Published in final edited form as:

J Chem Inf Model. 2019 May 28; 59(5): 2093–2102. doi:10.1021/acs.jcim.8b00959.

Ranking Reversible Covalent Drugs: From Free Energy Perturbation to Fragment Docking

Han Zhang, Wenjuan Jiang, Payal Chatterjee, and Yun Luo*

Western University of Health Sciences, College of Pharmacy, Pomona CA 91766

Abstract

Reversible covalent inhibitors have drawn increasing attention in drug design, as they are likely more potent than noncovalent inhibitors and less toxic than covalent inhibitors. Despite those advantages, the computational prediction of reversible covalent binding presents a formidable challenge because the binding process consists of multiple steps and quantum mechanics (QM) level calculation is needed to estimate the covalent binding free energy. It has been shown that the dissociation rates and the equilibrium dissociation constants vary significantly even with similar warheads, due to noncovalent interactions. We have previously used a simplistic two-state model for predicting the relative binding selectivity of reversible covalent inhibitors (*J. Am. Chem. Soc.* 2017, 139, 17945). Here we go beyond binding selectivity and demonstrate that it is possible to use free energy perturbation (FEP) molecular dynamics (MD) to calculate the overall reversible covalent binding using a specially designed thermodynamic cycle. We show that FEP can predict the varying binding free energies of the analogs sharing a common α -ketoamide warhead. More importantly, our results revealed that the chemical modification away from warhead changes the binding affinity at both noncovalent and covalent binding states, and the computational prediction can be improved by considering the binding free energy of both states. Furthermore, we explored the possibility of using a more rapid computational method, Site-Identification by Ligand Competitive Saturation (SILCS), to rank the same set of reversible covalent inhibitors. We found that the fragment docking to a set of precomputed SILCS FragMaps produces a reasonable ranking. In conclusion, two independent approaches provided consistent results that the covalent binding state is suitable for the initial ranking of the reversible covalent drug candidates. For lead-optimization, the FEP approach described here can provide more rigorous and detailed information regarding how much the covalent and noncovalent binding states are contributing to the overall binding affinity, thus offering a new avenue for fine tuning the noncovalent interactions for optimizing reversible covalent drugs.

Graphical Abstract

*corresponding author: luoy@westernu.edu.

SUPPORTING INFORMATION

Co-crystal structure of calpain-1 and a ketoamide compound is shown in Figure S1, with two calcium ions and EF-hand motifs highlighted. Difference in amino acid sequences between rat and human calpain-1 is shown on the protein backbone in Figure S2. Convergence of the total free energy for all ligands during FEP/ λ -REMD simulations at covalent bound state and noncovalent bound state is shown in Figure S3. Partial charge on ketoamide warhead is reported in Table S1. G_{solv} results from both 64 and 32 replicas of FEP/ λ -REMD simulations are reported in Table S2. The SILCS docking results are shown in Figure S4–S5.

INTRODUCTION

Reversibility is an essential covalent regulation mechanism in biology. Reversible covalent interactions pervade a broad range of cellular functions from enzyme reactions to DNA modifications. In the field of drug design, reversible covalent drugs have drawn increasing attention as they provide longer residence time, higher potency, and less drug resistance compared with conventional noncovalent drugs¹. In the meanwhile, a reversible covalent bond formation is under thermodynamic control and requires multiple noncovalent interactions to stabilize the complex. Hence the likelihood of forming an off-target complex is reduced relative to irreversible covalent drugs^{2–6}. Reversible covalent chemistry has been strategically applied to target proteins, especially kinases and proteases^{4–12}. Pioneering works have shown that modification of the warhead substituent can convert irreversible warheads (e.g., acrylate, acrylamide, vinylsulfone) to reversible warheads^{13–14}. Taking advantage of the thermodynamic equilibrium, chemists can now fine-tune the noncovalent binding to prolong the residence time of reversible covalent inhibitors from minutes to days¹⁵. Most interestingly, the dissociation rates and the equilibrium dissociation constants vary significantly even with similar warheads, due to noncovalent interactions¹³.

Conducting a perfect chemical orchestra between covalent and noncovalent binding states of reversible covalent drugs requires the ability to predict the reversible covalent binding affinity and how much each binding state is contributing to the overall affinity. Two strategies can be used to improve the binding affinity and specificity: modifying the electrophilicity of the warhead or modifying the noncovalent interaction without affecting the warhead¹⁶. A combination of those two strategies quickly increases the number of compounds that need to be synthesized, and thus a faster computational prediction is required. Unfortunately, computational prediction of reversible covalent binding is more challenging than for noncovalent drugs. This is because the binding process consists of multiple steps and the covalent binding step requires expensive quantum mechanics (QM)-based free energy calculations. Over the years, various covalent docking and fragment-based approaches have been proposed to address this urgent need for tools to facilitate covalent drug design^{17–19}.

We have previously studied reversible covalent binding process using a two-state binding model (Fig 1)⁶. According to Eq. 1, the overall binding affinity depends on both the noncovalent state binding free energy ΔG_{dm} and the covalent state binding free energy ΔG_{dc} . Only when the covalent binding is much stronger than noncovalent binding, the noncovalent state will be negligible (e.g., if $\Delta G_{\text{dc}} - \Delta G_{\text{dm}} < -5.5$ kcal/mol, $e^{-\beta\Delta G_{\text{dc}}}$ will be 10^4 times larger than $e^{-\beta\Delta G_{\text{dm}}}$). However, in order to know whether the noncovalent binding is sufficiently weaker to be negligible, one still has to use a QM-based method to calculate the covalent binding free energy. In this article, we propose a FEP-based approach that allows us to utilize Eq. 1 to estimate the reversible covalent binding free energy without QM calculation. This is achieved by designing a thermodynamic cycle illustrated in Fig. 2.

$$\frac{1}{K_{\text{d}}} = e^{-\beta\Delta G_{\text{tot}}} = \frac{[E \cdot I] + [EI]}{[E][I]} = e^{-\beta\Delta G_{\text{dc}}} + e^{-\beta\Delta G_{\text{dm}}} \quad \text{Eq. 1}$$

As a proof of concept, we applied this method on α -ketoamide analogs, which covalently bind to the catalytic site of calcium-dependent cysteine proteases, calpains. Calpains are involved in a wide range of physiological functions such as cell motility, cell cycling, and transcriptional regulation. Calpain overactivation is associated with cataract formation, heart attack, stroke, Alzheimer's disease, and traumatic spinal cord and brain injury²⁰. Therefore, potent and selective calpain inhibitors are highly sought after. In addition, α -ketoamide is one of the few warheads that can be extended to the prime side (C-terminus of the scissile peptide bond), providing additional chemical space to further improve the binding and selectivity.

Apart from using a two-state binding model and FEP calculation mentioned above, we also explored a more rapid computational method for covalent drug design. Site-Identification by Ligand Competitive Saturation (SILCS)^{21–23} method generates 3D probability binding distributions of chemical functional groups on protein surfaces, called “FragMaps”, by performing multiple MD simulations of the target protein in an aqueous solution of small molecule representatives of different chemical functional groups. It has been shown that the SILCS method is capable to efficiently rank the binding free energy of noncovalent drugs, with consideration of protein flexibility and desolvation in rational drug design^{24–28}. Here we generated calpain-1 FragMaps using different protonation states of the catalytic site and investigated how to use FragMaps to rank covalent drugs.

METHODS

The thermodynamic cycle for reversible covalent binding FEP calculation

The thermodynamic cycle is designed to calculate the overall reversible covalent binding free energies for a set of analogs sharing the same warhead. According to Figure 2, the absolute binding free energy at the noncovalent state can be first calculated for a reference structure $\Delta G_{\text{dm}}^{\text{ref}}$ (Eq. 2). The covalent binding free energy $\Delta G_{\text{dc}}^{\text{ref}}$ is then estimated using the

calculated $\Delta G_{\text{dm}}^{\text{ref}}$ and experimental K_d (Eq. 3). We chose nine analogs that contain structural modifications remote (non-conjugated) from the warhead so that the electrophilicity of the warhead remains the same as in the reference compound. We then calculated the relative binding free energy of each analog $\Delta\Delta G_{\text{dm}}$ (Eq. 4) and ${}^c\Delta\Delta G_{\text{dc}}$ (Eq. 5) in comparison with the reference compound. Using Eq. 6 and 7 we obtain the absolute binding free energy of each ligand at noncovalent ($\Delta G_{\text{dm}}^{\text{L}}$) and covalent state ($\Delta G_{\text{dc}}^{\text{L}}$). Here, absolute means the free energy of the ligand in binding site relative to in bulk solvent. Combining $\Delta G_{\text{dm}}^{\text{L}}$ and $\Delta G_{\text{dc}}^{\text{L}}$ using Eq. 8 yields the total reversible covalent binding free energy ($\Delta G_{\text{tot}}^{\text{L}}$) of each ligand.

$$\Delta G_{\text{dm}}^{\text{ref}} = \Delta G_{\text{noncov}}^{\text{ref}} - \Delta G_{\text{solv}}^{\text{ref}} \quad (2)$$

$$\Delta G_{\text{dc}}^{\text{ref}} = -\frac{1}{\beta} \ln \left(\frac{1}{K_d} - e^{-\beta \Delta G_{\text{dm}}^{\text{ref}}} \right) \quad (3)$$

$$\Delta\Delta G_{\text{dm}} = \Delta G_{\text{dm}}^{\text{L}} - \Delta G_{\text{dm}}^{\text{ref}} = \Delta G_{\text{noncov}} - \Delta G_{\text{solv}} \quad (4)$$

$$\Delta\Delta G_{\text{dc}} = \Delta G_{\text{dm}}^{\text{L}} + \Delta G_{\text{mc}}^{\text{L}} - (\Delta G_{\text{dm}}^{\text{ref}} + \Delta G_{\text{mc}}^{\text{ref}}) = \Delta G_{\text{cov}} - \Delta G_{\text{solv}} \quad (5)$$

$$\Delta G_{\text{dm}}^{\text{L}} = \Delta G_{\text{dm}}^{\text{ref}} + \Delta\Delta G_{\text{dm}} \quad (6)$$

$$\Delta G_{\text{dc}}^{\text{L}} = \Delta G_{\text{dm}}^{\text{L}} + \Delta G_{\text{mc}}^{\text{L}} = \Delta G_{\text{dc}}^{\text{ref}} + \Delta\Delta G_{\text{dc}} \quad (7)$$

$$\Delta G_{\text{tot}}^{\text{L}} = -\frac{1}{\beta} \ln \left(e^{-\beta \Delta G_{\text{dc}}^{\text{L}}} + e^{-\beta \Delta G_{\text{dm}}^{\text{L}}} \right) \quad (8)$$

System setup and simulation protocols

The initial covalent binding complex coordinates were obtained from a co-crystal structure of a rat calpain-1 with a covalently bound ketoamide compound (PDB 2R9C). In the co-

crystal structure, calpain-1 is in its active state with two calcium ions bound in each EF-hand motif (Figure S1). A hemi-thioetal complex is formed between a catalytic cysteine and the ketoamide. Although the rat calpain-1 was used in a previous study⁶, we mutated *in silico* the rat calpain-1 sequence to human's using *Molecular Operating Environment* (MOE)²⁹ to be consistent with the experimental assay (Figure S2)^{30–31}. To prepare the noncovalent binding complex, the covalent bond was cleaved using MOE. The catalytic cysteine was converted back to anion to form the Cys(115)-His(272) ion pair, and the ligand was converted back to ketoamide. All analogs were built using MOE for both covalent and noncovalent states.

CHARMM36 force field³² was used for all simulations. CHARMM36 force fields for deprotonated cysteine (resname CYM) and protonated histidine (resname HSP) were used as in the previous study⁶. All ligands were prepared using CHARMM General Force Field (CGenFF)^{33–34}. The compounds used in this study has the structural modification at least two carbon distance away from the warhead and is not conjugated with the warhead. The ketoamide warhead was re-parameterized previously⁶. All the MD simulation systems in explicit solvent were prepared by using the CHARMM-GUI^{35–36}. Each system was solvated into a rectangular water box consisted of CHARMM TIP3P water molecules³⁷ and 150 mM KCl, with an edge distance of 10 Å. All the MD simulations were performed using NAMD2.12³⁸ under periodic boundary conditions at a constant temperature of 300 K and pressure of 1 atm (NPT ensemble)³⁹. A smoothing function was applied to van der Waals forces between 10 and 12 Å. The solvated complexes were minimized and equilibrated using a stepwise procedure set up by CHARMM-GUI. The partial charge of the ketoamide warhead was also calculated at HF/6–31G* level for each noncovalently bound ligand to ensure that the warhead electrophilicity remains similar among analogs (Table S1).

Free energy perturbation/ λ -exchange molecular dynamics (FEP/ λ -REMD)

The thermodynamic cycle in Fig.2 can be in principle calculated using any alchemical free energy technique and simulation software⁴⁰. Here we applied FEP/ λ -REMD implemented in NAMD2.12 program using generalized scalable multiple copy algorithms^{41–42}. Each system was equilibrated for 20 ns before starting the FEP/ λ -REMD simulations. A thermodynamic coupling parameter λ was used to perturb the whole reference compound or the R group from a fully coupled ($\lambda = 1$) to a fully decoupled state ($\lambda = 0$). In the fully decoupled state, there is no non-bonded interaction between the perturbed atoms and the environment. The absolute binding free energy calculation follows the same protocol as a previous study.⁴³ A total number of 128 replicas ($96\lambda_{\text{dispersive}}, 16\lambda_{\text{repulsive}}, 16\lambda_{\text{electrostatic}}$) were used for absolute free energy $\Delta G_{\text{dm}}^{\text{ref}}$. 64 replicas ($48\lambda_{\text{dispersive}}, 8\lambda_{\text{repulsive}}, 8\lambda_{\text{electrostatic}}$) were used for relative free energy at noncovalent and covalent states ($\Delta\Delta G_{\text{dm}}$ and $\Delta\Delta G_{\text{dc}}$), as well as in solvent (ΔG_{solv}) for each ligand. Each replica in the FEP/ λ -REMD simulation represents a state along the coupling parameter λ , and a periodic swap is attempted between neighboring replicas every 100 steps (0.2 ps). The accuracy of FEP/ λ -REMD depends on the overlaps between two potential energy distributions, which can be measured by the acceptance ratio between replicas^{44–46}. The number of the replicas was thus chosen to ensure the average acceptance ratio range of 0.75 to 0.84 in all FEP/ λ -REMD simulations, indicating that the

overlap in the phase space is large. Convergence was monitored by the time dependence of each predicted free energy term. This sampling time dependence provides an asymptotically unbiased estimator for each ΔG . We considered the FEP/ λ -REMD simulation is converged when the last four 100 ps running average of free energy value fluctuates within 0.5 kcal/mol. With the sufficient number of replicas and frequent exchange across all replicas, the convergence for all tested ligands was achieved fairly fast, within 2 ns per replica (Figure S3). Within the available computational resources, we also duplicated all ΔG_{solv} calculations using 32 replicas ($20\lambda_{\text{dispersive}}, 6\lambda_{\text{repulsive}}, 6\lambda_{\text{electrostatic}}$) using different initial velocities and produced consistent results as 64 replicas ($48\lambda_{\text{dispersive}}, 8\lambda_{\text{repulsive}}, 8\lambda_{\text{electrostatic}}$) (Table S2). The average differences from two sets of FEPs are within reported standard deviation in Table 1.

Site identification by ligand competitive saturation (SILCS)

Site Identification by Ligand Competitive Saturation (SILCS) was performed using the MolCal program (SilcsBio, LLC) and GROMACS simulation program⁴⁷. SILCS were conducted for two sets of protonation states of calpain-1 catalytic site: deprotonated Cys(115)- protonated His(272) ion pair or uncharged Cys115 and His272. Position restraint of 1000 kJ mol⁻¹nm⁻² was used for Cys115 to make sure the sidechain of catalytic cysteine in SILCS remains the same position as in FEP simulation. The SILCS simulation uses an iterative Grand Canonical Monte Carlo (GCMC) and MD protocol, in which GCMC involves the sampling of water and various small molecules representing different functional groups, with the subsequent MD simulation for the sampling of protein conformational dynamics⁴⁸. In this approach, benzene, propane, methanol, formamide, imidazole, acetaldehyde, methylammonium and acetate were used as drug-like functional groups. During the simulations, these functional groups and water molecules compete with each other to bind with the protein. From ten copies of 100 ns trajectories, the three-dimensional (3D) probability distributions of the functional groups, called FragMaps, were calculated and then converted into a free energy representation, termed Grid Free Energy (GFE), based on a Boltzmann distribution for quantitative analysis. Force field for protein, water, ions and small molecules remains the same as in FEP calculations. The overlap coefficients are between 0.75 to 0.97 for all FragMaps obtained in current study.

Since the binding pocket is known, a local-sampling algorithm in SILCS was used to dock each ligand into the calpain-1 FragMaps generated above. For each ligand, two dockings were first performed using the whole ligand in original form (ketoamide) and in tertiary alcohol form (transition state mimic). One more fragment docking was performed for each ligand using its methyl-capped R group (Table 2). Initial ligand conformation is extracted from the last snapshot of the fully coupled ($\lambda = 1$) state of the FEP/ λ -REMD simulation. The Monte Carlo based SILCS (SILCS-MC) protocol⁴⁹ was applied to predict ligand binding poses using SILCS GFE as the scoring function. In this study, we utilized FragMaps based on generic apolar (benzene, propane), generic H-bond donor (methanol O, formamide N, imidazole (NH)) and acceptor (methanol O, formamide O, imidazole N, acetaldehyde O), negatively charged (acetate) and positively charged (methylammonium) for scoring. For each docking, five independent SILCS-MC runs were performed with a two-step mode.

Each SILCS-MC run involved 50 cycles of 100 Metropolis MC steps and 1000 steps of simulated annealing (SA). MC steps sample a wide range of binding poses with step-size of 1 Å for translations, 9° for maximal rigid rotations per step and 9° for maximal dihedral angle rotations per step. This setup is to ensure sufficient sampling of the covalent binding poses sampled by the fully coupled ($\lambda = 1$) state of the FEP/ λ -REMD simulation while preventing ligands from flipping at the binding pocket. SA steps are designed to find a local minimum with the acceptance criteria defined by the Ligand Grid Free Energies (LGFE). SA involved maximum ranges of step-size of 0.2 Å for translations, 9° for maximal rigid rotations per step and 9° for maximal dihedral angle rotations per step with gradual cooling from 300 to 0 K. The SILCS-MC runs were repeated for five times during which each cycle initiated with a different random seed and was repeated until it finished 10 cycles. The lowest LGFE and average LGFE scores from the minimum score in five independent jobs were obtained. SILCS-MC run with 50 cycles of 1000 MC steps and 5000 SA steps were also tested and yielded the similar LGFE values and produced the same correlation.

RESULTS and DISCUSSION

Free energy perturbation results

The reversible covalent binding scheme of α -ketoamide inhibitor in calpain-1 is shown in Figure 3 in the noncovalent binding state (left) and covalent binding state (right). Using those two states, we calculated the binding free energy of ten α -ketoamide reversible covalent inhibitors (Table 1), based on the thermodynamic cycle described in Figure 2. For the reference compound, we calculated the absolute binding free energy at noncovalent binding state $\Delta G_{\text{dm}}^{\text{ref}}$ to be -6.11 ± 0.39 kcal/mol. Using the experimentally reported inhibitory constant K_i 0.20 μM^{31} and the equation 3, we estimated the absolute binding free energy at covalent state $\Delta G_{\text{dc}}^{\text{ref}}$ to be -9.25 ± 0.34 kcal/mol. The binding free energy of other nine ligands at noncovalent and covalent states are calculated relative to this reference compound. All free energy calculation results are reported in Table 1.

For each ligand, we then plotted the noncovalent binding free energy $\Delta G_{\text{dm}}^{\text{L}}$, covalent binding free energy $\Delta G_{\text{dc}}^{\text{L}}$, and the total reversible covalent binding free energy $\Delta G_{\text{tot}}^{\text{L}}$ against their experimental free energy converted from reported K_i values³¹. Figure 4B shows that there is no correlation between experimental and noncovalent state binding free energy ΔG_{dm} (Spearman's ranking correlation of 0.40, Pearson correlation coefficient of 0.41). By contrast, we obtained a fairly strong correlation between experimental and covalent binding free energy ΔG_{dc} (Spearman's ranking correlation of 0.81, Pearson correlation coefficient of 0.84) (Fig. 4C), suggesting that the relative binding affinity of the covalent binding state alone can be used to rank the overall binding affinity of the current compound set. This relationship between ΔG_{dc} and ΔG_{tot} can be easily explained from the data in Table 1. For nine out of ten compounds calculated, the free energy of noncovalent binding has negligible contribution to the ΔG_{tot} , compared with covalent binding (${}^{\text{c}}\Delta G_{\text{dc}} - \Delta G_{\text{dm}} < -3.5$ kcal/mol, $e^{-\beta\Delta G_{\text{dc}}}$ is at least 340 times larger than $e^{-\beta\Delta G_{\text{dm}}}$). The data clearly indicate that covalent

ligands with the same warhead and electrophilicity show varying binding free energies in the covalent binding state. Comparing among compound 1–9, the aromatic R group is more favorable than aliphatic R group, except for compound 7. Adding a methoxy group (compound 3 and 4) or a hydroxyl group (compound 2) on the benzene stabilizes the covalent product with higher binding affinity. However, compound 5 suggests that shorten the linker between benzene and the scaffold may reduce the binding affinity. Comparing the three pyridine-containing R groups, nitrogen at 2-position (compound 6 and 8) is more favorable than nitrogen at 4-position (compound 7). Interestingly, none of the observed structure-affinity relationship described above holds for the noncovalent binding state. Considering the α -ketoamide backbone conformation is different from the covalently bonded hemithioacetal, which has the sp^3 carbon center (Fig. 3), it is not surprising that a slight difference in binding pose between the two states leads to the difference in protein-ligand interaction.

Perhaps the most exciting result is the compound 1. Replacing the benzyl group in the reference compound by ethanol group dramatically altered the two-state free energy landscape. The binding affinity at noncovalent binding state increases and the affinity at covalent state decreases. The difference between the two states is nearly identical ($\Delta G_{dc} - \Delta G_{dm} = -0.15 \pm 0.35$ kcal/mol). Hence, compound 1's binding free energy at noncovalent state ΔG_{dm} has a significant contribution to its ΔG_{tot} . The slight improvement in the correlation obtained from ΔG_{tot} (Spearman correlation improved from 0.83 to 0.85 and Pearson coefficient improved from 0.84 to 0.85) may well reflect the benefit of combining both states (Fig. 4D). One may argue that only one compound is not sufficient to make such a conclusion, however, it is easier to understand that, under thermodynamic equilibrium, a competitive inhibitor at the enzyme catalytic site will inhibit the enzyme no matter with the covalent bond or not. Therefore, when the binding affinity is comparable, both states will contribute to the inhibition constant K_i .

SILCS Results

SILCS simulations were conducted for two types of protonation states of the calpain-1 catalytic site: type A contains uncharged Cys115 and His272, and type B contains deprotonated Cys115- protonated His272 ionic pair (Table 2). Figure 5 shows the type A FragMaps overlaid on the calpain-1 surface. A Monte Carlo based local-sampling algorithm in SILCS is suitable to dock ligands with known binding pocket. We first attempted to dock each ligand using the whole ligand in original form (ketoamide) and tertiary alcohol form (transition state mimic) (Table 2). To ensure the docking poses are close to the covalent binding pose, a flat-bottom harmonic distance restraint of 2 Å was applied between the sulfur of the Cys115 and the proximal carbonyl carbon of the ketoamide warhead. The predicted LGFE scores ($LGFE_L$ for the lowest LGFE scores, $LGFE_A$ for the average LGFE scores) were obtained for each ligand. The LGFE values are calculated as the LGFE scores of each ligand relative to the reference compound. However, unlike FEP/ λ -REMD results above, no correlation was obtained between experimental relative binding free energies and LGFE scores, no matter which FragMaps or which form of the ligand was used (Figure S4 and S5).

Through visual inspection, we found that even with the distance restraint, the ketoamide or the tertiary alcohol form do not entirely reproduce the conformation of the covalently bound hemithioketal structure, regardless of the protonation state of the catalytic residues. Since we have learned from our FEP/ λ -REMD results that the structural modifications in R groups do not alter the binding pose of the ligand scaffold and the covalent binding state alone is sufficient to rank the overall binding affinity of the current compound set, we reasoned that ranking the R group using the FragMaps may correlate better with the covalent binding affinity and avoid the discrepancy at the warhead during the docking. We hence docked the methyl-capped R groups instead of the whole ligand into the two pre-computed FragMaps. While the fragment docking using FragMaps type B did not produce any correlation (Figure S5 AB), the results from FragMaps type A is quite promising. The LGFE scores of each R groups relative to the reference compound generated using FragMaps type A are shown in Table 3. The correlations between LGFE values and experimental relative binding free energies are still lower than the ones obtained from FEP/ λ -REMD calculation above (Pearson 0.58 for lowest LGFE and 0.60 for average LGFE). However, after eliminating the R group of compound 5, a strong correlation (Pearson 0.86) was obtained for both lowest LGFE values and average LGFE values (Figure 6AB). We speculated that compound 5 contains the bulkiest R group of all analogs that may induce protein conformational change upon binding. Such ligand-induced conformation change is of course not captured by pre-computed FragMaps.

The result from non-covalent docking of R-group might appear to contradict the covalent binding. However, the fact that only R-group docking, not the whole ligand docking, yields reasonably good correlation is actually consistent with our FEP finding. First, because the covalent bond cannot be modeled in current SILCS docking, the docking of the whole ligand would mimic the non-covalent binding state, which has no correlation with the experimental data, as shown by the FEP results above. When only the R-group is docked, the binding pose of the R-group is not altered by the warhead conformation, thus avoiding the problem caused by the missing covalent bond. Secondly, between two different protonation states of the catalytic residues, type A FragMaps containing uncharged Cys115 and His272, similar as the charge at the covalent bond state. While type B FragMaps are generated using Cys115-His272 ionic pair, representing the charge at noncovalent bond state. Only type A FragMaps was able to produce reasonable correlation further confirmed the R-group docking is consistent with the FEP calculation at covalent binding state. Lastly, the large size and the flexibility of the peptidomimetic compounds we used here makes it more challenging to predict binding pose and affinity using LGFE score. When using a small R-group, the prediction is more reliable, especially here we already knew that the R-group binds at the prime region of the catalytic site.

CONCLUSION AND PERSPECTIVES

To our knowledge, this is the first time that the total binding free energy of reversible covalent binder is calculated using FEP MD simulation. Our free energy results are in general consistent with the previous report that the binding of the covalent state is stronger than the noncovalent state. However, using a small set of compounds in the current study, it is sufficient to demonstrate that the noncovalent binding state may play a nontrivial role

towards the overall binding, even though the ΔG_{dm} can be safely neglected for the majority of the analogs. The most popular covalent drug design strategy nowadays is to first identify a lead compound that binds to the protein target through noncovalent interaction. Then a chosen warhead will be added to the lead compound through chemical synthesis with the hope that a more potent lead compound can be obtained through an additional covalent bond formation. This type of strategy is promising but also challenging as we show here that noncovalent binding state does not correlate with the total reversible covalent binding affinity. The calculated free energy of both states also pointed out the fact that structural modification away from the warhead can change the covalent binding affinity. This is because the change in the orientation of the noncovalent binding pose can subsequently change the following covalent binding step, even though the warhead electrophilicity remains the same. Therefore, the current study further emphasized that the knowledge of the two binding states is important for designing reversible covalent drugs.

On the other hand, we demonstrated that a set of precomputed FragMaps could be used to rank reversible covalent binders sharing the same warhead. Unlike for conventional noncovalent drugs, we have shown that extra care is needed regarding choosing the protonation states of the catalytic residues and the similarity of the fragments. Despite those limitations, our data suggest that fragment docking provides a rapid method for ranking the reversible covalent analogs. The two independent approaches provided consistent results that the covalent binding state is suitable for the initial ranking of the reversible covalent drug candidates. Notably, the SILCS approach has the unique advantage for rapid screening because the FragMaps only need to be computed once. For lead-optimization, the FEP approach proposed here is not only a more rigorous prediction but also provides detailed information regarding how much the covalent and noncovalent binding states are contributing to the overall binding affinity, offering a new avenue for fine tuning the noncovalent interactions to obtain highly selective reversible covalent drugs.

One of the limitations of free energy approach is the assumption of thermodynamic equilibrium. However, with the interplay of the rate of distribution and rate of elimination, the equilibrium at protein target may not be achieved. This would explain why certain drugs' residence time ($1/k_{\text{off}}$) correlates drug efficacy better than the dissociation constant (K_{d})⁵⁰. The situation is even more complicated for reversible covalent drugs. The stability of the covalent complex depends on the reversibility of the warhead as well as the noncovalent interactions that stabilize/destabilize the reversible covalent binding. Due to the barrier of the chemical reaction, slow binding might affect the formation of the chemical equilibrium. Therefore, investigating the time dependence of the reversible covalent binding process is desirable. In the current study, the experimental inhibition constants were measured without taking into account the possible slow binding or enzyme autolysis. If the initial velocities of the substrate hydrolysis were determined before reaching the chemical equilibrium, the inhibition constant would be overestimated, thus underestimate binding affinity. It is perhaps one of the reasons that our calculated binding free energies from FEP, which assume the chemical equilibrium, are more negative than the experimental binding free energy converted from K_{i} . Unfortunately, such speculation can only be examined if the time dependency can be experimentally measured. Further investigation is needed, and the

assumption of thermodynamic equilibrium will depend on the type of warhead as well. A covalent inhibitor can sometimes be considered irreversible in one target but reversible in another⁵¹. When the covalent binding becomes irreversible, the thermodynamic equilibrium is not satisfied; instead, the free energy barrier between the dissociation state and the transition state will be the dominant factor in determining the drug efficacies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENT

This work was supported by NIH NS090397-01A1 and NSF XSEDE research allocation MCB160119. Authors thank the SilcsBio team, especially Prof. Alexander MacKerell and Dr. Sunhwan Jo for their support and suggestions.

REFERENCES

1. Baillie TA, Targeted Covalent Inhibitors for Drug Design. *Angew Chem Int Ed Engl* 2016, 55 (43), 13408–13421. [PubMed: 27539547]
2. Mah R; Thomas JR; Shafer CM, Drug discovery considerations in the development of covalent inhibitors. *Bioorg. Med. Chem. Lett* 2014, 24 (1), 33–9. [PubMed: 24314671]
3. Backus KM; Correia BE; Lum KM; Forli S; Horning BD; Gonzalez-Paez GE; Chatterjee S; Lanning BR; Teijaro JR; Olson AJ; Wolan DW; Cravatt BF, Proteome-wide covalent ligand discovery in native biological systems. *Nature* 2016, 534 (7608), 570–4. [PubMed: 27309814]
4. Singh J; Petter RC; Baillie TA; Whitty A, The resurgence of covalent drugs. *Nat. Rev. Drug Discov* 2011, 10 (4), 307–17. [PubMed: 21455239]
5. Mehdi S, Covalent Enzyme Inhibition in Drug Discovery and Development. In *Enzyme Technologies: Pluripotent Players in Discovering Therapeutic Agents*, Yang H-C, W.-K. Y., and McCarthy JR, Ed. John Wiley and Sons: Hoboken, New Jersey, 2014; pp 81–129.
6. Chatterjee P; Botello-Smith WM; Zhang H; Qian L; Alsamarah A; Kent D; Lacroix JJ; Baudry M; Luo Y, Can Relative Binding Free Energy Predict Selectivity of Reversible Covalent Inhibitors? *J Am Chem Soc* 2017, 139 (49), 17945–17952. [PubMed: 29124934]
7. Forster M; Chaikuad A; Dimitrov T; Doring E; Holstein J; Berger BT; Gehring M; Ghoreschi K; Muller S; Knapp S; Laufer SA, Development, Optimization, and Structure-Activity Relationships of Covalent-Reversible JAK3 Inhibitors Based on a Tricyclic Imidazo[5,4- d]pyrrolo[2,3- b]pyridine Scaffold. *J. Med. Chem* 2018, 61 (12), 5350–5366. [PubMed: 29852068]
8. Smith S; Keul M; Engel J; Basu D; Eppmann S; Rauh D, Characterization of Covalent-Reversible EGFR Inhibitors. *ACS Omega* 2017, 2 (4), 1563–1575. [PubMed: 29963651]
9. Bandyopadhyay A; Gao J, Targeting biomolecules with reversible covalent chemistry. *Curr. Opin. Chem. Biol* 2016, 34, 110–116. [PubMed: 27599186]
10. Ticha A; Stanchev S; Vinothkumar KR; Mikles DC; Pachi P; Began J; Skerle J; Svehlova K; Nguyen MTN; Verhelst SHL; Johnson DC; Bachovchin DA; Lepsik M; Majer P; Strisovsky K, General and Modular Strategy for Designing Potent, Selective, and Pharmacologically Compliant Inhibitors of Rhomboid Proteases. *Cell Chem Biol* 2017, 24 (12), 1523–1536 e4. [PubMed: 29107700]
11. Awoonor-Williams E; Walsh AG; Rowley CN, Modeling covalent-modifier drugs. *Biochim. Biophys. Acta* 2017.
12. Lonsdale R; Burgess J; Colclough N; Davies NL; Lenz EM; Orton AL; Ward RA, Expanding the Armory: Predicting and Tuning Covalent Warhead Reactivity. *J. Chem. Inf. Model* 2017.
13. Serafimova IM; Pufall MA; Krishnan S; Duda K; Cohen MS; Maglathlin RL; McFarland JM; Miller RM; Frodin M; Taunton J, Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. *Nat. Chem. Biol* 2012, 8 (5), 471–6. [PubMed: 22466421]

14. Schirmeister T; Kesselring J; Jung S; Schneider TH; Weickert A; Becker J; Lee W; Bamberger D; Wich PR; Distler U; Tenzer S; Johe P; Hellmich UA; Engels B, Quantum Chemical-Based Protocol for the Rational Design of Covalent Inhibitors. *J. Am. Chem. Soc* 2016, 138 (27), 8332–5. [PubMed: 27347738]
15. Bradshaw JM; McFarland JM; Paavilainen VO; Bisconte A; Tam D; Phan VT; Romanov S; Finkle D; Shu J; Patel V; Ton T; Li X; Loughhead DG; Nunn PA; Karr DE; Gerritsen ME; Funk JO; Owens TD; Verner E; Brameld KA; Hill RJ; Goldstein DM; Taunton J, Prolonged and tunable residence time using reversible covalent kinase inhibitors. *Nat Chem Biol* 2015, 11 (7), 525–31. [PubMed: 26006010]
16. Hallenbeck KK; Turner DM; Renslo AR; Arkin MR, Targeting Non-Catalytic Cysteine Residues Through Structure-Guided Drug Discovery. *Curr Top Med Chem* 2017, 17 (1), 4–15. [PubMed: 27449257]
17. London N; Miller RM; Krishnan S; Uchida K; Irwin JJ; Eidam O; Gibold L; Cimermancic P; Bonnet R; Shoichet BK; Taunton J, Covalent docking of large libraries for the discovery of chemical probes. *Nat. Chem. Biol* 2014, 10 (12), 1066–72. [PubMed: 25344815]
18. Scarpino A; Ferenczy GG; Keseru GM, Comparative Evaluation of Covalent Docking Tools. *J. Chem. Inf. Model* 2018, 58 (7), 1441–1458. [PubMed: 29890081]
19. Miller RM; Paavilainen VO; Krishnan S; Serafimova IM; Taunton J, Electrophilic fragment-based design of reversible covalent kinase inhibitors. *J. Am. Chem. Soc* 2013, 135 (14), 5298–301. [PubMed: 23540679]
20. Ono Y; Saido TC; Sorimachi H, Calpain research for drug discovery: challenges and potential. *Nat. Rev. Drug Discov* 2016, 15 (12), 854–876. [PubMed: 27833121]
21. Guvench O; MacKerell AD Jr., Computational fragment-based binding site identification by ligand competitive saturation. *PLoS Comput Biol* 2009, 5 (7), e1000435. [PubMed: 19593374]
22. Raman EP; Yu W; Lakkaraju SK; MacKerell AD Jr., Inclusion of multiple fragment types in the site identification by ligand competitive saturation (SILCS) approach. *J. Chem. Inf. Model* 2013, 53 (12), 3384–98. [PubMed: 24245913]
23. Yu W; MacKerell AD Jr., Computer-Aided Drug Design Methods. *Methods Mol Biol* 2017, 1520, 85–106. [PubMed: 27873247]
24. Cardenas MG; Yu W; Beguelin W; Teater MR; Geng H; Goldstein RL; Oswald E; Hatzi K; Yang SN; Cohen J; Shakhovich R; Vanommeslaeghe K; Cheng H; Liang D; Cho HJ; Abbott J; Tam W; Du W; Leonard JP; Elemento O; Cerchiatti L; Cierpicki T; Xue F; MacKerell AD Jr.; Melnick AM, Rationally designed BCL6 inhibitors target activated B cell diffuse large B cell lymphoma. *J Clin Invest* 2016, 126 (9), 3351–62. [PubMed: 27482887]
25. Cheng H; Linhares BM; Yu W; Cardenas MG; Ai Y; Jiang W; Winkler A; Cohen S; Melnick A; MacKerell A Jr.; Cierpicki T; Xue F, Identification of Thiourea-Based Inhibitors of the B-Cell Lymphoma 6 BTB Domain via NMR-Based Fragment Screening and Computer-Aided Drug Design. *J Med Chem* 2018, 61 (17), 7573–7588. [PubMed: 29969259]
26. Yu WB; Lakkaraju SK; Raman EP; Fang L; MacKerell AD, Pharmacophore Modeling Using Site-Identification by Ligand Competitive Saturation (SILCS) with Multiple Probe Molecules. *Journal of Chemical Information and Modeling* 2015, 55 (2), 407–420. [PubMed: 25622696]
27. Lanning ME; Yu WB; Yap JL; Chauhan J; Chen LJ; Whiting E; Pidugu LS; Atkinson T; Bailey H; Li W; Roth BM; Hynicka L; Chesko K; Toth EA; Shapiro P; MacKerell AD; Wilder PT; Fletcher S, Structure-based design of N-substituted 1-hydroxy-4-sulfamoyl-2-naphthoates as selective inhibitors of the Mcl-1 oncoprotein. *Eur J Med Chem* 2016, 113, 273–292. [PubMed: 26985630]
28. Liang DD; Robinson E; Hom K; Yu WB; Nguyen N; Li Y; Zong QS; Wilks A; Xue FT, Structure-based design and biological evaluation of inhibitors of the pseudomonas aeruginosa heme oxygenase (pa-HemO). *Bioorg Med Chem Lett* 2018, 28 (6), 1024–1029. [PubMed: 29459206]
29. Inc., C. C. G. Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc: Montreal, QC, Canada, 2016.
30. Li Z; Patil GS; Golubski ZE; Hori H; Tehrani K; Foreman JE; Eveleth DD; Bartus RT; Powers JC, Peptide alpha-keto ester, alpha-keto amide, and alpha-keto acid inhibitors of calpains and other cysteine proteases. *J. Med. Chem* 1993, 36 (22), 3472–80. [PubMed: 8230139]

31. Li Z; Ortega-Vilain AC; Patil GS; Chu DL; Foreman JE; Eveleth DD; Powers JC, Novel peptidyl alpha-keto amide inhibitors of calpains and other cysteine proteases. *J. Med. Chem* 1996, 39 (20), 4089–98. [PubMed: 8831774]
32. Best RB; Zhu X; Shim J; Lopes PE; Mittal J; Feig M; Mackerell AD Jr., Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J Chem Theory Comput* 2012, 8 (9), 3257–3273. [PubMed: 23341755]
33. Vanommeslaeghe K; Hatcher E; Acharya C; Kundu S; Zhong S; Shim J; Darian E; Guvench O; Lopes P; Vorobyov I; Mackerell AD Jr., CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem* 2010, 31 (4), 671–90. [PubMed: 19575467]
34. Yu W; He X; Vanommeslaeghe K; MacKerell AD Jr., Extension of the CHARMM General Force Field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comput. Chem* 2012, 33 (31), 2451–68. [PubMed: 22821581]
35. Jo S; Kim T; Iyer VG; Im W, CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem* 2008, 29 (11), 1859–65. [PubMed: 18351591]
36. Lee J; Cheng X; Swails JM; Yeom MS; Eastman PK; Lemkul JA; Wei S; Buckner J; Jeong JC; Qi Y; Jo S; Pande VS; Case DA; Brooks CL 3rd; MacKerell AD Jr.; Klauda JB; Im W, CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J Chem Theory Comput* 2016, 12 (1), 405–13. [PubMed: 26631602]
37. Jorgensen WL; Chandrasekhar J; Madura JD; Impey RW; Klein ML, Comparison of Simple Potential Functions for Simulating Liquid Water. *Journal of Chemical Physics* 1983, 79 (2), 926–935.
38. James CP; Rosemary B; Wei W; James G; Emad T; Elizabeth V; Christophe C; Robert DS; Laxmikant K; Klaus S, Scalable molecular dynamics with NAMD. *J. Comput. Chem* 2005, 26 (16), 1781–1802. [PubMed: 16222654]
39. Adelman S; Garrison B, Generalized Langevin theory for gas/solid processes: Dynamical solid models. *J Chem Phys* 1976, 65 (9), 3751–3761.
40. Loeffler HH; Bosio S; Duarte Ramos Matos G; Suh D; Roux B; Mobley DL; Michel J, Reproducibility of Free Energy Calculations across Different Molecular Simulation Software Packages. *J. Chem. Theory Comput* 2018, 14 (11), 5567–5582. [PubMed: 30289712]
41. Jiang W; Phillips JC; Huang L; Fajer M; Meng Y; Gumbart JC; Luo Y; Schulten K; Roux B, Generalized Scalable Multiple Copy Algorithms for Molecular Dynamics Simulations in NAMD. *Comput Phys Commun* 2014, 185 (3), 908–916. [PubMed: 24944348]
42. Jiang W; Hodosscek M; Roux B, Computation of Absolute Hydration and Binding Free Energy with Free Energy Perturbation Distributed Replica-Exchange Molecular Dynamics (FEP/REMD). *J Chem Theory Comput* 2009, 5 (10), 2583–2588. [PubMed: 21857812]
43. Alsamarah A; LaCuran AE; Oelschlaeger P; Hao JJ; Luo Y, Uncovering Molecular Bases Underlying Bone Morphogenetic Protein Receptor Inhibitor Selectivity. *PLoS One* 2015, 10 (7), e0132221. [PubMed: 26133550]
44. Pohorille A; Jarzynski C; Chipot C, Good practices in free-energy calculations. *J. Phys. Chem. B* 2010, 114 (32), 10235–53. [PubMed: 20701361]
45. Shirts MR; Chodera JD, Statistically optimal analysis of samples from multiple equilibrium states. *J Chem Phys* 2008, 129 (12), 124105. [PubMed: 19045004]
46. Bennett CH, Efficient Estimation of Free-Energy Differences from Monte-Carlo Data. *J. Comput. Phys* 1976, 22 (2), 245–268.
47. Hess B; Kutzner C; van der Spoel D; Lindahl E, GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation* 2008, 4 (3), 435–447. [PubMed: 26620784]
48. Lakkaraju SK; Raman EP; Yu W; MacKerell AD Jr., Sampling of Organic Solutes in Aqueous and Heterogeneous Environments Using Oscillating Excess Chemical Potentials in Grand Canonical-like Monte Carlo-Molecular Dynamics Simulations. *J. Chem. Theory Comput* 2014, 10 (6), 2281–2290. [PubMed: 24932136]

49. Faller CE; Raman EP; MacKerell AD Jr.; Guvench O, Site Identification by Ligand Competitive Saturation (SILCS) simulations for fragment-based drug design. *Methods Mol Biol* 2015, 1289, 75–87. [PubMed: 25709034]
50. Copeland RA; Pompliano DL; Meek TD, Drug-target residence time and its implications for lead optimization. *Nat. Rev. Drug Discov* 2006, 5 (9), 730–9. [PubMed: 16888652]
51. Chandrasekaran A; Shen L; Lockhead S; Oganessian A; Wang J; Scatina J, Reversible covalent binding of neratinib to human serum albumin in vitro. *Drug Metab. Lett* 2010, 4 (4), 220–7. [PubMed: 20690900]

Figure 1. A simplified free energy profile of two-state binding for a reversible covalent binder. Equation 1 shows the relation between the association constant, $1/K_d$, and the total binding free energy G_{tot} , the binding free energy of the covalent state G_{dc} and noncovalent state G_{dm} . Each binding state and the corresponding free energy term are illustrated on the free energy profile.

Figure 2. The thermodynamic cycle of reversible covalent binding.

The cycle illustrates the steps to calculate the ligand binding free energy at noncovalent state $\Delta G_{\text{dm}}^{\text{L}}$ and covalent state $\Delta G_{\text{dc}}^{\text{L}}$ using five free energy terms in the figure colored in megenta or blue. The relative free energy terms between a ligand and the reference compound are in blue, which include free energy of transforming R to R' in solvent ΔG_{solv} , in noncovalent state ΔG_{noncov} , and in covalent state ΔG_{cov} . The absolute binding free energy terms (magenta) are only calculated for a reference compound at noncovalent state, which includes the free energy of decoupling whole reference compound from its environment in solvent $\Delta G_{\text{solv}}^{\text{ref}}$ and at noncovalent binding state $\Delta G_{\text{noncov}}^{\text{ref}}$. The same color code is used in Eq. 2, 4, and 5.

Figure 3.

(A) The reversible covalent binding scheme of α -ketoamide catalyzed by calpain-1. (B) Binding pose of the reference ligand in calpain-1 catalytic site is shown in noncovalent (left) and covalent state (right). Protein backbone is shown in white. The reference compound is shown in licorice and catalytic Cys115 and His272 are shown in CPK mode in atom color code: cyan carbon, blue nitrogen, red oxygen, and yellow sulfur. All hydrogen atoms are omitted for clarity.

Figure 4.

(A) The functional groups of nine ligands are shown in black, and the one for the reference ligand is shown in magenta. The scaffold structure is shown in Figure 3. (B, C, D)

Correlation between experimental and calculated noncovalent state binding free energies, covalent state binding free energies, and total binding free energies. All free energy data are in kcal/mol unit. Experimental data were obtained using $G(\text{exp}) = RT \ln(K_i)$, in which K_i is the inhibitory constants for each compound in human calpain-1 from ref³¹. R is gas constant, and T is room temperature at 300 Kelvin. For each plot, R^2 , Pearson R , and Spearman ρ are reported. The error bars represent the standard deviation reported in Table 1.

Figure 5. SILCS FragMaps and the calpain-1 crystal structure. Apolar (green), H-bond donor (blue) and acceptor (red), negatively charged (yellow) and positively charged (cyan) SILCS FragMaps are shown. All the FragMaps are set to a cutoff of -1.2 kcal/mol. Residue Cys115 of the calpain-1 is shown in CPK mode. All hydrogen atoms are omitted for clarity.

Figure 6.

(A) Correlations between experimental relative binding free energy and calculated lowest LGFE values. All free energy data are in kcal/mol unit. Experimental data were obtained same as described before. **(B)** Correlations between experimental relative binding free energy and calculated average LGFE values based on five independent runs. The error bar of average LGFE was calculated using the minimum LGFE from five independent runs. For each plot, R^2 , Pearson R, and Spearman ρ . The compound5 is excluded in both plots as an outlier.

Table 1.Calculated free energies^a for each ligand in human calpain-1.

Ligand	R group	ΔG_{solv}	ΔG_{dm}	ΔG_{dc}	ΔG_{tot}^b	$\Delta G_{dc} - \Delta G_{dm}^c$
Reference		-6.84(0.17)	-6.11(0.39)	-9.25(0.34)	-9.25(-0.34)	-3.14(0.51)
1		4.60(0.07)	-8.65(0.33)	-8.80(0.10)	-9.15(-0.16)	-0.15(0.35)
2		-11.61(0.32)	-7.54(0.49)	-13.81(0.24)	-13.81(-0.24)	-6.27(0.55)
3		-7.42(0.23)	-8.72(0.31)	-17.32(0.28)	-17.32(-0.28)	-8.60(0.42)
4		-9.59(0.24)	-5.10(0.35)	-16.03(0.25)	-16.03(-0.25)	-10.93(0.43)
5		0.17(0.23)	-3.14(0.53)	-9.48(0.57)	-9.48(-0.57)	-6.34(0.70)
6		1.56(0.27)	-4.45(0.41)	-15.02(0.48)	-15.02(-0.48)	-10.57(0.63)
7		-0.33(0.18)	-5.89(0.41)	-9.12(0.30)	-9.12(-0.30)	-3.23(0.51)
8		-8.03(0.17)	-4.76(0.44)	-11.37(0.23)	-11.37(-0.23)	-6.61(0.49)
9		-1.51(0.18)	-5.89(0.38)	-9.90(0.24)	-9.90(-0.24)	-4.01(0.45)

^a Each free energy term is illustrated in the thermodynamic cycle in Figure 2. The unit is kcal/mol. The computed values and their standard deviation in parentheses were calculated via block averages using last four blocks of 100 ps.

^{b.} $\Delta G_{\text{tot}}^L = -\frac{1}{\beta} \ln \left(e^{-\beta \Delta G_{\text{dc}}^L} + e^{-\beta \Delta G_{\text{dm}}^L} \right)$. The standard deviation of ΔG_{tot} is calculated based on uncertainty propagation using equation

$$-\frac{1}{\beta} \sqrt{\frac{\left(e^{-\beta \Delta G_{\text{dc}}^L} * \beta \sigma_{\Delta G_{\text{dc}}} \right)^2 + \left(e^{-\beta \Delta G_{\text{dm}}^L} * \beta \sigma_{\Delta G_{\text{dm}}} \right)^2}{e^{-\beta \Delta G_{\text{dc}}^L} + e^{-\beta \Delta G_{\text{dm}}^L}}}.$$

^{c.} Standard deviation of $\Delta G_{\text{dc}} - \Delta G_{\text{dm}}$ is calculated using $\sqrt{\sum_i \sigma_i^2}$, assuming the uncertainties are uncorrelated.

Table 2.

Two types of FragMaps generated based on different protonation states of catalytic residues in calpain-1 protease and the ligand scaffolds used in SILCS docking.

FragMaps	Cys115 protonation	His272 protonation	SILCS docking ligand scaffold
Type A (Figure 5)			Methyl-capped R group
			Ketoamide ligand
			Tertiary alcohol ligand
Type B			Methyl-capped R group
			Ketoamide ligand
			Tertiary alcohol ligand

Table 3.

Calculated SILCS LGFE values (kcal/mol) for each R group in human calpain-1.

Ligand	ΔLGFE_L	ΔLGFE_L^a	ΔLGFE_A^b	ΔLGFE_A	$\Delta\Delta G_{\text{exp}}^c$
Reference	-1.75	0.00	-1.61 (0.08)	0.00(0.00)	0.00
1	-0.86	0.89	-0.84(0.02)	0.77(0.07)	0.83
2	-2.47	-0.71	-2.44(0.02)	-0.83(0.07)	0.38
3	-2.26	-0.50	-2.22(0.03)	-0.61(0.09)	-0.26
4	-2.56	-0.81	-2.49(0.05)	-0.88(0.06)	-0.36
5	-1.97	-0.22	1.87(0.08)	-0.26(0.12)	1.46
6	-0.81	0.94	-0.74(0.07)	0.87(0.08)	0.69
7	-0.91	0.84	-0.84(0.04)	0.77(0.05)	1.02
8	-1.75	0.00	-1.64(0.09)	-0.03(0.09)	0.43
9	-1.46	0.29	-1.35(0.07)	0.26(0.03)	0.70

^a LGFE values are the relative LGFE between each R group and reference R group.

^b The reported average LGFE values were calculated by utilizing the minimum LGFE value in five independent runs for each R group.

^c G_{exp} values are the experimental relative binding affinity between each ligand and reference ligand. The number in parentheses are standard deviations.