JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Accurate Binding Free Energy Predictions in Fragment Optimization

Thomas B. Steinbrecher,*[,†] Markus Dahlgren,[‡] Daniel Cappel,[†] Teng Lin,[‡] Lingle Wang,[‡] Goran Krilov,[‡] Robert Abel,[‡] Richard Friesner,[§] and Woody Sherman[‡]

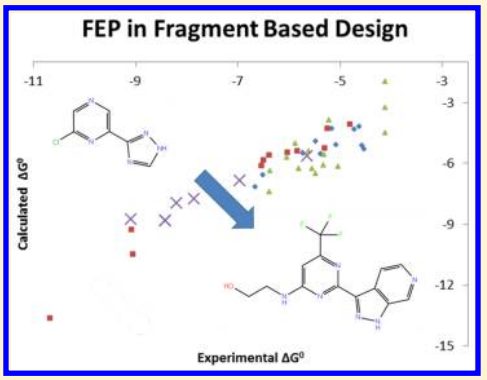
[†]Schrödinger GmbH, Dynamostrasse 13, 68165 Mannheim, Baden-Württemberg, Germany

[‡]Schrödinger Inc., 120 West 45th Street, 17th Floor, New York, New York 10036, United States

[§]Department of Chemistry, Columbia University, 3000 Broadway New York, New York 10027, United States

Supporting Information

ABSTRACT: Predicting protein−ligand binding free energies is a central aim of computational structure-based drug design (SBDD) — improved accuracy in binding free energy predictions could significantly reduce costs and accelerate project timelines in lead discovery and optimization. The recent development and validation of advanced free energy calculation methods represents a major step toward this goal. Accurately predicting the relative binding free energy changes of modifications to ligands is especially valuable in the field of fragment-based drug design, since fragment screens tend to deliver initial hits of low binding affinity that require multiple rounds of synthesis to gain the requisite potency for a project. In this study, we show that a free energy perturbation protocol, FEP+, which was previously validated on drug-like lead compounds, is suitable for the calculation of relative binding strengths of fragment-sized compounds as well. We study several pharmaceutically relevant targets with a total of more than 90 fragments and find that the FEP+ methodology, which uses explicit solvent molecular dynamics and physics-based scoring with no parameters adjusted, can accurately predict relative fragment binding affinities. The calculations afford R^2-values on average greater than 0.5 compared to experimental data and RMS errors of ca. 1.1 kcal/mol overall, demonstrating significant improvements over the docking and MM-GBSA methods tested in this work and indicating that FEP+ has the requisite predictive power to impact fragment-based affinity optimization projects.



INTRODUCTION

Fragment-based drug design (FBDD) has evolved from a niche technique 20 years ago into a powerful approach used throughout the pharmaceutical and biotech industry.[1−6] Multiple fragment derived compounds are currently in clinical trials, with one compound, vemurafenib, gaining FDA approval in 2011.[7] In brief, the FBDD approach uses libraries of small molecules (fragments, often defined as following the "rule of three"[8] with molecular weights < 300 and ClopP < 3 instead of the "rule of five" in use for drug-like compounds[9]) to find hits that bind with a high degree of ligand efficiency, which can then be optimized. Due to the extraordinarily rapid increase in the size of the accessible chemical space as the molecular size is increased,[10] a library of thousands of fragments will offer much higher coverage of the chemical space in its size class than an HTS library with millions of drug-like compounds.

FBDD typically aims for the discovery of weak (e.g., millimolar to high micromolar) inhibitors via biochemical assays or biophysical techniques such as SPR, TSA, ITC, and NMR. Weak initial binders can then be optimized through multiple rounds of chemical modifications, ideally resulting in larger and more drug-like lead compounds with high affinity and specificity. Two possible strategies for fragment optimization involve either *growing* a single fragment via introduction of chemical R-groups to fill empty regions of a binding site or *linking* two or more fragment inhibitors with nonoverlapping binding modes via appropriate linking groups. Both techniques make heavy use of structural information, which explains why FBDD programs typically contain an aggressive structural aspect via NMR or X-ray crystallography. The FBDD approach has produced an impressive and growing number of success stories in recent years, with dozens of compounds from fragment studies in clinical trials[11] and multiple mature commercial and academic FBDD programs in existence today.

While the idea to use computational techniques to predict the binding of fragment probes to binding sites was explored from the very start,[12−15] modern *in silico* structure-based drug design approaches to optimize the affinity of fragment hits have not played a major role in most FBDD efforts. There are several potential reasons for this, including the following: 1) ligand docking algorithms and scoring functions that have been parametrized to describe the binding of drug-like compounds may not perform as well for fragment-sized compound; 2) the binding of fragment inhibitors, which often exhibit weak directional interactions and more dynamic binding modes, may pose particularly challenging for an accurate description using fast, empirical methods; 3) since FBDD programs have

Received: August 26, 2015
Published: October 12, 2015

DOI: 10.1021/acs.jcim.5b00538
J. Chem. Inf. Model. 2015, 55, 2411−2420

typically grown around the application of particular experimental techniques, their underutilization of computational approaches might be based on their evolutionary history.

Here, we will apply free energy calculation techniques (the FEP methodology in particular) to the study of fragment-sized molecules. FEP calculations are based on molecular dynamics simulations that explicitly consider conformational flexibility and entropy effects through the use of a physics-based force field to describe molecular interactions and explicit solvent to model the real environment of the protein binding site. One of the goals of this study is to explore whether a rigorous free energy approach, such as FEP, can provide predictive power and insights for FBDD without any special parametrization or customization.

Since attempts to improve ligand potency usually go hand in hand with increases in molecular weight, FBDD studies often use various metrics of ligand efficiency to judge if a gain in potency is worth the required addition of new functional groups.[16,17] This concept of focusing on the effect that small chemical changes have on the thermodynamics of binding suggests that the results of FEP calculations, which likewise deal with the effect that small chemical perturbations have on the binding free energy, could be well suited for practical FBDD applications.

For the case of applying FEP to conventional lead optimization, we have recently published a study of advanced FEP calculations for ca. 200 ligands over multiple target classes that showed highly significant correlation coefficients between calculated and experimental binding free energies with average errors in the range of 1 kcal/mol.[18] Given this demonstrated accuracy, coupled with the expanding capabilities and accessibility of large-scale computational hardware (e.g., Cloud and GPU), free energy calculations are poised to be more routinely applied in lead optimization efforts, where increased design efficiency is critical. The present study will extend the validation of this calculation protocol, called FEP+, to fragment optimization studies.

RESULTS AND DISCUSSION

System Selection. We selected eight different targets where high-quality binding data for a series of fragment sized ligands and high-quality crystal structures are available. The selection was based on a literature review of recent FBDD studies and, while necessarily subjective, aims at representing a diverse set of biochemical systems and ligand chemical structures that are within the realm of applicability of FEP+. Our data set comprises both well-established test systems of fragment-receptor binding and FBDD case studies on pharmaceutically relevant targets (see Table 1). A brief description of each system is provided below:

Enterobacteria Phage T4 Lysozyme (Lysozyme) L99A mutant is an engineered form of Lysozyme that contains an internal cavity able to bind small molecules.[19] The only artificial receptor in our list, this system has become a widely used test bed of both experimental and theoretical ligand binding studies[20−26] and represents a rather unusual case of specific binding into a near featureless binding site. While not of pharmaceutical interest, this system was chosen as a difficult test case where other computational techniques often fail to accurately predict binding energies.

Bacterial DNA Ligase (DNA Ligase) is an NAD-dependent enzyme responsible for fusing 3′ and 5′ ends of DNA strands.[27] It is essential for bacterial growth and structurally different from the equivalent mammalian enzyme, making it an interesting attack point for novel antibiotic compounds.[28]

Myeloid Cell Leukemia 1 (Mcl-1) is a member of the Bcl-2 protein family that is commonly overexpressed in human cancer[29−31] and allows aberrant cells to avoid programmed cell death.[32,33] Inhibitors of Mcl-1 would be highly appealing for cancer treatment, but so far no promising compounds have entered clinical trials. This is considered a particularly difficult system for small molecule drug design, since Mcl-1 fulfills its role in the complex interactions of the Bcl-2 proteins via protein−protein interactions (PPIs).[34]

Major urinary protein I (MUP-I) is a member of the MUP complex of protein isoforms found in mice.[35,36] It is involved in pheromone binding and release.[37] Unlike the remaining systems studied here, this is an example of a receptor that is naturally evolved to bind fragment sized ligands. Beyond an interest in its allergenic properties, MUP-I has been used as a general test case for the thermodynamics of protein ligand binding.[38−43]

Human lactate dehydrogenase isoform A (LDH) is a key enzyme in glycolysis. Since aggressive tumors heavily rely on glycolysis for energy production (the Warburg effect[44]), LDH inhibitors could interfere with their growth and offer a currently underutilized cancer treatment. The enzyme exists in two types of subunits making up various isoforms, and a promising inhibitor should bind to LDH subunit A but not B.[45]

Janus Kinase 2 (JAK-2) is one of a family of intracellular tyrosine kinases responsible for cell signaling through the JAK-STAT pathway.[46] Gain-of-function mutations in JAK-2 have been implicated in a variety of myeloproliferative disorders,[47] spurring considerable interest in JAK-2 inhibitors as therapeutic agents.

Heat shock protein 90 (HSP90) is a ca. 90 kDa sized member of the HSP molecular chaperone and stress response family and one of the most common proteins in human cytoplasm. Its function is essential for cellular viability, but HSP90 inhibitors are nevertheless considered as therapeutic agents, since HSP90 aids the folding of several protein known to be overexpressed in tumor cells.[48]

Table 1. Structural Data for the Eight Systems Studied in This Work[b]

	species	PDB ID and ref	resolution [Å]	cocrystallized fragment
Lysozyme	enterobacteria phage T4	186L[22]	1.8	*n*-butyl benzene
DNA Ligase	Staph. aureus	4CC5[28]	1.9	triazole-pyrazine fragment
Mcl-1	*Homo sapiens*	4HW3[51]	2.4	fragment-based lead compound
MUP-I	*Mus musculus*	1I06[52]	1.9	2-(*sec*-butyl) thiazole
LDH	*Homo sapiens*	1I10[53]	2.1	NADH, oxamate
JAK-2	*Homo sapiens*	3E64[54a]	1.9	indazole sulfonamide
HSP90	*Homo sapiens*	3FT8[55]	2.0	fragment-based lead compound
p38	*Homo sapiens*	1W7H[56]	2.2	benzyloxy pyridin fragment

[a]For JAK-2, calculations were repeated with another X-ray crystal structure, 3E62, containing a smaller fragment ligand, with overall comparable results. See the Supporting Information for details. [b]See the Supporting Information for fragment chemical structures.

DOI: 10.1021/acs.jcim.5b00538
J. Chem. Inf. Model. 2015, 55, 2411−2420

p38α MAP kinase (p38) belongs to a family of serine/threonine specific kinases mediating cellular stress response.[49] Active p38 is involved in regulating the biosynthesis of TNF and other inflammatory cytokines, making it a promising target to address inflammatory diseases.[50]

Our validation set therefore comprises several diverse pharmaceutically relevant targets (a DNA binder, a PPI, two kinases, a chaperone, and a metabolic enzyme) and two established test systems for fragment-receptor interactions in general (T4 Lysozyme and MUP-I). The X-ray crystal structures used here are relatively high resolution and clearly show the binding mode of one fragment inhibitor in each series (all except LDH, where the bound NADH cofactor and oxamate molecule, while not chemically similar to the studied fragments, at least clearly define their binding site). The systems range from small to medium sized proteins (ca. 150 to 350 residues), but larger receptors could be treated with this approach as well. The nature of the receptor binding sites range from hydrophobic to hydrophilic and from fully occluded to solvent accessible (see the Supporting Information figures). Ligand series consisting of compounds with conserved molecular charge were chosen to avoid potential problems with changing the simulated system's total charge.

Summary of Calculation Results. Succinctly, all free energy results below are obtained by enhanced sampling molecular dynamics based FEP+ calculations perturbing between individual ligand structures, conducted without restraints in explicit aqueous solvation. Systems were prepared by placing compounds into the known receptor binding sites using X-ray crystallography data and molecular docking. For details, see the Computational Methods section below and the detailed system-specific preparation description in the Supporting Information.

Relative binding free energies for ligands within each series were computed directly from the FEP+ calculations. Absolute binding free energy values were derived from the FEP+ results (relative free energies, i.e. $\Delta\Delta G^0$) by adding a single free energy offset to all members of each series (this offset is chosen to minimize the mean unsigned error of the absolute free energy predictions and is added merely to make comparison of experimental and computational results easier; see refs 57 and 58 for details). This uniform offset value facilitates plotting of the predicted versus measured free energies of binding of the individual ligands and does not alter the mean unsigned error or the R^2-value of the predictions. Since quantitative experimental binding free energies were not available for the LDH series (ref 45 reports percent inhibition values), the calculated values for LDH have been omitted from all RMSE averages below. The average correlation between calculated and experimentally determined binding free energies across all ligands in this study is high, with an R^2-value of 0.65 (weighted average over all series) and a root-mean-square (RMS) error of 1.14 kcal/mol. The RMS error is strongly influenced by a small number of outliers, as can be inferred based on the relatively low median absolute error of 0.70 kcal/mol, and two-thirds of absolute error values are below 1 kcal/mol. Overall, the distribution of absolute errors appears approximately Gaussian in shape (Figure 1) with some asymmetry toward larger errors, as observed before for FEP+.[18]

The slope of the theoretical vs experimental free energy plot is close to 1.0 (1.19, Figure 2), suggesting that the underlying physics is being accurately modeled. This is in agreement with the idea that explicit solvent molecular dynamics simulations run at 300 K with molecular mechanics force fields offers an accurate description of the physics associated with molecular recognition.

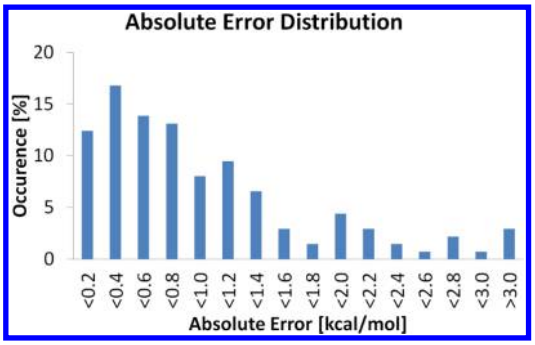


Figure 1. Distribution of absolute errors for all predictions.

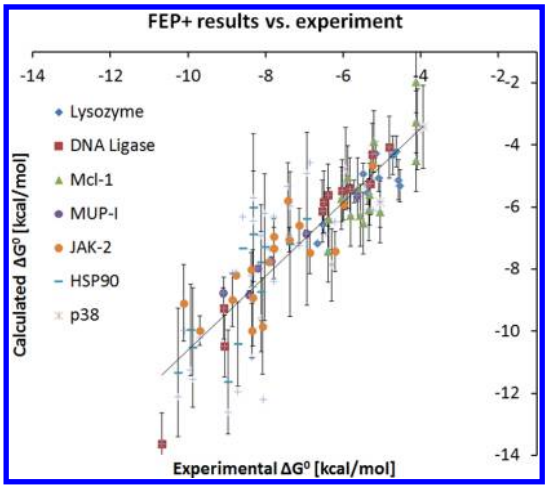


Figure 2. FEP+ results for seven systems compared to experimental data. Vertical error bars give estimated prediction errors from cycle closure analysis. No values for the LDH system are shown as no quantitative experimental data is available. The R^2-value is 0.73 when calculated over all the data shown, while it is 0.65 when the weighted average over all series is taken.

Using the approach described to compute absolute binding free energies, the calculated affinities can be directly compared between different systems, unlike many other SBDD methods (e.g., ligand docking scores and approximate free energy methods). The ability to compute a true binding free energy has significant implications, such as being able to compare energies of ligands binding to different targets (binding selectivity) or the effects of putative drug resistance mutations on ligand binding. In addition, accurately modeling the underling physics should result in transferability across a broad range of targets without the need for parametrization. Tables for all calculated free energy results and comparison to binding free energies derived from experimental data for all series are contained in the Supporting Information.

The average binding free energy for the fragments studied in this work is −6.8 kcal/mol (equivalent to a binding constant in the intermediate micromolar range), which is significantly weaker than for typical lead series but in line with expectations for initial optimization of fragment hits. The weakest binders in the list (for Mcl-1) exceed 1 mM binding and might not even be considered specific inhibitors in typical screening campaigns; while on the other side of the activity range the most active MUP-I ligand is of high nanomolar activity, and the most active

DOI: 10.1021/acs.jcim.5b00538
J. Chem. Inf. Model. 2015, 55, 2411−2420

Table 2. FEP+ Results Overview and Comparison to Experimental Data and Other Binding Energy Prediction Approaches[a]

System	# ligands	RMSE FEP+ [kcal/mol]	Correct Sign	R²-value FEP+	R²-value Glide SP Docking	R²-value MM-GBSA default	R²-value MM-GBSA flexible	R²-value Molecular Weight
Lysozyme	12	0.7	74%	**0.75**	0.32	0.40	0.30	0.24
DNA Ligase	11	0.9	95%	**0.96**	0.36	0.01	0.36	0.91
Mcl-1	15	1.0	96%	**0.62**	0.03	0.34	0.58	0.45
MUP-I	7	0.4	91%	**0.94**	0.92	0.86	0.75	0.93
LDH	14	n.a.*	89%	0.27*	0.02	0.09	**0.37**	0.21
JAK-2	18	1.15	85%	**0.64**	0.50	0.50	0.21	0.32
HSP90	13	2.03	84%	0.60	0.70	**0.71**	0.69	0.64
p38	6	1.25	80%	**0.69**	0.09	0.01	0.00	0.63
Weighted Ave.	**96**	**1.14**	**89%**	**0.65**	0.35	0.37	0.41	0.49

[a]RMS errors are calculated based on the predicted and experimental $\Delta\Delta G^0$-values for all ligand transformations. "Correct Sign" gives the percentage of direct ligand to ligand FEP transformations where the sign of the affinity change was predicted correctly. The table entries give the coefficient of determination R^2 with respect to experimental data for FEP+, two alternative computational techniques, ligand docking and MM-GBSA, as well as the compound molecular weight. Two R^2-values for MM-GBSA results are given, first for the default protocol and second for calculations with additional binding site flexibility. R^2-values are color coded to indicate excellent, good, acceptable, or low prediction qualities (value > 0.60 shown in green, > 0.40 in yellow, > 0.20 in orange, else in red). For each system, the R^2-value of the most successful prediction tool is printed in bold.* For LDH, the coefficient of determination, R^2 (marked with an asterisk), compares the calculated $\Delta\Delta G^0$ from FEP+ with a qualitative experimental activity measure, so no RMSE is given.

(and largest) DNA Ligase ligands are high nanomolar inhibitors as well. Over this entire range of 5 orders of magnitude differences in binding constants, all calculated affinities lie close to the linear trendline in Figure 2, mostly within the range of the estimated error bars. The accuracy of FEP+ results therefore seems not to decline when dealing with exceptionally high or low affinity fragments.

R^2-values and RMSE vary substantially between the test systems, suggesting that the quality of FEP+ results is at least somewhat system-dependent (this was observed in refs 8 and59 as well). For DNA Ligase and MUP-I, exceptionally high R^2-values are found, higher than would theoretically be expected for even two repetitions of experimental binding affinity determination.[60] This indicates that expected errors in both experimental and theoretical binding affinities are exceptionally low or correlated for both systems (both experimental studies use the highly accurate ITC method for at least some affinity measurements).

The calculation results presented here have been obtained using default protocol settings intended to imitate how a comparable free energy calculation study would be conducted in industrial practice. However, there are multiple parameters and setup options that could be varied for each system to obtain better retrospective results. Prospective studies do not benefit from such calculation parameter tuning, but, occasionally, they can lead to a better understanding of the simulated system. Therefore, we give details for some alternative results obtained with different setup conditions in the Supporting Information. Topics explored cover the impact of the MD ensemble used, tested for DNA Ligase, the omission of crystal water molecules, tested for MUP-I, and the difference in results for two different docked starting poses for LDH. The Supporting Information further explores an alternative setup for JAK-2, in which a different protein X-ray structure had to be refined before placing ligands into the binding site, and a possible ligand size bias of the HSP90 binding site.

Comparison to Other Tools. To compare the performance of FEP+ to other common estimators of ligand binding affinities, we have conducted ligand docking calculations using Glide[61,62] as well as MM-GBSA calculations based on the FEP+ starting geometries. For ligand starting geometries obtained from Glide docking, the resulting docking scores were used directly, while for manually placed ligands, Glide's module for scoring-in-place after geometry optimization was used. The correlation coefficients for all three methods as well as a pure molecular weight correlation are also given in Table 2. For seven ligands no Glide or MM-GBSA score was found, and these cases have been omitted from calculating the R^2-values. For MM-GBSA, binding affinities were computed both using the default protocol without binding site conformational search and using a more extensive minimization of the binding site. The comparison of both results indicates a system-dependent large effect of variations in the protocol for MM-GBSA calculations.

FEP+ consistently outperforms the alternative scoring methods studied in this work, yielding the most accurate predictions for six of the systems and the second best for another. For MUP-I and HSP90, all methods perform comparably well. For JAK-2, Glide and MM-GBSA results are comparably predictive to FEP+, as is MM-GBSA for Mcl-1. MM-GBSA outperforms FEP+ for the LDH case if the extended minimization protocol is used, although none of the methods perform particularly well on this system. For LDH only qualitative activity measurements instead of binding free energies were experimentally determined, and uncertainty about the fragment binding modes is largest (see the Supporting Information).

A detailed comparison of the magnitude of the binding energy errors (MUE or RMSE) was not conducted for Glide or MM-GBSA, as these methods do not aim at computing binding free energies and tend to have significantly higher binding energy errors. For example, across all systems, the weighted

DOI: 10.1021/acs.jcim.5b00538
J. Chem. Inf. Model. 2015, 55, 2411−2420

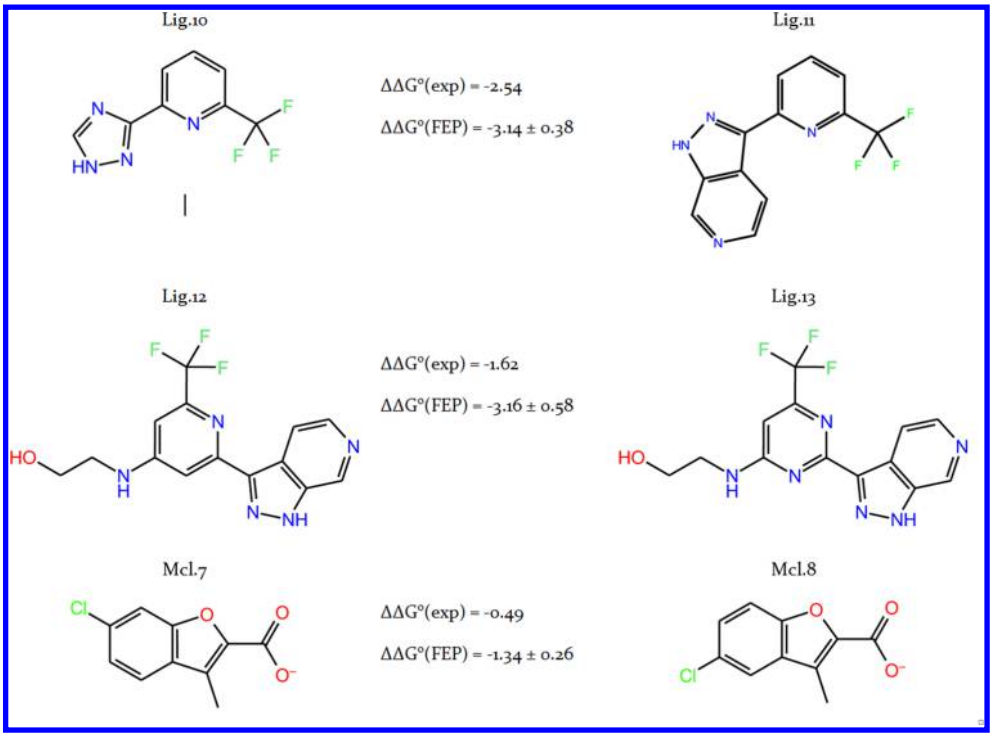


Figure 3. Example ligand transformations with experimental and predicted $\Delta\Delta G^0$-values. In all three cases, FEP+ calculates the sign of the potency effect for a chemical change in the ligand correctly. Ligand modifications, from top to bottom, include replacing a small heterocycle with a large one, a single element substitution with large impact on binding and two substitutional isomers. The examples are for the targets DNA Ligase and Mcl-1.

average RMSE for FEP+ is 1.1 kcal/mol (it is 1.2 kcal/mol averaging over all individual transformations), whereas for MM-GBSA and Glide it is 4.0 and 1.5 kcal/mol, respectively. The RMSE for FEP+ is ca. 25% smaller than for Glide, and the difference in the median absolute error is even more significant (1.45 kcal/mol for Glide versus 0.70 kcal/mol for FEP+). For MM-GBSA the results for JAK-2 and HSP90, where average RMSE-values exceed 15 kcal/mol, were excluded as outliers, otherwise the RMSE for MM-GBSA is 9.3 kcal/mol.

A paired *t* test analysis, taking the R^2-values for each system as observables, indicates FEP to provide significantly (5% threshold) higher correlations to experiment than Glide or either MMGBSA protocol and an improvement just outside of significance comparing to molecular weight (p-value 0.08), which is discussed in the following.

Molecular Weight Correlation. It is interesting to note that a simple molecular weight correlation performs relatively well (better than Glide or MM-GBSA on average for the cases studied here) in ranking ligands for some of the systems. This is likely the result of publication bias, since only fragments that were successfully optimized to some extent result in publication, whereas additions to a fragment that render it inactive are generally not published. A typical fragment-based discovery project starts with a weakly active small compound and improves its affinity by growing it into a larger more potent molecule while maintaining ligand efficiency. For a successful fragment optimization study, an activity-molecular weight correlation would therefore be expected for the published results, while potentially hundreds of inactive, higher molecular weight compounds may have been synthesized but not reported.

For example, in the study on HSP90 inhibitors reported in ref 55, the SAR for an additional ten compound series was explored, but after larger derivatives of the initial bound fragment did not show increases in potency, this series was abandoned in favor of the compounds described below and ultimately published (personal communication).[63] We believe this case to be typical for fragment optimization projects (and lead optimization in general), where negative results are probably common but rarely published.

As an additional example, we have conducted FEP+ calculations on another series of DNA Ligase inhibitors from ref 64, namely the 15 compounds described in Table 3 of that reference. The compound series exhibits no activity-molecular weight correlation (R^2-value of 0.02), while FEP+ predicts their relative activities with a moderate R^2-value of 0.23. While this R^2-value is lower than for the other data set presented above, a better agreement to experiment is not expected for this series, since most of the compounds have very similar or approximate activities, e.g. seven out of 15 have a measured IC_{50} value of "$>200\ \mu M$". The expected R^2-value for predictions on this set of compounds is ca. 0.2 based on guidelines by Brown et al.[60] Another example can be taken from ref 18, where FEP+ calculations for a series of benzamides binding to TYK2 show an R^2-value of 0.79, while no molecular weight correlation (R^2-value <0.01) is observed. Since these additional inhibitors are not from a fragment based drug design study, they have not been added to the eight test systems of this study and are only presented to show that FEP+ is predictive for ligand series with no MW correlation as well.

In practice, the molecular weight of a compound is clearly of little use as an activity predictor in any prospective fragment optimization, as the question to be answered in design is not "*Should substituents be added*" but rather "*What substituent should be added, and where*". To look at this another way, we analyzed pairs of compounds with small differences in molecular weight but significant differences in binding energies. For Lysozyme ligands *o*-xylene/ethylbenzene, DNA Ligase 12/

DOI: 10.1021/acs.jcim.5b00538
J. Chem. Inf. Model. 2015, 55, 2411−2420

13, Mcl-1 3/15, LDH 1g/1h, HSP 15/16, and JAK-2 12/13, the compound pairs have MW within 5 amu but more than 1 kcal/mol activity differences. As such, using MW as a predictor would have no value, whereas FEP+ correctly predicts the more active compound for all six ligand pairs.

Application in Practical Fragment Optimization. The validation data presented here and summarized in Figure 2 show that FEP+ results agree well with experimental data over the entire set of compounds studied. While an aggregate statistical analysis of the data, as presented above, is useful for method validation, it is not necessarily the way such calculations will be employed in practice. Instead, we envision FEP+ to be a tool for guiding idea molecule generation and triaging potential compounds to synthesize. As such, in this section we explore the ability of FEP+ to predict the binding energy difference between pairs of molecules. Ignoring the Lysozyme and MUP-I series, which were not derived from drug discovery campaigns, the remaining systems offer multiple examples of FEP+ results being useful to prioritize ligand modifications.

For DNA Ligase, the improved binding affinity of Lig.4 over Lig.3, due to replacing a ring nitrogen by a carbon atom, is correctly predicted, as well as the loss of affinity when the R1-chloro group of Lig.4 is exchanged for a hydrogen, methyl, or methoxy group (Ligs.5−7). The large gain in potency from exchanging the triazole ring with a 6-*N*-indazole is also correctly predicted (see Figure 3). In addition, the FEP+ results would also have suggested the unintuitive reintroducing of a nitrogen atom in the ortho-position to the indazole (Lig.12 to Lig.13, see Figure 3) to generate the most active ligand in the series. However, the addition of an ethanol-amine side chain to the ligand (Lig.11 to Lig.12) was incorrectly predicted to yield a substantial gain in binding free energy, while the actual effect on potency was negligible. The substituent was likely introduced to tune solvation properties but forms a bifurcated H-bond to the backbone of residue Leu80 in the simulations instead. It is possible that the underlying force field overestimates the strength of this interaction or underestimates the desolvation penalty of the ethanol amine (amines are a group of compounds that exhibited comparably large errors for solvation free energy calculations in OPLS2.1[65]). Overall, the predictions for all pairs of DNA Ligase ligands had the correct sign in 95% (18/19) of the cases.

For the Mcl-1 system, the benefit of introducing any small substituent in the 3, 4, 5, 6, and 7-position compared to the unsubstituted benzofuran, benzothiophen, and indole compounds is uniformly predicted correctly. Also, the greater improvement in binding affinity obtained from the introduction of a second chloro group to the 3-Cl substituted benzothiophene at the 4-position instead of the 6-position is in agreement to experiment (Mcl.2−4). FEP+ provides similar guidance with the benzofuran fragment improvement, where the addition of a chloro group to a 3-methyl substituent yields a better inhibitor in the 6- than the 5-position (Mcl.7 and Mcl.8, see Figure 3), and where a third methyl group in the 6-position in addition to 3- and 7-methyl groups yields a decrease in potency (Mcl.9 and Mcl.10). The calculations mispredict the loss in potency that a 4-bromo substituent would cause and rank Mcl.11 incorrectly as one of the best benzofuran fragments. Likewise, the benefit of introducing a 5-bromo substituent in the indole fragments (Mcl.13 to Mcl.14) is somewhat overestimated, while still ranking the two compounds correctly. Nonetheless, 96% (24/25) of the fragment pairs for Mcl-1 are predicted with the correct sign using FEP+.

The qualitative nature of the LDH activity measurements does not allow such a detailed comparison of single substitution effects, but, in general, FEP+ clearly rewards the introduction of bulky substituents in the 5- or 6-position, which leads to higher experimentally observed inhibition. On the other hand, the loss of activity when the *N*-hydroxy group is removed (LDH.1i to LDH.8) is not reflected in the FEP+ results. Still, 89% (8/9) of the predictions for pairs of LDH fragments have the correct sign when considering pairs in which the experimental activity changes by at least 10%.

For JAK-2, the difference between a phenyl residue in the 5- instead of the 6-position (JAK.2 and JAK.3) is predicted by FEP+ as well as the notable gain in potency from replacing a dimethyl-isoxazolyl (JAK.6) with a dimethyl-thiazolyl moiety (JAK.7) in position 5. For further optimization of the 5-aryl substituent of the aminoimidazole ligands, FEP+ correctly suggests that a sulfonamide in the para-position leads to a highly potent ligand (JAK.13) instead of either the respective amide (JAK.14) or a sulfonamide in the meta-position (JAK.12). The most potent inhibitor (JAK.22) is ranked only fourth best out of the 18 ligands, after three similar compounds differing by small substituents (JAK.13, JAK.20, and JAK.21). Overall FEP+ performs acceptably well for directly comparing the potency of ligand pairs. For the 28 ligand pairs connected by direct transformations in the calculations, 85% (24/28) have the correct sign of $\Delta\Delta G^0$ predicted, a slightly lower percentage than for the previous systems.

A somewhat worse behavior is found for the HSP90 system. FEP+ does not correctly identify the potency effect of several ligand modifications, such as the gain in affinity from adding a 4-methyl group (HSP.2 vs HSP.3, as well as HSP.9 vs HSP.10). On the other hand, FEP+ predicts the potency gain that both a meta-phenyl (HSP.15) or meta-3-pyridyl (HSP.17) substituent have over a meta-4-pyridyl (HSP.16) group, clearly demonstrating the ability to distinguish between chemically similar residues. Ranking the affinity for all HSP90 ligand pairs involved in direct computational transformations is successful in 84% (21/25) of the cases. This worse performance in ranking may be caused by the large RMS error for the HSP90 system predictions (see Table 2), which makes it more likely to incorrectly predict the sign of small changes in binding free energy.

For p38, calculations on only a small number of ligands have been conducted. FEP+ correctly identifies the potency gain in adding two meta-chloro substituents (p38.4) or a naphthyl group (p38.5) to the initial fragment (p38.1) and predicts dimethyl-ethanolamine to be a better substituent than ethanolamine (p38,6 vs p38.7). Of the ten direct ligand transformations conducted, the sign of $\Delta\Delta G^0$ is correctly calculated in 80% (8/10) cases.

Overall, the FEP+ calculations performed here show great promise for prioritizing fragment modifications, even in cases with small changes, and application of the technology in the idea generation stage of fragment optimization should lead to a substantial reduction in the synthetic efforts necessary to create improved inhibitors. For the 116 direct transformations described in this work, FEP+ could identify the better inhibitor in >85% of cases. While the method occasionally ranks compounds incorrectly, the overall effect of employing FEP+ prior to synthesis should eliminate many fragments that do not

DOI: 10.1021/acs.jcim.5b00538
J. Chem. Inf. Model. 2015, 55, 2411−2420

improve activity and substantially improve the quality of newly made compounds.

Discussion. Free energy calculations have long been considered the most rigorous approach to computing protein−ligand binding free energies.[66−70] The FEP+ protocol presented here has previously been found to offer an accurate and predictive way to compute binding free energies for protein−ligand interactions in a general drug design context.[18] Here we show that the good performance of FEP+ extends to the case of fragment binding, as expected, since the underlying physics of protein-fragment binding is not fundamentally different from the binding of drug-like ligands. However, to judge the impact an approach like FEP+ could have on practical fragment-based design, the particularities of such studies need to be considered. For example, unlike experimental assays, free energy calculations can be applied over a large range of ligand activities (i.e., binding can always be detected no matter how weak and the assay will not "bottom out" for tight binders). This is beneficial, as the same protocol can be consistently applied over the entire range of a ligand optimization project from weakest initial binders to highly optimized lead compound, as was demonstrated for the DNA Ligase case in our study. Furthermore, fragment-based drug design efforts typically require greater ligand potency improvements compared to standard drug design studies, which often begin with low micromolar or high nanomolar inhibitors after an initial HTS. In lead optimization, potency is only one of the multiple ligand properties that need to be optimized simultaneously. In contrast, for growing and optimizing fragment hits, potency improvement is initially the primary concern, making FEP+ especially valuable.

FEP+, like all molecular dynamics-based techniques, has high requirements on the quality of the input structural models. For reliable predictions, ideally a high resolution X-ray crystal structure, cocrystallized with a compound representative of molecules within the chemical series to be predicted, is preferred. Uncertainties in ligand tautomer or protonation states would translate into higher uncertainties of computed binding energies. Furthermore, any type of ligand protein interaction beyond the scope of molecular mechanics force fields, such as covalent binding or receptor bound metal ion ligation, could not be treated explicitly without including appropriate corrections. However, systems containing covalent bonds or metal interactions can still be predicted accurately within the FEP+ framework if said interactions are relatively constant throughout the chemical series (data not shown here).

Any structure-based drug design technique requires a good prediction of ligand binding modes before scoring their affinity can be attempted. As the MD calculations in FEP+ offer some conformational sampling, initial ligand poses need only be accurate enough so that the equilibrium binding mode is reached during the course of the simulation. Since routine MD simulations are currently limited to multinanosecond lengths, in practice this does require ligand binding modes either from X-ray crystallography or from a combination of modeling and other experimental information. Therefore, it is not surprising that the case of LDH, where the ligand binding modes are least certain and based on ambiguous docking calculation results, is also the case where free energy calculations performed the worst in predicting ligand activity. Fortunately, the error estimates from FEP+ typically indicate when such convergence issues are present, which can be used to determine whether the calculation results should be considered when prioritizing experimental compounds to be made.

To have a significant impact on drug design projects, computational tools need to offer throughput that is complementary with experimental timelines. Any of the calculations presented in this work can be set up, conducted, and analyzed in the course of 24 h with moderate computational resources. For example, the calculations on Mcl-1 with 15 ligands could be completed in 24 h running on 30 graphics processing units (GPUs). Even for molecules that are accessible with established synthetic pathways, this makes FEP+ predictions an appealing option for compound prioritization. However, perhaps the greatest value in FEP+ calculations will come from cases that are synthetically highly challenging to make but from the computational perspective are easily accessible (e.g., core hops, multistep syntheses, etc.).

With most of the $\Delta\Delta G^0$ prediction errors below 1 kcal/mol for the cases presented here, FEP+ calculations offer a level of uncertainty approximately twice as high as that of experimental data (which is typically 0.3−0.5 kcal/mol[60]). As a rough guideline based on our results, this means that if a compound has a predicted $\Delta\Delta G^0$ of −1 kcal/mol compared to a known inhibitor, it can be estimated to be 64% likely to truly be a better binder (assuming an error distribution similar to Figure 1). Similarly, a compound that is predicted to be 2 kcal/mol better would have an 88% chance of being better. As such, if a primary objective of the calculations is to confidently predict compounds that will bind more tightly than a reference compound, then one should be able to continue running calculations until compounds are found with substantially better predicted binding energies. Thus, depending on project objectives, timelines, and risk profiles, the mode of deploying the FEP+ calculations can be tuned such that the expected probability of success for the energy directionality of perturbations (i.e., will molecule B be better than molecule A) will be known in advance. Understanding the prediction accuracy, as described here, will certainly prove valuable in picking new compounds for synthesis and testing.

In summary, we show that FEP+ calculations are highly predictive in calculating relative binding free energies for fragment-sized inhibitors, and this approach outperforms alternative affinity prediction methods (namely, docking and MM-GBSA). The FEP+ protocol described here has been designed so that it can easily be integrated into typical drug design workflows, and the throughput with a modest GPU cluster should be complementary with experimental timelines. Despite high computational costs as compared with most other binding energy prediction methods, FEP+ appears to offer a valuable approach for structurally enabled FBDD projects. Future work will go into expanding our validation work to explore protein−protein interactions (PPIs), allosteric modulators, macrocycles, covalent inhibitors, and other areas of pharmaceutical interest. Such projects are currently underway in our lab and will be the focus of future publications. In addition, we are exploring the sensitivity of the results to the input structure of the target and whether homology models could be used when crystal structures are not available.

Computational Procedures. All calculations have been conducted using the Schrödinger molecular modeling suite.[71] The FEP+ methodology combines an accurate modern force field (OPLS2.1),[65,72] efficient GPU-enabled parallel molecular dynamics with Desmond version 3.9,[73] REST enhanced sampling,[74] cycle-closure correction[57] to incorporate redundant

DOI: 10.1021/acs.jcim.5b00538
J. Chem. Inf. Model. 2015, 55, 2411−2420

information into free energy estimates, and the FEP Mapper (Figure 4) to automate setup and analysis of the calculations.

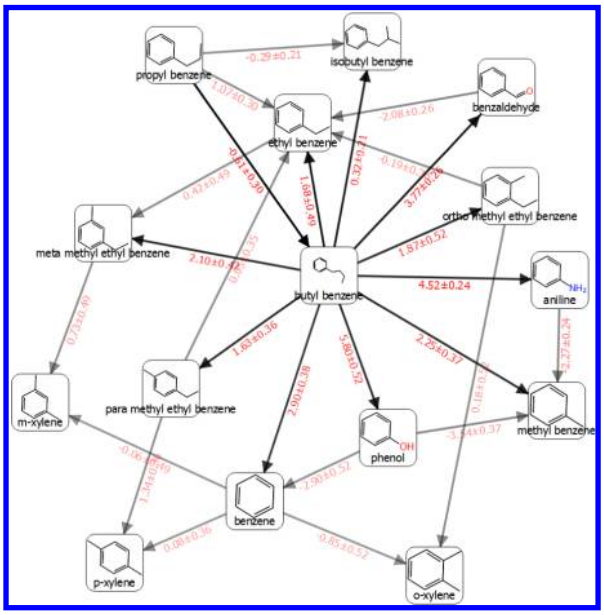


Figure 4. Typical calculation output representation for the Lysozyme system, as provided by the FEP+ Mapper. Each line represents one free energy calculation conducted both in solvated and receptor bound state resulting in the labeled $\Delta\Delta G^0$-value.

The force field builder tool was used to automatically derive accurate OPLS2.1 force field torsional parameters for all ligands containing substructures not fully covered by the standard parameters. Calculations were conducted using the default protocols unless indicated otherwise in the Results section. In this protocol, the systems were solvated in an orthogonal box of water molecules with buffer width of 5 Å for the complex and 10 Å for the solvent simulations. The full systems were relaxed and equilibrated using the default Desmond relaxation protocol, consisting of an energy-minimization with restraints on the solute, then 12 ps length simulations at 10 K using an NVT ensemble followed by an NPT ensemble. After that the restrained system was equilibrated at room temperature using the NPT ensemble. Finally a 240 ps room temperature NPT ensemble simulation was conducted. The NPT production stage lasted 5 ns for both the complex and the solvent simulations. A total of 12 λ windows were used for all the FEP/REST calculations. Replica exchanges between neighboring λ windows were attempted every 1.2 ps. For a more detailed description of the free energy calculation protocol employed, consult the Supporting Information of refs 57 and 18. All calculations were run on Nvidia GeForce GTX-780 GPUs.

All protein structures investigated have been downloaded from the Protein Data Bank (PDB)[75] and prepared with the Protein Preparation Wizard,[76] which adds missing atoms, assigns tautomer/ionization states, samples water orientations, flips Asn, Gln, and His residues to optimize the H-bond network, and performs energy minimization prior to conducting FEP+ calculations. All resolved crystal water molecules were maintained. Ligand 3D-structures where sketched by hand, energy-minimized, and placed in their corresponding receptor binding sites by ligand docking calculations using Glide version 6.4 with default parameters or were placed manually where a cocrystallized bound compound provided an unambiguous binding mode.

For comparison purposes, Glide ligand docking and Prime[77,78] MM-GBSA binding energy calculations were conducted. Both methods were used with their default parameters, using no constraints, standard precision (SP) docking in Glide with a binding site defined by the position of the cocrystallized ligand. For MM-GBSA calculations, the VSGB2 implicit solvent model[79] was used, and either the MM-GBSA default protocol, which minimizes the ligand only while keeping the protein structure rigid, or an alternative protocol with a 5 Å active region around the ligands for full molecular mechanics minimization (using the same active region for each system) was tested.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.5b00538.

Calculation details and results for all eight systems, together with experimental binding free energies and ligand structures (PDF)

Excel files of the raw calculation output, indicating absolute and relative binding free energies calculated for all compounds (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: thomas.steinbrecher@schrodinger.com.

Notes

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

■ ACKNOWLEDGMENTS

The authors express thanks to the Schrodinger support team for facilitating these calculations.

■ ABBREVIATIONS

FBDD, Fragment-based Drug Design; FEP, Free Energy Perturbation; GPU, Graphics Processing Unit; HTS, High Throughput Screening; ITC, Isothermal Titration Calorimetry; LDH, Human lactate dehydrogenase isoform A; Mcl-1, Induced myeloid leukemia cell differentiation protein; MD, Molecular Dynamics; MM-GBSA, Molecular Mechanics - Generalized Born, Surface Area; MUE, Mean Unsigned Error; MUP, Mouse urinary protein; PPI, Protein−protein interaction; QM/MM, Quantum Mechanical/Molecular Mechanical; RMSE, Root Mean Squared Error; SAR, Structure Activity Relationship; SPR, Surface Plasmon Resonance; TSA, Thermal Shift Assay

■ REFERENCES

(1) Loving, K.; Alberts, I.; Sherman, W. Computational approaches for fragment-based and de novo design. *Curr. Top. Med. Chem.* **2010**, *10*, 14−32.

(2) Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discovery* **2007**, *6*, 211−219.

(3) Erlanson, D. A.; McDowell, R. S.; O'Brien, T. Fragment-Based Drug Discovery. *J. Med. Chem.* **2004**, *47*, 3463−3482.

DOI: 10.1021/acs.jcim.5b00538
J. Chem. Inf. Model. 2015, 55, 2411−2420

(4) Carr, R. A. E.; Congreve, M.; Murray, C. W.; Rees, D. C. Fragment-based lead discovery: leads by design. *Drug Discovery Today* **2005**, *10*, 987−992.
(5) Murray, C. W.; Blundell, T. L. Structural biology in fragment-based drug design. *Curr. Opin. Struct. Biol.* **2010**, *20*, 497−507.
(6) Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C. Fragment-based approaches in drug discovery and chemical biology. *Biochemistry* **2012**, *51*, 4990−5003.
(7) Erlanson, D. Practical Fragments. http://practicalfragments.blogspot.de/2015/01/fragments-in-clinic-2015-edition.html (accessed 13.03.2015).
(8) Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A 'rule of three'for fragment-based lead discovery? *Drug Discovery Today* **2003**, *8*, 876−877.
(9) Lipinski, C. A. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today: Technol.* **2004**, *1*, 337−341.
(10) Ruddigkeit, L.; van Deursen, R.; Blum, L. C.; Reymond, J. L. Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17. *J. Chem. Inf. Model.* **2012**, *52*, 2864−75.
(11) Baker, M. Fragment-based lead discovery grows up. *Nat. Rev. Drug Discovery* **2013**, *12*, 5−7.
(12) Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, *28*, 849−857.
(13) Lauri, G.; Bartlett, P. CAVEAT: A program to facilitate the design of organic molecules. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 51−66.
(14) Miranker, A.; Karplus, M. Functionality maps of binding sites: a multiple copy simultaneous search method. *Proteins: Struct., Funct., Genet.* **1991**, *11*, 29−34.
(15) Eisen, M. B.; Wiley, D. C.; Karplus, M.; Hubbard, R. E. HOOK: a program for finding novel molecular architectures that satisfy the chemical and steric requirements of a macromolecule binding site. *Proteins: Struct., Funct., Genet.* **1994**, *19*, 199−221.
(16) Murray, C. W.; Erlanson, D. A.; Hopkins, A. L.; Keserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H.; Richmond, N. J. Validity of Ligand Efficiency Metrics. *ACS Med. Chem. Lett.* **2014**, *5*, 616−618.
(17) Shultz, M. D. Improving the Plausibility of Success with Inefficient Metrics. *ACS Med. Chem. Lett.* **2014**, *5*, 2−5.
(18) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by way of a Modern Free Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* **2015**, *137*, 2695−2703.
(19) Eriksson, A. E.; Baase, W. A.; Wozniak, J. A.; Matthews, B. W. A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene. *Nature* **1992**, *355*, 371−373.
(20) Graves, A. P.; Brenk, R.; Shoichet, B. K. Decoys for docking. *J. Med. Chem.* **2005**, *48*, 3714−3728.
(21) Morton, A.; Baase, W. A.; Matthews, B. W. Energetic origins of specificity of ligand binding in an interior nonpolar cavity of T4 lysozyme. *Biochemistry* **1995**, *34*, 8564−8575.
(22) Morton, A.; Matthews, B. W. Specificity of ligand binding in a buried nonpolar cavity of T4 lysozyme: linkage of dynamics and structural plasticity. *Biochemistry* **1995**, *34*, 8576−8588.
(23) Mobley, D. L.; Dill, K. A. Binding of Small-Molecule Ligands to Proteins: "What You See" Is Not Always "What You Get. *Structure* **2009**, *17*, 489−498.
(24) Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. Predicting Absolute Ligand Binding Free Energies to a Simple Model Site. *J. Mol. Biol.* **2007**, *371*, 1118−1134.
(25) Mobley, D. L.; Chodera, J. D.; Dill, K. A. On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J. Chem. Phys.* **2006**, *125*, 084902.
(26) Graves, A. P.; Shivakumar, D. M.; Boyce, S. E.; Jacobson, M. P.; Case, D. A.; Shoichet, B. K. Rescoring Docking Hit Lists for Model Cavity Sites: Predictions and Experimental Testing. *J. Mol. Biol.* **2008**, *377*, 914−934.
(27) Shuman, S. DNA ligases: progress and prospects. *J. Biol. Chem.* **2009**, *284*, 17365−17369.
(28) Howard, S.; Amin, N.; Benowitz, A. B.; Chiarparin, E.; Cui, H.; Deng, X.; Heightman, T. D.; Holmes, D. J.; Hopkins, A.; Huang, J.; Jin, Q.; Kreatsoulas, C.; Martin, A. C.; Massey, F.; McCloskey, L.; Mortenson, P. N.; Pathuri, P.; Tisi, D.; Williams, P. A. Fragment-based discovery of 6-azaindazoles as inhibitors of bacterial DNA ligase. *ACS Med. Chem. Lett.* **2013**, *4*, 1208−1212.
(29) Wei, G.; Margolin, A. A.; Haery, L.; Brown, E.; Cucolo, L.; Julian, B.; Shehata, S.; Kung, A. L.; Beroukhim, R.; Golub, T. R. Chemical Genomics Identifies Small-Molecule MCL1 Repressors and BCL-xL as a Predictor of MCL1 Dependency. *Cancer Cell* **2012**, *21*, 547−562.
(30) Beroukhim, R.; Mermel, C. H.; Porter, D.; Wei, G.; Raychaudhuri, S.; Donovan, J.; Barretina, J.; Boehm, J. S.; Dobson, J.; Urashima, M. The landscape of somatic copy-number alteration across human cancers. *Nature* **2010**, *463*, 899−905.
(31) Danial, N. N.; Korsmeyer, S. J. Cell death: critical control points. *Cell* **2004**, *116*, 205−219.
(32) Willis, S. N.; Fletcher, J. I.; Kaufmann, T.; van Delft, M. F.; Chen, L.; Czabotar, P. E.; Ierino, H.; Lee, E. F.; Fairlie, W. D.; Bouillet, P. Apoptosis initiated when BH3 ligands engage multiple Bcl-2 homologs, not Bax or Bak. *Science* **2007**, *315*, 856−859.
(33) Adams, J.; Cory, S. The Bcl-2 apoptotic switch in cancer development and therapy. *Oncogene* **2007**, *26*, 1324−1337.
(34) Thomas, L. W.; Lam, C.; Edwards, S. W. Mcl-1; the molecular regulation of protein function. *FEBS Lett.* **2010**, *584*, 2981−2989.
(35) Al-Shawi, R.; Ghazal, P.; Clark, A. J.; Bishop, J. O. Intraspecific evolution of a gene family coding for urinary proteins. *J. Mol. Evol.* **1989**, *29*, 302−313.
(36) Finlayson, J.; Potter, M.; Runner, C. R. Electrophoretic variation and sex dimorphism of the major urinary protein complex in inbred mice: a new genetic marker. *J. Natl. Cancer Inst.* **1963**, *31*, 91−107.
(37) Robertson, D.; Hurst, J.; Hubbard, S.; Gaskell, S. J.; Beynon, R. Ligands of urinary lipocalins from the mouse: uptake of environmentally derived chemicals. *J. Chem. Ecol.* **1998**, *24*, 1127−1140.
(38) Sharrow, S. D.; Vaughn, J. L.; Žídek, L.; Novotny, M. V.; Stone, M. J. Pheromone binding by polymorphic mouse major urinary proteins. *Protein Sci.* **2002**, *11*, 2247−2256.
(39) Homans, S. W. Water, water everywhere−except where it matters? *Drug Discovery Today* **2007**, *12*, 534−539.
(40) Barratt, E.; Bronowska, A.; Vondrasek, J.; Cerny, J.; Bingham, R.; Phillips, S.; Homans, S. W. Thermodynamic penalty arising from burial of a ligand polar group within a hydrophobic pocket of a protein receptor. *J. Mol. Biol.* **2006**, *362*, 994−1003.
(41) Malham, R.; Johnstone, S.; Bingham, R. J.; Barratt, E.; Phillips, S. E.; Laughton, C. A.; Homans, S. W. Strong solute-solute dispersive interactions in a protein-ligand complex. *J. Am. Chem. Soc.* **2005**, *127*, 17061−17067.
(42) Barratt, E.; Bingham, R. J.; Warner, D. J.; Laughton, C. A.; Phillips, S. E.; Homans, S. W. Van der Waals interactions dominate ligand-protein association in a protein binding site occluded from solvent water. *J. Am. Chem. Soc.* **2005**, *127*, 11827−11834.
(43) Bingham, R. J.; Findlay, J. B.; Hsieh, S. Y.; Kalverda, A. P.; Kjellberg, A.; Perazzolo, C.; Phillips, S. E.; Seshadri, K.; Trinh, C. H.; Turnbull, W. B.; Bodenhausen, G.; Homans, S. W. Thermodynamics of binding of 2-methoxy-3-isopropylpyrazine and 2-methoxy-3-isobutylpyrazine to the major urinary protein. *J. Am. Chem. Soc.* **2004**, *126*, 1675−1681.
(44) Warburg, O. On the origin of cancer cells. *Science* **1956**, *123*, 309−314.
(45) Granchi, C.; Roy, S.; Giacomelli, C.; Macchia, M.; Tuccinardi, T.; Martinelli, A.; Lanza, M.; Betti, L.; Giannaccini, G.; Lucacchini, A.; Funel, N.; Leon, L. G.; Giovannetti, E.; Peters, G. J.; Palchaudhuri, R.; Calvaresi, E. C.; Hergenrother, P. J.; Minutolo, F. Discovery of N-

DOI: 10.1021/acs.jcim.5b00538
J. Chem. Inf. Model. 2015, 55, 2411−2420

hydroxyindole-based inhibitors of human lactate dehydrogenase isoform A (LDH-A) as starvation agents against cancer cells. *J. Med. Chem.* **2011**, *54*, 1599−612.
(46) Leonard, W. J.; O'Shea, J. J. Jaks and STATs: biological implications*. *Annu. Rev. Immunol.* **1998**, *16*, 293−322.
(47) Delhommeau, F.; Pisani, D. F.; James, C.; Casadevall, N.; Constantinescu, S.; Vainchenker, W. Oncogenic mechanisms in myeloproliferative disorders. *Cell. Mol. Life Sci.* **2006**, *63*, 2939−2953.
(48) Kamal, A.; Thao, L.; Sensintaffar, J.; Zhang, L.; Boehm, M. F.; Fritz, L. C.; Burrows, F. J. A high-affinity conformation of Hsp90 confers tumour selectivity on Hsp90 inhibitors. *Nature* **2003**, *425*, 407−410.
(49) Han, J.; Lee, J.; Bibbs, L.; Ulevitch, R. A MAP kinase targeted by endotoxin and hyperosmolarity in mammalian cells. *Science* **1994**, *265*, 808−811.
(50) Lee, J. C.; Laydon, J. T.; McDonnell, P. C.; Gallagher, T. F.; Kumar, S.; Green, D.; McNulty, D.; Blumenthal, M. J.; Keys, J. R.; Land vatter, S. W.; Strickler, J. E.; McLaughlin, M. M.; Siemens, I. R.; Fisher, S. M.; Livi, G. P.; White, J. R.; Adams, J. L.; Young, P. R. A protein kinase involved in the regulation of inflammatory cytokine biosynthesis. *Nature* **1994**, *372*, 739−746.
(51) Friberg, A.; Vigil, D.; Zhao, B.; Daniels, R. N.; Burke, J. P.; Garcia-Barrantes, P. M.; Camper, D.; Chauder, B. A.; Lee, T.; Olejniczak, E. T.; Fesik, S. W. Discovery of potent myeloid cell leukemia 1 (Mcl-1) inhibitors using fragment-based methods and structure-based design. *J. Med. Chem.* **2013**, *56*, 15−30.
(52) Timm, D. E.; Baker, L. J.; Mueller, H.; Zidek, L.; Novotny, M. V. Structural basis of pheromone binding to mouse major urinary protein (MUP-I). *Protein Sci.* **2001**, *10*, 997−1004.
(53) Read, J. A.; Winter, V. J.; Eszes, C. M.; Sessions, R. B.; Brady, R. L. Structural basis for altered activity of M- and H-isozyme forms of human lactate dehydrogenase. *Proteins: Struct., Funct., Genet.* **2001**, *43*, 175−85.
(54) Antonysamy, S.; Hirst, G.; Park, F.; Sprengeler, P.; Stappenbeck, F.; Steensma, R.; Wilson, M.; Wong, M. Fragment-based discovery of JAK-2 inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 279−282.
(55) Barker, J. J.; Barker, O.; Boggio, R.; Chauhan, V.; Cheng, R. K. Y.; Corden, V.; Courtney, S. M.; Edwards, N.; Falque, V. M.; Fusar, F.; Gardiner, M.; Hamelin, E. M. N.; Hesterkamp, T.; Ichihara, O.; Jones, R. S.; Mather, O.; Mercurio, C.; Minucci, S.; Montalbetti, C. A. G. N.; Müller, A.; Patel, D.; Phillips, B. G.; Varasi, M.; Whittaker, M.; Winkler, D.; Yarnold, C. J. Fragment-based Identification of Hsp90 Inhibitors. *ChemMedChem* **2009**, *4*, 963−966.
(56) Hartshorn, M. J.; Murray, C. W.; Cleasby, A.; Frederickson, M.; Tickle, I. J.; Jhoti, H. Fragment-Based Lead Discovery Using X-ray Crystallography. *J. Med. Chem.* **2005**, *48*, 403−413.
(57) Wang, L.; Deng, Y.; Knight, J. L.; Wu, Y.; Kim, B.; Sherman, W.; Shelley, J. C.; Lin, T.; Abel, R. Modeling Local Structural Rearrangements Using FEP/REST: Application to Relative Binding Affinity Predictions of CDK2 Inhibitors. *J. Chem. Theory Comput.* **2013**, *9*, 1282−1293.
(58) Wang, L.; Berne, B. J.; Friesner, R. A. On achieving high accuracy and reliability in the calculation of relative protein−ligand binding affinities. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 1937−1942.
(59) Christ, C. D.; Fox, T. Accuracy Assessment and Automation of Free Energy Calculations for Drug Design. *J. Chem. Inf. Model.* **2014**, *54*, 108−120.
(60) Brown, S. P.; Muchmore, S. W.; Hajduk, P. J. Healthy skepticism: assessing realistic model performance. *Drug Discovery Today* **2009**, *14*, 420−427.
(61) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177−6196.
(62) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750−1759.
(63) Varasi, M. *personal communication*, 2015.
(64) Murphy-Benenato, K.; Wang, H.; McGuire, H. M.; Davis, H. E.; Gao, N.; Prince, D. B.; Jahic, H.; Stokes, S. S.; Boriack-Sjodin, P. A. Identification through structure-based methods of a bacterial NAD +-dependent DNA ligase inhibitor that avoids known resistance mutations. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 360−366.
(65) Shivakumar, D.; Harder, E.; Damm, W.; Friesner, R. A.; Sherman, W. Improving the prediction of absolute solvation free energies using the next generation OPLS force field. *J. Chem. Theory Comput.* **2012**, *8*, 2553−2558.
(66) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. Alchemical free energy methods for drug discovery: progress and challenges. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150−160.
(67) Chipot, C.; Pohorille, A. *Free energy calculations: theory and applications in chemistry and biology*; Springer: 2007; Vol. *86*.
(68) Jorgensen, W. L. The many roles of computation in drug discovery. *Science* **2004**, *303*, 1813−1818.
(69) Jorgensen, W. L. Efficient drug lead discovery and optimization. *Acc. Chem. Res.* **2009**, *42*, 724−733.
(70) Steinbrecher, T.; Labahn, A. Towards accurate free energy calculations in ligand protein-binding studies. *Curr. Med. Chem.* **2010**, *17*, 767−85.
(71) *Schrödinger Suite*, 2014-3; Schrödinger, LLC: New York, NY, 2014.
(72) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *J. Chem. Theory Comput.* **2010**, *6*, 1509−1519.
(73) 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. In *Scalable algorithms for molecular dynamics simulations on commodity clusters*, SC 2006 Conference, Proceedings of the ACM/IEEE, IEEE: 2006; 43−43.
(74) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. Replica exchange with solute tempering: A method for sampling biological systems in explicit water. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 13749−13754.
(75) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(76) Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221−234.
(77) Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B. On the role of the crystal environment in determining protein side-chain conformations. *J. Mol. Biol.* **2002**, *320*, 597−608.
(78) Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J.; Honig, B.; Shaw, D. E.; Friesner, R. A. A hierarchical approach to all-atom protein loop prediction. *Proteins: Struct., Funct., Genet.* **2004**, *55*, 351−367.
(79) Li, J.; Abel, R.; Zhu, K.; Cao, Y.; Zhao, S.; Friesner, R. A. The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins: Struct., Funct., Genet.* **2011**, *79*, 2794−2812.

DOI: 10.1021/acs.jcim.5b00538
J. Chem. Inf. Model. 2015, 55, 2411−2420