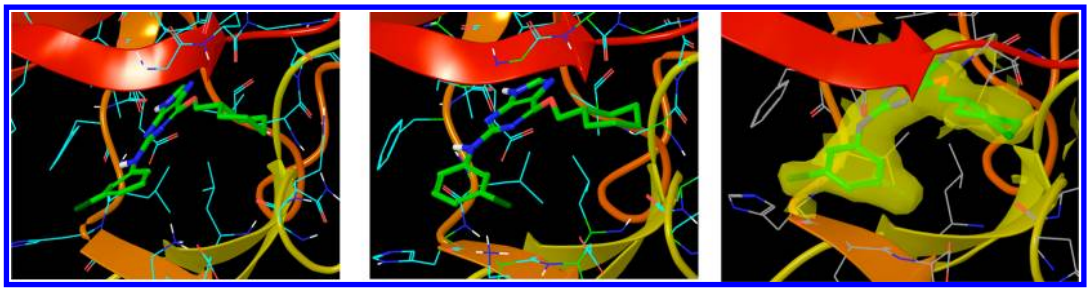


Figure 3. The starting conformation corresponding to the crystal structure for the FEP/REST simulation for ligand 1h1r (left), the alternative binding mode, with the chloro- substituted phenyl ring flipped, observed in the FEP/REST trajectory (middle), and the electron density of ligand 1h1r in the X-ray crystallography. (right) The proteins are represented by colored ribbons for the backbone and stick representation for the side-chains, while the ligands are depicted in tube representation, and the electron density is displayed as a solid yellow surface. The sigma cutoff used for density visualization is 1.0. For ligand 1h1r, the electron density has two peaks for the chloro-substituted benzene ring, indicating that the ligand has two binding modes, consistent with the FEP/REST simulations.

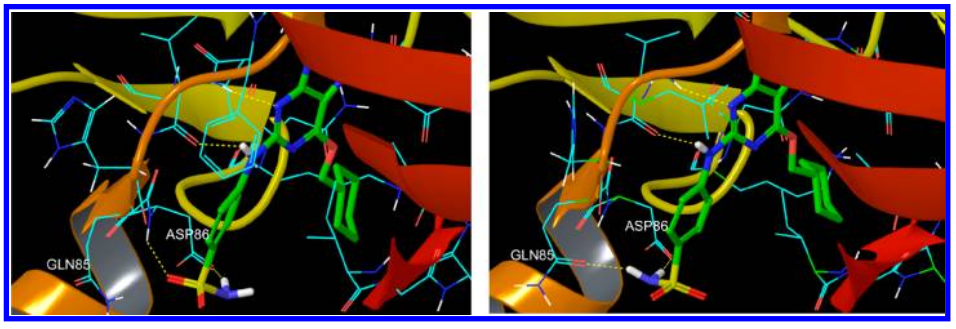


Figure 4. The starting conformation corresponding to the crystal structure for FEP/REST simulation for ligand 1h1s (left) and the alternative binding mode observed in the FEP/REST trajectory (right). The proteins are represented by colored ribbons for the backbone and stick representation for the side-chains, while the ligands are depicted in tube representation. Compared to the starting conformation, the sulfonamide group rotates about 120° around the C−S bond, and the residues surrounding the binding pocket, especially Gln85 and Asp86, rearrange accordingly. Two hydrogen bonds are formed between the sulfonamide group and the protein residues.

of 0.32. By comparison, the FEP/REST results are highly correlated with the experimental values (R^2 = 0.91), and the deviations of the FEP/REST results from the experimental values are less than 0.8 kcal/mol for all of the mutation pairs.

From comparisons of the FEP/MD and FEP/REST results, it appears that the predictions for compounds 1oi9, 1oiu, and 1ioy from the two methods are essentially the same and that the main reason for the improved correlation with the experimental results comes from the better binding affinity estimates for ligands 1h1s and 1h1r using FEP/REST. Below, we will discuss why FEP/REST significantly outperforms FEP/MD for these two cases.

A review of the FEP/MD and FEP/REST trajectories for ligand 1h1r showed that the 2-chloro substituted benzene ring flips several times in the FEP/REST trajectory, while that ring remains trapped in the initial conformation in the FEP/MD trajectory. The initial conformation, corresponding to one of the conformations in the crystal structure, and the alternative flipped conformation, observed in the FEP/REST calculation, are displayed in Figure 3. Interestingly, for ligand 1h1r, the electron density in the deposited X-ray crystal structure also indicated the presence of an alternative ligand binding mode (Figure 3 right panel) that corresponds to the alternate one found in the FEP/REST calculations.

For 1h1s, in addition to the binding mode modeled in the crystal structure, another conformation is sampled in the FEP/REST trajectory, but not in the FEP/MD simulations. Figure 4 displays the binding mode found in the crystal structure and the alternative binding mode sampled in the FEP/REST calculation. In the crystal structure, the sulfonamide group makes two hydrogen bonds (HB) with Asp86: one with the backbone and another with the side chain. In the alternative binding mode sampled in the FEP/REST simulation, the sulfonamide group rotates about 120° around the CS bond relative to the crystal structure. The protein residues surrounding the binding pocket, particularly Gln85 and Asp86, rearrange accordingly, and two hydrogen bonds are formed between the sulfonamide group and the protein residues: one with the Asp86 side chain and the other with the Gln85 side chain. The R factors of these few residues surrounding the binding pocket and the ligand in the X-ray crystallography indicate that these groups are quite mobile.[18] Therefore, it is quite possible that the ligand can switch between the two binding modes in solution. In the FEP/REST simulation trajectory, the occupancy ratio between the binding mode observed crystallographically and the alternative binding mode is roughly 60:40, in favor of the crystallographically determined mode. However, sampling this alternative binding mode appears to be very important for the correct prediction of the binding affinity and is responsible for the greater efficacy of FEP/REST in accurately computing the relative binding free energy of this species compared with FEP/MD.

Comparison between Ligands 1h1s and 1oiu. Ligands 1h1s and 1oiu can be generated by adding a sulfonamide group to the para- and meta- positions of the benzene ring of 1h1q, respectively. Interestingly, for the 1h1s case, adding the sulfonamide group increases the binding affinity by about 3 kcal/mol compared to ligand 1h1q, while for 1oiu, adding the

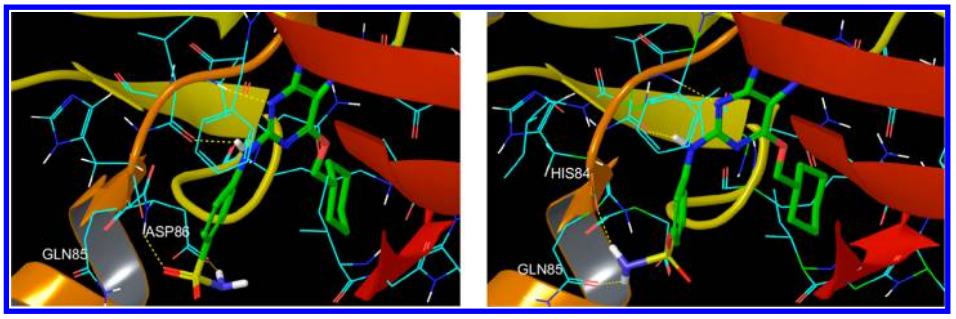


Figure 5. The crystal structures of ligand 1h1s (left) and 1oiu (right). The proteins are displayed in colored ribbon representation for the backbone and stick representation for the side chains, while the ligands are displayed in tube representation. In both cases, the sulfonamide group makes two hydrogen bonds with the protein residues.

Table 2. Decomposition of the Relative Binding Free Energies for Ligands 1h1s and 1oiu into Components[a]

	ΔΔ*G* breakdown by components					
ligands	ligand-self	protein−ligand interaction energy	ligand-water interaction energy	sum of each terms	total ΔΔ*G*	Δ(|total-Σcomponents|)
1h1s	0.24	−13.24	10.25	**−2.75**	**−2.72**	0.03
1oiu	1.49	−13.76	10.51	**−1.76**	**−1.71**	0.05
difference	***−1.25***	0.52	−0.26	**−0.99**	**1.01**	0.02

[a]The largest binding free energy difference between the two ligands comes from the ligand itself (***bold italic*** in the table), indicating the ligand strain or ligand conformational entropy difference between the two ligands is responsible for the tighter binding of ligand 1h1s. For comparison, the sum of the free energy differences due to each component and the total free energy difference from BAR estimate (**bold** in the table) are also listed in the table. The deviations between the two methods are very small, indicating that the decomposition of the free energy is thus expected to be very reliable. Free energies are reported in units of kcal/mol.

sulfonamide group increases the binding affinity by only about 1 kcal/mol. The 1h1s and 1oiu binding sites are shown in Figure 5. In both cases, the sulfonamide group makes two hydrogen bonds with the protein residues. In the 1h1s case, two hydrogen bonds are formed between the sulfonamide group and the Asp86 (one with its backbone NH and another with its side chain carboxylate group), while in the case of 1oiu, the sulfonamide group forms a hydrogen bond with the backbone of His84 and another hydrogen bond with the side chain of Gln85. So, hydrogen bonding interactions are expected to make similarly favorable enthalpic contributions to the binding of either species. However, for the 1h1s case, the sulfonamide group has multiple modes in which to make hydrogen bonds with protein residues (discussed in the previous section), while in the 1oiu case, there is only one conformation that allows the sulfonamide group to simultaneously make two hydrogen bonds with the protein residues. Therefore, compared to ligand 1h1s, ligand 1oiu would be expected to lose more entropy when binding to CDK2 since it is restricted to only a single predominant binding mode while 1h1s has two. Since the favorable enthalpic contribution of the sulfonamide group in 1oiu is offset by an unfavorable entropic contribution, the net contribution to the binding is modest. By comparison, the 1h1s sulfonamide group forms favorable enthalpic contributions due to hydrogen bonding but also induces less entropic penalties since it has multiple modes by which to make hydrogen bonds with the surrounding protein residues. So adding the sulfonamide group at the para- position of benzene ring in ligand 1h1s increases the binding affinity by a large amount, about 3 kcal/mol in this case.

To quantify the effect of ligand conformational entropy to the binding, we decompose the free energy into components. For a given λ schedule, the energy difference between neighboring Hamiltonians for each frame sampled in the trajectories may be decomposed into contributions from different sources and components, and assuming that the component contributions are independent, the Bennett acceptance ratio method can be used to calculate the component-based free energy difference between neighboring λ windows. Summing over the free energy difference between all neighboring λ windows gives the net contribution to the free energy for each component. Appendix B provides the details about how to decompose the free energy into each physically interpretable component in FEP. Since the component free energies differences are not state functions, the contribution from each component depends on the λ schedule used to do the mutation. Since we use the same λ schedule for all the simulations, the relative contribution from each free energy component for different ligands provides information regarding the underlying driving forces determining relative binding potencies of the ligands. Furthermore, the reliability of such an approach can be validated by comparing the total free energy with the sum of the component free energies (they may differ due to nonadditive ensemble effects).

Table 2 lists the contribution from each component to the relative binding free energies of ligand 1h1s and ligand 1oiu. The sum of the component free energy differences agrees well with the total free energy difference, suggesting that the decomposition of the free energy may have a consistent meaning for these two mutations. The largest difference between the two ligands comes from the ligand self-term. This term can be interpreted as the ligand conformational entropic and enthalpic contributions to the free energy difference or more simply stated the difference in the ligand strain free energy. It is clear that the ligand conformational strain free energy favors the ligand 1h1s by about 1.2 kcal/mol over the ligand 1oiu, suggesting that the additional binding mode for ligand 1h1s is the main reason for the tighter binding of ligand 1h1s compared to ligand 1oiu. The ligand−protein interaction energy favors ligand 1oiu, but this effect is partially

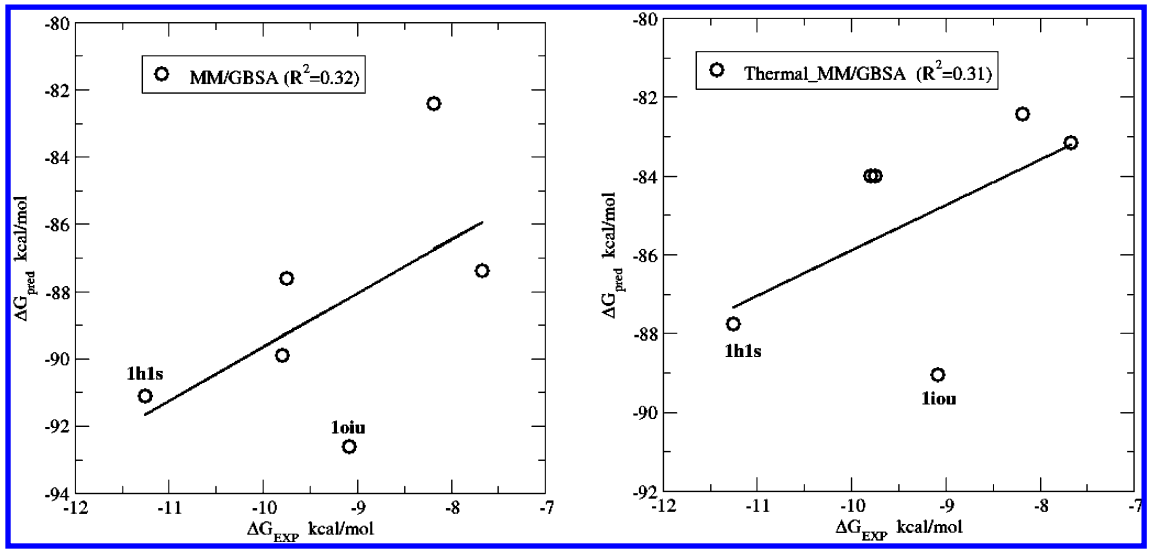


Figure 6. Left: The MM/GBSA predicted binding affinities versus the experimental values. The MM/GBSA method cannot distinguish the relative binding affinities between ligands 1h1s and 1oiu. Right: The thermal MM/GBSA predicted results versus experimental data. Averaging the MM/GBSA scores for all the frames saved from FEP/REST trajectory does not correctly predict the relative binding affinities of ligands 1oiu and 1h1s, which verifies that the alternative binding mode for ligand 1h1s is the main reason for the tighter binding of ligand 1h1s.

offset by the ligand−water interactions. Overall, the FEP/REST results predict that the binding of the ligand 1h1s to CDK2 is about 1 kcal/mol more favorable than ligand 1oiu.

Our hypothesis that a relatively favorable entropic contribution to the binding for ligand 1h1s compared to ligand 1oiu was further supported through single snapshot MM/GBSA calculations. The MM/GBSA predicted binding affinities versus the experimental data are given in Figure 6. While the MM/GBSA predictions have some correlation with the experimental data for ligands 1h1q, 1h1r, 1h1s, 1oiy, and 1oi9, it severely overestimates the binding affinity of ligand 1oiu. In fact, the MM/GBSA calculations predict ligand 1oiu to be a slightly stronger binder than ligand 1h1s. This result agrees with the decomposition of the free energy listed in Table 2 since the protein−ligand interaction energy slightly favors ligand 1oiu and the MM/GBSA model employed here does not include the conformational entropy of the ligand, which is here suspected to be the primary driving force behind the tighter binding of ligand 1h1s. Thus, the MM/GBSA calculations cannot correctly rank order the relative binding affinity between ligands 1h1s and 1oiu.

It has been shown that single snapshot MM/GBSA calculations have difficulty reproducing experimental free energies when multiple conformations exist in the bound or solvent environments. For example, Mobley et al.[6] modeled small molecule hydration free energies using implicit solvent models and found relatively large differences in hydration free energy estimates depending on the ligand conformation used, with errors up to 2 kcal/mol. In protein−ligand simulations, the contributions arising from conformational changes between the bound and the solvated state may be significant.[1,5,43,44] To assess the conformational enthalpic contributions, we also performed MM/GBSA calculations for each frame saved in the FEP/REST trajectories to obtain a thermal MM/GBSA score. However, the thermal MM/GBSA predictions still indicate that ligand 1oiu is a tighter binder than ligand 1h1s.

We believe this lends significant support to our hypothesis that the relatively tighter binding of 1h1s is due to a primarily entropic effect. If the binding affinity difference between ligand 1h1s and ligand 1oiu is due to the ligand strain energy not the entropy, then the thermal MM/GBSA method should be able to correctly rank order the relative binding affinities between these two ligands. Therefore, the stronger binding of ligand 1h1s compared to ligand 1oiu must be due to the favorable ligand conformational entropy, further validating that the additional binding mode for ligand 1h1s is important for the stronger binding. The trends remain the same whether the frames are minimized or not and whether the energy contributions for the free ligand are evaluated from the FEP trajectories modeling the bound or the solvent environment.

Consistency and Reliability of the Predictions. To assess the consistency and reliability of the FEP/REST predictions, we performed three more simulations forming three cycle closures in the mutation path (Figure 1). The FEP/REST results for the three mutations forming these cycles are given in the first three rows of Table 3. The final predictions for the binding affinities of these ligands using the estimator derived in Appendix A are given in Table 4. The cycles are closed self-consistently with discrepancies less than 0.8 kcal/mol from different paths for all three cycles, indicating that the free energy is well converged and the predictions are highly reliable within the limits of the underlying force field.

However, even though the cycles are well closed and the free energies calculated from different paths agree to within a rather tight threshold value, they are not exactly the same. Thus, for the same pair of ligands, we have multiple different estimates for the free energy differences depending upon how one traverses the paths between the ligand. One is then left with the question of how to best incorporate all the information that is generated in the calculations to obtain the statistically best estimates for the free energies? We use the model and formulas derived in Appendix A to get the optimal estimates of the free energies and the errors associated with these predictions. Here, all the cycles are well closed, and the final predictions with cycle closures agree well with the above predictions without cycle closures, further validating the high reliability of these calculations. The deviation between the final predictions and the experimental results are all less than 0.7 kcal/mol, and they

Table 3. The FEP/REST Predicted Relative Binding Free Energies for the Additional 10 Compounds without Crystal Structures[a]

ligand	ΔΔG(EXP)	ΔΔG(FEP/REST)	Δ(|FEP/REST-EXP|)
1h1r→1oiu	−1.41	−1.25 ± 0.25	0.16
1oi9→1h1s	−1.51	−0.81 ± 0.30	0.70
1oi9→1oiy	−0.04	0.49 ± 0.24	0.53
1h1q→17	1.14	0.19 ± 0.11	0.95
1h1q→20	−0.54	−0.52 ± 0.18	0.02
1h1q→21	0.35	0.64 ± 0.15	0.29
1h1q→22	0.42	0.44 ± 0.15	0.02
1h1q→26	−0.25	−1.39 ± 0.14	1.14
1h1q→31	−1.36	−0.57 ± 0.19	0.79
1h1s→28	0.14	0.14 ± 0.20	0.00
1h1s→30	1.44	1.53 ± 0.15	0.09
1h1s→32	1.50	0.78 ± 0.17	0.72
28→29	1.23	1.05 ± 0.18	0.18
1h1r→17	0.63	0.13 ± 0.11	0.50
21→22	0.07	0.62 ± 0.24	0.55
1h1s→29	1.37	2.19 ± 0.19	0.82
30→29	−0.07	−0.27 ± 0.19	0.20
1oiu→26	0.65	1.99 ± 0.19	1.34

[a]The results for the three mutations forming three cycle closures for the six ligands with crystal structures are also include in the table. The errors for the calculated free energies using BAR analytical error estimation are also reported in the table. Free energies are reported in units of kcal/mol.

Table 4. The Final Free Energy Predictions for the Whole Set of Ligands Using the Formula Derived in Appendix A[a]

	ΔG(EXP)	ΔG(FEP/REST)	Δ(|FEP/REST-EXP|)	Err(cycle_closure)
1h1q	−8.18	−8.18		
1h1r	−7.67	−8.25	0.58	0.35
1h1s	−11.25	−10.91	0.34	0.29
1oi9	−9.74	−10.11	0.37	0.29
1oiu	−9.08	−9.71	0.63	0.35
1oiy	−9.79	−9.86	0.07	0.29
17	−7.04	−8.06	1.02	0.26
20	−8.72	−8.70	0.02	
21	−7.83	−7.81	0.02	0.47
22	−7.76	−7.47	0.29	0.47
26	−8.43	−8.65	0.22	0.69
28	−11.11	−10.51	0.60	0.58
29	−9.88	−9.20	0.68	0.58
30	−9.81	−9.16	0.65	0.54
31	−9.54	−8.75	0.79	
32	−9.75	−10.13	0.38	
R^2		0.82		

[a]The error estimates using the formula derived in Appendix A for the final free energy prediction using cycle closure are also included in the table. Ligand 1h1q is used as the reference, so the deviation for ligand 1h1q is 0 by design. Note, a cycle closure error bar is not reported for ligand 1h1q since it is the reference ligand used for this work, and its binding free energy is set identically equal to its experimental value for plotting purposes. Cycle closure error bars are also not reported for ligands 20, 31 and 32, since those species were technically challenging to connect to other ligands other than the reference ligand, as required to form closed cycles for the error. Free energies are reported in units of kcal/mol.

have very high correlation with the experimental data (R^2 = 0.92).

Performance on Ligands without Crystal Structures. To further validate our model, additional FEP/REST calculations were performed for another 10 ligands binding to the same receptor without a crystal structure available for the holo complex. These 10 ligands were randomly selected from the same publication by Hardcastle.[18] The binding affinities were measured using the same method as the above-mentioned ligands with crystal structures, and they conserve the same core. The structures of the ligands and their corresponding compound numbers as in the original publication[18] are given in the Supporting Information. The FEP/REST predicted relative binding affinities for these ligands are reported in Table 3 along with the experimental data. Cycle closures are constructed in the mutation paths for these ligands, and the hysteresis of the cycle closures are all less than 1.0 kcal/mol, indicating that the calculations are highly converged. The final predicted relative binding affinities using the formula derived in Appendix A are reported in Table 4. For ligands with cycle closures constructed in the mutation paths, the errors associated with the free energy predictions using the formula derived in Appendix A are also reported in Table 4. The correlation between the predicted relative binding affinities and the experimental data for the whole set of ligands (six ligands with crystal structure and the 10 ligands without crystal structure) is displayed in Figure 7. It is clear that our particular

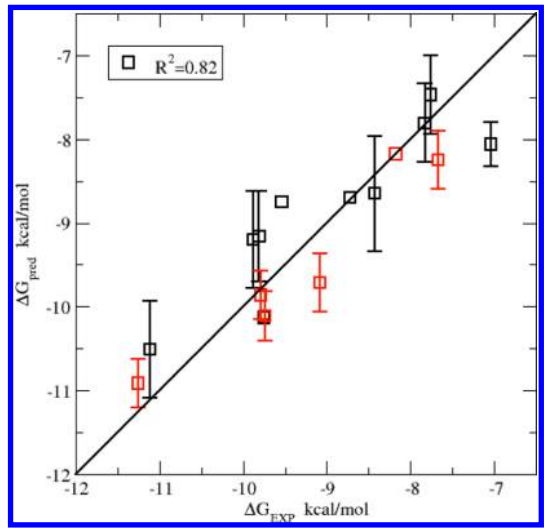


Figure 7. The FEP/REST predicted results versus the experimental binding affinity data for the whole set of CDK2 inhibitors including both the six ligands with crystal structures and the 10 ligands without crystal structures. The six ligands with crystal structures (displayed in Figure 1) are colored red in the figure, and the 10 ligands without crystal structures are colored black. The error bars for the FEP/REST predictions reported in this figure are calculated based on the method detailed in Appendix A using cycle closures.

FEP/REST protocol performs well for these ligands without crystal structures, further demonstrating the transferability of the approach. The average deviation between the predicted binding affinities compared to experimental data is 0.44 kcal/mol, and the R^2 value is 0.82.

CONCLUSIONS

In this paper, the recently developed FEP/REST method was used to calculate the relative binding affinities between a series of congeneric ligands binding to the CDK2-cyclin A receptor,

and comparisons were made between FEP/REST and FEP/MD predictions. A transferable set of parameters was introduced for FEP/REST calculations, including, the size of the "hot" region, the effective temperature profile, the number of λ windows, and the alchemical λ schedule, and they worked well for all these calculations. While the normal FEP/MD method performed poorly in rank-ordering the relative binding affinities for the six ligands with crystal structures, the FEP/REST calculated relative binding affinities were in excellent agreement with the experimental data. Application of the FEP/REST method on a larger data set that includes 10 ligands without crystal structures further verified its ability to correctly rank order the relative binding affinities of congeneric ligands.

For two of the ligands, 1h1r and 1h1s, multiple binding modes were identified from the FEP/REST trajectories. While the ligands were trapped in their initial conformation using FEP/MD, all these different conformations were sampled using FEP/REST within relatively short simulations (2 ns), resulting in better convergence of the free energy and improved agreement with experimental data. Comparison between ligand 1h1s and another ligand in this series (ligand 1oiu), both containing the sulfonamide group but located in different positions of the benzene ring, indicated that the multiple binding modes for ligand 1h1s are responsible for its tighter binding.

We have introduced an approximate method to decompose the relative free energy in FEP calculations into its different components. Decomposition of the relative free energy between ligand 1h1s and 1oiu suggest that the tighter binding of ligand 1h1s compared to ligand 1oiu comes from conformational entropic contributions to the ligand strain free energy. Results from MM/GBSA calculations also verified that the stronger binding of ligand 1h1s compared to ligand 1oiu appears to be an entropic effect, largely due to the multiple accessible conformations of ligand 1h1s when bound.

We have also derived a rigorous framework to assess the consistency and reliability of free energy calculations using information derived from cycle closure calculations along mutation paths. This framework provides clear criteria by which to judge the consistency and reliability of the predictions, as well as optimal estimates of the free energies with well-formed error bounds. Application of this model to the CDK2-cyclin A series indicates that our predictions using FEP/REST are highly reliable, which is also demonstrated by the excellent agreement of the predictions with the experimental data.

APPENDIX A

Consistency Validation and the Best Free Energy Estimator Using Cycle Closure in FEP Mutation Path

Let us consider a simple model, with a set of three ligands, L_1, L_2, and L_3. Suppose the experimentally measured relative binding free energy differences between the three ligands are $F_{21}^{exp} = F_2^{exp} - F_1^{exp}$, $F_{32}^{exp} = F_3^{exp} - F_2^{exp}$, and $F_{13}^{exp} = F_1^{exp} - F_3^{exp}$, where F_1^{exp}, F_2^{exp}, and F_3^{exp} are the experimentally measured binding free energies for the ligands L_1, L_2, and L_3, respectively. The free energy is a thermodynamic property; therefore $F_{13}^{exp} + F_{32}^{exp} + F_{21}^{exp} = 0$ (see Figure 8). Suppose three FEP simulations are performed, from L_1 to L_2, L_2 to L_3, and L_3 to L_1, and the calculated free energy differences for the three mutations are E_{21}, E_{32}, and E_{13}, respectively. If the simulations are perfectly converged and the force field is perfect, then ideally $E_{21} = F_{21}^{exp}$, $E_{32} = F_{32}^{exp}$, $E_{13} = F_{13}^{exp}$, and $E_{13} + E_{21} + E_{32} = 0$. In practice,

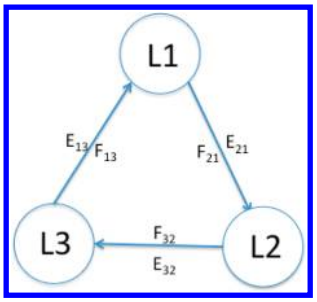


Figure 8. A simple model illustrating the problem of consistency and reliability checking in FEP calculations. For a set of three ligands L_1, L_2, and L_3, three FEP calculations were performed with estimated free energy differences E_{21}, E_{32}, and E_{13}, respectively. How do we get the best estimates of the relative free energies (F_{21}, F_{32}, and F_{13}) from the three FEP calculations, and how do we judge whether the calculations are reliable?

however, there are errors associated with the calculated free energies, and in general, $E_{21} \neq F_{21}^{exp}$, $E_{32} \neq F_{32}^{exp}$, $E_{13} \neq F_{13}^{exp}$, and $E_{13} + E_{32} + E_{21} = \Delta \neq 0$. We call Δ the hysteresis of the free energy calculation associated with this cycle closure. These errors include the unbiased statistical errors due to the random fluctuations and the biased errors due to the incomplete sampling of the phase space (the protein and/or the ligand are trapped in a local minimal in the conformation space) and the error in the force field. How do we assess the consistency and reliability of these calculations?

The errors in FEP calculations compared to experimental values can be separated into two categories: the systematic error coming from the difference of the force field used in simulation compared to the true potential energy surface of the system and the error coming from the unconverged calculation, either due to randomly or systematically incomplete sampling of the phase space, or from the free energy estimator itself. Let F_{21} denote the theoretical free energy difference between two thermodynamic states (mutation from ligand L_1 to ligand L_2) for the underlying force field from an infinitely long unbiased simulation and unbiased free energy estimator. If there is no systematic error in the force field, then $F_{21} = F_{21}^{exp}$. In practical FEP calculations, for example, for mutation from L_1 to L_2, the simulation is run with a finite amount of time and the sampling may have some bias; thus the calculated free energy E_{21} might deviate from its theoretical value F_{21}, and it depends on the initial configuration of the simulation. Suppose we repeat the same FEP calculation for an infinite number of times starting from different initial configurations and different random seeds for the velocities, then the calculated free energies have a distribution. Let us assume that the calculated free energies are Gaussian distributed with average F_{21} (no systematic bias) and standard deviation σ_{21}. Then the probability density that one single FEP calculation gives a value of E_{21} for this mutation is

$$\rho_{21}(E_{21}|F_{21}) = \frac{1}{\sqrt{2\pi}\,\sigma_{21}} \exp\left(-\frac{(E_{21} - F_{21})^2}{2\sigma_{21}^2}\right)$$

Similarly, for mutations L_2 to L_3, and L_3 to L_1, the probability density that the FEP calculations give values of E_{32} and E_{13} are respectively

$$\rho_{32}(E_{32}|F_{32}) = \frac{1}{\sqrt{2\pi}\sigma_{32}} \exp\left(-\frac{(E_{32} - F_{32})^2}{2\sigma_{32}^2}\right)$$

$$\rho_{13}(E_{13}|F_{13}) = \frac{1}{\sqrt{2\pi}\sigma_{13}} \exp\left(-\frac{(E_{13} - F_{13})^2}{2\sigma_{13}^2}\right)$$

For a given set of theoretical free energy differences F_{21}, F_{32}, and F_{13}, the overall likelihood, L, that the three FEP simulations give values of E_{21}, E_{32}, and E_{13} is

$$L(E_{21}, E_{32}, E_{13}|F_{21}, F_{32}, F_{13}) = \frac{1}{(2\pi)^{3/2}\sigma_{21}\sigma_{32}\sigma_{13}} \exp\left(-\sum_i \frac{(E_i - F_i)^2}{2\sigma_i^2}\right) \quad \text{where } i \in (21, 32, 13)$$

According to the maximum likelihood method, the most likely values of F_{21}, F_{32}, and F_{13} are the set of values that maximize the above likelihood. Taking the log of the above expression, it is clear that the set of values which maximizes the likelihood is the set of values that minimize the following function:

$$\sum_i \frac{(E_i - F_i)^2}{2\sigma_i^2} \quad \text{where } i \in (21, 32, 13)$$

with the constraint

$$F_{21} + F_{32} + F_{13} = 0$$

Using Lagrange multipliers, one may show that the set of values which maximizes the likelihood is

$$\hat{F}_{21} = E_{21} - \frac{\sigma_{21}^2}{\sigma_{21}^2 + \sigma_{32}^2 + \sigma_{13}^2}(E_{21} + E_{32} + E_{13}) = \frac{\sigma_{32}^2 + \sigma_{13}^2}{\sigma_{21}^2 + \sigma_{32}^2 + \sigma_{13}^2}E_{21} - \frac{\sigma_{21}^2}{\sigma_{21}^2 + \sigma_{32}^2 + \sigma_{13}^2}(E_{32} + E_{13})$$

$$\hat{F}_{32} = E_{32} - \frac{\sigma_{32}^2}{\sigma_{21}^2 + \sigma_{32}^2 + \sigma_{13}^2}(E_{21} + E_{32} + E_{13}) = \frac{\sigma_{21}^2 + \sigma_{13}^2}{\sigma_{21}^2 + \sigma_{32}^2 + \sigma_{13}^2}E_{32} - \frac{\sigma_{32}^2}{\sigma_{21}^2 + \sigma_{32}^2 + \sigma_{13}^2}(E_{21} + E_{13})$$

$$\hat{F}_{13} = E_{13} - \frac{\sigma_{13}^2}{\sigma_{21}^2 + \sigma_{32}^2 + \sigma_{13}^2}(E_{21} + E_{32} + E_{13}) = \frac{\sigma_{32}^2 + \sigma_{21}^2}{\sigma_{21}^2 + \sigma_{32}^2 + \sigma_{13}^2}E_{13} - \frac{\sigma_{13}^2}{\sigma_{21}^2 + \sigma_{32}^2 + \sigma_{13}^2}(E_{32} + E_{21})$$

It is also straightforward to prove that the above estimators have no systematic bias, and there will be no discrepancy between the free energy predictions from different paths. We can interpret the above estimators as the weighted average from the two paths. For example, the free energy difference between ligand L_1 and L_2, F_{21}, can be estimated from E_{21}, or from $-(E_{32} + E_{13})$, and the best estimator is a weighted average of the two predictions. The smaller the standard deviation for the mutation is, the larger the weight to the best estimator, and vice versa.

In addition, according to the above model, the hysteresis of the cycle closure, $E_{21} + E_{32} + E_{13} = \Delta$, is also Gaussian distributed with an average of 0 and standard deviation of $s = (\sigma_{21}^2 + \sigma_{32}^2 + \sigma_{13}^2)^{1/2}$. If the sum of the three calculated free energies deviates by more than $2s$ from 0, it is almost certain ($P = 0.95$) that the calculations are not converged and the results are not reliable; if the sum of the three calculated free energies deviates by more than s from 0, it is highly probable ($P = 0.68$) that the calculation is not converged and the results may not be reliable. Thus, the above model provides a simple way to determine whether the predictions are reliable.

In practical FEP calculations, due to the high computational expense, one usually only performs a given calculation for a single time, and thus one cannot estimate a standard deviation for each prediction. In this case, the rule of thumb assumption is that the standard deviation for each calculation is the same, for example, 0.8 kcal/mol for each mutation. Then, if the discrepancy of the free energy results from two paths are larger than 1.4 kcal/mol, it indicates that the calculation is not converged and probably not reliable. Under this assumption, the best free energy estimator has the following simple expression:

$$\hat{F}_{21} = E_{21} - \frac{1}{3}(E_{21} + E_{32} + E_{13}) = \frac{2}{3}E_{21} - \frac{1}{3}(E_{32} + E_{13})$$

$$\hat{F}_{32} = E_{32} - \frac{1}{3}(E_{21} + E_{32} + E_{13}) = \frac{2}{3}E_{32} - \frac{1}{3}(E_{21} + E_{13})$$

$$\hat{F}_{13} = E_{13} - \frac{1}{3}(E_{21} + E_{32} + E_{13}) = \frac{2}{3}E_{13} - \frac{1}{3}(E_{32} + E_{21})$$

In addition, under the assumption that the standard deviation for each calculation is the same, the hysteresis of the cycle closure itself provides an estimate of the error associated with each free energy prediction. As discussed above, the hysteresis Δ of the cycle closure is Gaussian distributed with an average of 0 and standard deviation of $s = \sqrt{3}\sigma$. So the probability that a given calculation generates a hysteresis with value Δ for the cycle closure is

$$P = \frac{1}{\sqrt{2\pi}\sqrt{3\sigma^2}} \exp\left(-\frac{-\Delta^2}{6\sigma^2}\right)$$

The value of σ which maximizes the above probability gives the maximum likelihood estimate of the standard deviation associated with each free energy calculation:

$$\hat{\sigma} = \frac{\Delta}{\sqrt{3}}$$

Note that in the above derivation, the estimated free energies and the associated errors are the optimal estimates based on all of the information gained from FEP calculations. If the underlying force field in the calculations has systematic bias, the perfectly converged FEP calculation results may still deviate from the experimental measured values, which is something that cannot be corrected based on the above analysis.

The above model can be easily generalized to more complicated cases with more members in the cycle closure and several cycles sharing the same edge. Here, we give the best free energy estimators for the mutation path used for the CDK2 ligands calculation (Figure 9) and leave the derivation as an exercise for the reader.

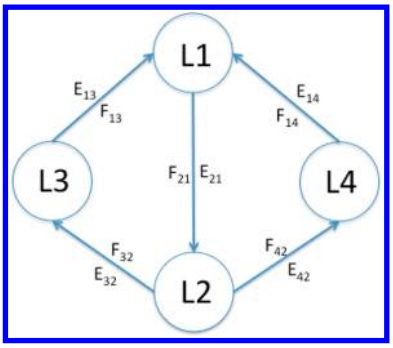


Figure 9. A more complicated mutation path with two cycles sharing the same edge used in the CDK2 inhibitors binding affinity calculation.

$$\hat{F}_{21} = \frac{1}{2}E_{21} - \frac{1}{4}(E_{42} + E_{14}) - \frac{1}{4}(E_{32} + E_{13})$$

$$\hat{F}_{32} = \frac{5}{8}E_{32} - \frac{3}{8}\left(E_{13} + \frac{2}{3}E_{21} - \frac{1}{3}(E_{14} + E_{42})\right)$$

$$\hat{F}_{42} = \frac{5}{8}E_{42} - \frac{3}{8}\left(E_{14} + \frac{2}{3}E_{21} - \frac{1}{3}(E_{13} + E_{32})\right)$$

$$\hat{F}_{13} = \frac{5}{8}E_{13} - \frac{3}{8}\left(E_{32} + \frac{2}{3}E_{21} - \frac{1}{3}(E_{14} + E_{42})\right)$$

$$\hat{F}_{14} = \frac{5}{8}E_{14} - \frac{3}{8}\left(E_{42} + \frac{2}{3}E_{21} - \frac{1}{3}(E_{13} + E_{32})\right)$$

$$\hat{F}_{43} = \frac{1}{2}(E_{13} - E_{14}) + \frac{1}{2}(E_{42} - E_{32})$$

APPENDIX B

Decomposition of Free Energy into Different Components

With a coupling parameter t, the free energy change from reactant ($t = 0$) and product ($t = 1$) is

$$\Delta G = \int_0^1 \left\langle \frac{\partial U}{\partial t} \right\rangle_t \mathrm{d}t$$

To make the alchemical change more flexible, a multicomponent coupling parameter, $\lambda(t)$ is introduced:

$$\Delta G = \int_0^1 \left\langle \nabla_\lambda U \cdot \frac{\mathrm{d}\lambda}{\mathrm{d}t} \right\rangle_\lambda \mathrm{d}t = \sum_i \int_0^i \left\langle \frac{\partial U}{\partial \lambda_i} \right\rangle_\lambda \mathrm{d}\lambda_i$$

Usually, different components of λ act on different energy terms, bonded, angle, electrostatic, van der Waals, etc.; the decomposition of the total free energy onto energy terms follows

$$\Delta G_i = \int_0^1 \left\langle \frac{\partial U}{\partial \lambda_i} \right\rangle_\lambda \mathrm{d}\lambda_i$$

Because the potential energy is a sum of parts, atom based decomposition also works. If we split the system into ligand (l), protein (p), and solvent (s)

$$U = U_\mathrm{l} + U_\mathrm{p} + U_\mathrm{s} + U_\mathrm{lp} + U_\mathrm{ls} + U_\mathrm{ps}$$

The free energy decomposition follows

$$\Delta G = \Delta G_\mathrm{l} + \Delta G_\mathrm{lp} + \Delta G_\mathrm{ls}$$

$$\Delta G_j = \sum_i \int_0^1 \left\langle \frac{\partial U_j}{\partial \lambda_i} \right\rangle_\lambda \mathrm{d}\lambda_i$$

The terms not involving the ligand do not have dependence on the coupling parameters; therefore they do not appear in the components.

Given converged energy perturbations, we use the Bennett-Acceptance-Ratio method to approximate the thermodynamic integrals.[45] If the path of coupling progression is chosen appropriately (neighboring λ windows have substantial overlapped regions in phase space) and the correlations among the exponential averages cancel, then the sum of the Bennett free energy components should be equal to the total free energy. In practical applications, we compare the total free energy computed using BAR with the sum of the free energy due to all components, and if they are close enough, it indicates that the approximation is reasonable. While this still does not guarantee per term equivalence of the Bennett and TI values, it is good enough in practice—we only use the components to gain insight, and no matter how the free energy composition is computed, it does not change the total free energy change.[45] These decomposition schemes ensure the parts sum up to the total.

ASSOCIATED CONTENT

Supporting Information

The structures of CDK2 inhibitors and their compound IDs used in this study. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: robert.abel@schrodinger.com, lingle.wang@schrodinger.com.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank David Mobley, Michael Shirts, John Chodera, Vijay Pande, Richard Friesner, and Bruce Berne for many helpful discussions.

REFERENCES

(1) Gilson, M. K.; Zhou, H. X. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21.
(2) Jorgensen, W. L. *Acc. Chem. Res.* **2009**, *42*, 733.
(3) Mobley, D. L.; Dill, K. A. *Structure* **2009**, *17*, 498.
(4) Deng, Y.; Roux, B. *J. Phys. Chem. B* **2009**, *113*, 2234.
(5) Gallicchio, E.; Levy, R. M. *Curr. Opin. Struct. Biol.* **2011**, *21*, 161.
(6) Mobley, D. L.; Dill, K. A.; Chodera, J. D. *J. Phys. Chem. B* **2008**, *112*, 938.
(7) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 813.
(8) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2008**, *130*, 2831.
(9) Wang, L.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 1326.
(10) de Amorim, H. L.; Caceres, R. A.; Netz, P. A. *Curr. Drug Targets* **2008**, *9*, 1100.
(11) Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2009**, *131*, 15403.

(12) Chipot, C.; Rozanska, X.; Dixit, S. *J. Comput.-Aided Mol. Des.* **2005**, *19*, 765.
(13) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. *Curr. Opin. Struct. Biol.* **2011**, 21.
(14) Wang, L.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 1937.
(15) Huang, X.; Hagen, M.; Kim, B.; Friesner, R. A.; Zhou, R.; Berne, B. J. *J. Phys. Chem. B* **2007**, *111*, 5405.
(16) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 13749.
(17) Wang, L.; Friesner, R. A.; Berne, B. J. *J. Phys. Chem. B* **2011**, *115*, 9431.
(18) Hardcastle, I. R.; Arris, C. E.; Bentley, J.; Boyle, F. T.; Chen, Y.; Curtin, N. J.; Endicott, J. A.; Gibson, A. E.; Golding, B. T.; Griffin, R. J.; Jewsbury, P.; Menyerol, J.; Mesguiche, V.; Newell, D. R.; Noble, M. E.; Pratt, D. J.; Wang, L. Z.; Whitfield, H. J. *J. Med. Chem.* **2004**, *47*, 3710.
(19) Malumbres, M.; Barbacid, M. *Nat. Rev. Cancer* **2009**, *9*, 153.
(20) Chen, Y. N.; Sharma, S. K.; Ramsey, T. M.; Jiang, L.; Martin, M. S.; Baker, K.; Adams, P. D.; Bair, K. W.; Kaelin, W. G., Jr. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 4325.
(21) Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10253.
(22) Steiner, D.; Oostenbrink, C.; Diederich, F.; Zurcher, M.; van Gunsteren, W. F. *J. Comput. Chem.* **2011**, *32*, 1801.
(23) Lai, B.; Oostenbrink, C. *Theor. Chem. Acc.* **2012**, *131*, 1.
(24) Jiang, W.; Roux, B. *J. Chem. Theory Comput.* **2010**, *6*, 2565.
(25) Woods, C. J.; Essex, J. W.; King, M. A. *J. Phys. Chem. B* **2003**, *107*, 13703.
(26) Patriksson, A.; van der Spoel, D. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2073.
(27) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529.
(28) *Desmond*; 3.4 ed.; Schrodinger, LLC: New York, 2012.
(29) Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. In *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*; ACM: Tampa, FL, 2006; p 84.
(30) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. *J. Chem. Theory Comput.* **2010**, *6*, 1509.
(31) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
(32) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474.
(33) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981; p 331.
(34) *Maestro*; 9.3 ed.; Schrodinger, LLC: New York, 2012.
(35) *Schrodinger Suite 2012 Protein Preparation Wizard*; Schrodinger, LLC: New York, 2012.
(36) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 525.
(37) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245.
(38) Paliwal, H.; Shirts, M. R. *J. Chem. Theory Comput.* **2011**, *7*, 4115.
(39) Hahn, A. M.; Then, H. *Phys. Rev. E* **2009**, *80*, 031111.
(40) Shenfeld, D. K.; Xu, H.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. *Phys. Rev. E* **2009**, *80*, 046705.
(41) Li, J.; Abel, R.; Zhu, K.; Cao, Y.; Zhao, S.; Friesner, R. A. *Proteins* **2011**, *79*, 2794.
(42) Borrelli, K. W.; Cossins, B.; Guallar, V. *J. Comput. Chem.* **2010**, *31*, 1224.
(43) Yang, C. Y.; Sun, H.; Chen, J.; Nikolovska-Coleska, Z.; Wang, S. *J. Am. Chem. Soc.* **2009**, *131*, 13709.
(44) Gao, C.; Park, M. S.; Stern, H. A. *Biophys. J.* **2010**, *98*, 901.
(45) Deng, Y.; Roux, B. *J. Chem. Phys.* **2008**, *128*, 115103.