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

Supporting Information

Authors: Lingle Wang,[1] Yujie Wu,[1] Yuqing Deng,[1] Byungchan Kim,[1] Levi Pierce,[1] Goran Krilov,[1] Dmitry Lupyan,[1] Shaughnessy Robinson,[1] Markus K. Dahlgren,[1] Jeremy Greenwood,[1] Donna L. Romero,[2] Craig Masse,[2] Jennifer L. Knight,[1] Thomas Steinbrecher,[1] Thijs Beuming,[1] Wolfgang Damm,[1] Ed Harder,[1] Woody Sherman,[1] Mark Brewer,[1] Ron Wester,[2] Mark Murcko,[1] Leah Frye,[1] Ramy Farid,[1] Teng Lin,[1] David L. Mobley,[5] William L. Jorgensen,[4] Bruce J. Berne,[3] Richard A. Friesner,[3] Robert Abel[1,*]

[1] Schrödinger Inc, 120 West 45th Street, New York, NY 10036, United States

[2] Nimbus Disocvery, 25 First Street, Suite 404 Cambridge, MA 02141 United States

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

[4] Department of Chemistry, Yale University, New Haven, CT 06520, United States

[5] Departments of Pharmaceutical Sciences and Chemistry, University of California—Irvine, Irvine, CA 92697, United States

[*] Corresponding author email: robert.abel@schrodinger.com

Methods

Details about the REST region selection algorithm

A crucial aspect of the use of FEP/REST in practical calculations is the selection of the REST region. Heuristic rules must be developed with regard to which atoms in the ligand and the protein environment should be included in REST enhanced sampling region. We have developed an automated algorithm to select the REST region, and this algorithm was employed in a uniform fashion in all the studies reported in the article. The details about the REST region selection algorithm are described in the following paragraph.

In FEP simulations for a certain receptor, if there are multiple crystal structures available, comparisons between the structures are made, and the protein residues close to the binding pocking having significantly different conformations in the different crystal structures are included in the REST region. For the retrospective validation set, no protein residues were included in the REST region. For the prospective set Project II, a uniform set of key mobile protein residues inferred from extensive in-house crystallography were added to the REST region. For the prospective Project I, two protein residues close to the binding pocket forming a salt bridge were included in the REST region for mutations perturbing that region of the binding pocket. In all systems, for a mutation from ligand A to ligand B, the region of the ligand involved directly in the perturbation is included in the REST region. If the

number of heavy atoms of the perturbed functional group is less than the cutoff value (25 heavy atoms), then the algorithm tries to include more ligand atoms in the REST region so that the REST region would include one more rotatable bond until the total number of heavy atoms exceeds the cutoff value.

Details of the simulations

The FEP mapper module implemented in Desmond was used to set up the calculations. The starting structures for the simulations were taken from the PDB structures with IDs 4DJW for BACE, 1H1Q for CDK2, 2GMX for JNK1, 4HW3 for MCL1, 3FLY for P38, 2QBS for PTP1B, 2ZFF for Thrombin and 4GIH for Tyk2. The OPLS2.1 force field was used for the proteins and the ligands along with the SPC water model.[1,2] The torsional parameters for those torsions that are not included in the default OPLS2.1 force field were generated using the Force Field builder.

The proteins were prepared using the Protein Preparation Wizard[3] during which the protonation states were assigned assuming a pH of 7.0. The crystallogaphic waters in the binding pocket of the BACE receptor were included in the protein for further FEP simulations, which was found to be important for accurate modeling of the binding affinities of the ligands. Cyclin A was retained with CDK2 and treated as part of the receptor for all calculations. The systems were solvated in a water box with buffer width of 5 Å for the complex simulations and 10 Å for the solvent simulations. The systems were relaxed and equilibrated using the default Desmond

relaxation protocol. The whole system with the solute molecules restrained to their initial positions was first minimized using the Brownie integrator and then simulated at 10 K using an NVT ensemble followed by an NPT ensemble. After that the system was simulated at room temperature using the NPT ensemble with the restraints retained. Then the whole system without any restraint was simulated at room temperature using the NPT ensemble for 240 ps followed by the production simulation. A total of 12 λ windows were used for all the FEP/REST calculations. The production stage lasted 5 ns for both the complex and the solvent simulations using NPT ensemble conditions. Replica exchanges between neighboring λ windows were attempted every 1.2 ps. The Bennett acceptance ratio method (BAR)[4] was used to calculate the free energy. Errors were estimated for each free energy calculation using both bootstrapping and the BAR analytical error prediction.[5] After the simulations were complete, the hysteresis along closed thermodynamic cycles were calculated and best estimates of the free energies and the errors for the predictions were calculated using the cycle closure algorithm reported in a prior publication.[6]

Details about the OPLS 2.1 force field

An analogous figure to Figure 1 in the main text illustrating conformational energy errors associated with the force field is provided here as well (Figure S2) with an additional comparison to the previous iteration of the model (OPLS2.0).   Quantum mechanical energies are calculated using the program Jaguar[7] while force field energies are based on the molecular mechanics program MacroModel.[8] Key

improvements made to the force field between OPLS2.0 and OPLS2.1 include the following:

- Approximately 1200 new torsion profiles were added to the OPLS2 training set and approximately 10000 new torsion parameters were introduced to the model.

- Utilizing a set of approximately 8000 geometry-optimized conjugated and saturated rings, an additional 7000 new bend types were introduced into the model.

- To address excessive strain seen in biaryl conformations[9], bend force constants assigned to biaryl moieties were optimized based on a normal mode frequency comparison to QM.[10,11]

Before running the FEP calculations missing torsions were identified and fit using an automated workflow. An outline of the workflows basic steps is as follows:

- A model fragment compound that retains the original torsion substructure is generated for each missing torsion.

- Multiple force field based candidate torsion scans are developed and one is selected using criteria that rewards low conformational energies and penalizes non-bonded variability along the developed coordinates.

- B3LYP/6-31G* geometry optimizations are run subject to a constraint applied to the targeted missing torsion.

- The QM energy surface is resolved based on LMP2/cc-pvtz(-f) single point energy calculations calculated on the B3LYP optimized structures.

- Torsion restrained force field based minimizations followed by fitting against the QM energy surface is used to derive new parameters.

- The OPLS2.1 force field is augmented, including the newly developed parameters.

Over the full set of 199 ligands included in this study 89 included at least one missing torsion parameter.   A comparison is provided in Table S3 contrasting the fraction of missing torsions found in the set with the fraction identified in a set of approved drugs and that measured for a large database of drug-like compounds. The average root mean square error after fitting these missing torsions is 0.6 kcal/mol, consistent with the reported OPLS2.1 error over well-covered fragment molecules shown in Figure 1 and Figure S2. The average OPLS2.1 root mean square error on those same torsions before fitting is 1.6 kcal/mol, consistent with the errors seen for OPLS2005 and MMFF.

How to convert the calculated ΔΔG values to ΔG values

Through FEP/REST simulations, we directly compute the relative binding free energies between pairs of ligands directly connected in the mutation graph. For mutations which form a closed thermodynamic cycle, for example, A→B, B→C, C→A, the sum of the calculated free energies from FEP/REST along the cycle might in general deviate from the theoretical value of 0, which is denoted as the hysteresis of the cycle. Using the cycle closure correction algorithm developed in a prior publication,[6] the free energies after correction do not have hysteresis.

To convert the relative binding free energies to absolute binding free energies, we need to know the absolute binding free energy for at least one of the molecules of the series. In retrospective studies, in principle, we can choose any molecule as the reference compound, and force the predicted ΔG value of the reference compound to be equal to the experimental value ΔG_{exp}. However, if the experimental measurement for the reference compound has a large error, the error will propagate to the estimated binding free energies for all the other molecules using this method. Thus, to minimize the error introduced in choosing an arbitrary compound as the reference compound, we instead choose all the compounds with experimental binding free energies available as the reference. To clarify, if there are N molecules with experimental binding free energies available, we use the following equation to calculate the offset to convert the relative binding free energies to absolute free energies:

$$\sum_{i=1}^{N} \Delta G_{exp}^{i} = \sum_{i=1}^{N} \Delta G_{pred}^{i} \tag{1.1}$$

In retrospective studies, all the molecules are effectively used as the reference to convert relative binding free energies into absolute free energies. In prospective studies, if multiple molecules are available with experimental binding free energy data, all such available molecules are used as the reference to calculate the absolute binding free energy for the idea molecule.

Expected FEP prediction accuracy in prospective studies

It appears that the distribution of errors between the FEP predicted binding affinities versus the experimental data for the 337 perturbations is roughly a Gaussian distribution with a sigma parameter of ca. 1.14 kcal/mol, albeit with a fat tail beyond 2.2 kcal/mol. (Table 3 in main text) One can infer from this distribution that the expected fraction of affinities correctly computed to within some threshold X, will approximately be erf(X/(1.1*sqrt(2))), as per Gaussian statistics. Likewise when comparing the relative binding affinities of two molecules, the probability that two molecules are correctly rank-ordered with respect to each other becomes an approximate function of the computed free energy difference between them—i.e., a 1.0 kcal/mol ΔΔG corresponds to 82% probability of correct rank ordering between the pair, a 1.5 kcal/mol ΔΔG corresponds to 91% probability of correct rank ordering between the pair, and in the general case a ΔΔG of X kcal/mol corresponds to 0.5*(1+erf(|X|/(1.1*sqrt(2)))) approximate probability of correct rank ordering between the pair. Knowledge of the error distribution and robust ligand rank ordering facilitates confident use of FEP predictions in drug discovery in a way that has been challenging with other physics-based affinity calculation techniques.

Expected correlation coefficient between FEP predicted binding affinities and experimental values, and between two independent experimental measurements

As pointed out by Brown et al. in a prior publication,[12] the errors in the experimental binding affinity assay set a limit on how well a model agrees with experimental results. [12] The correlation coefficient of the binding free energies for a set of compounds from two independent measurements is generally smaller than 1, depending on the errors associated with the measurements, the number of data points, and the range of the binding free energies.

In this paper, we use the following model to estimate the expected correlation coefficient of the binding free energies between two independent experimental measurements. The error for careful binding free energy measurements is about 0.3-0.5 kcal/mol according to the analysis of Brown et al.[12] Therefore, we assume the error of the experimental binding free energy is Gaussian distributed with standard deviation of 0.4 kcal/mol. For a given set of N compounds with binding free energies of ΔG^i (i=1,2,...N), the experimentally measured binding free energy for the i-th compound has the following distribution:

$$P(\Delta G_{exp}^i) = \frac{1}{\sqrt{2\pi}\sigma_{exp}} \exp\left(-\frac{(\Delta G_{exp}^i - \Delta G^i)^2}{2\sigma_{exp}^2}\right) \tag{1.2}$$

For each set of ligands studied in the present paper, we generated 1,000 sets of

binding free energies according to the above Gaussian distribution, and the

correlation coefficient between all pairs of sets of free energies are calculated. The

average and the standard deviation of the resulting correlation coefficients are

reported as the expected correlation coefficient between two independent

experimental measurements and their uncertainty.

Similarly, the FEP predicted binding free energies also have errors associated with

the predictions. Assuming the error for the FEP predicted binding free energies is

also Gaussian distributed with an average of 0 (no systematic error) and a standard

deviation of 1.1 kcal/mol, (i.e., the RMSE of the FEP predicted binding free energies

compared to experimental results). Then, the FEP predicted binding free energy for

compound i with binding free energy of ΔG^i has the following distribution:

$$P(\Delta G_{FEP}^i) = \frac{1}{\sqrt{2\pi}\sigma_{FEP}}\exp(-\frac{(\Delta G_{FEP}^i - \Delta G^i)^2}{2\sigma_{FEP}^2}) \qquad \textbf{(1.3)}$$

For each set of ligands studied in the present paper, 1,000 sets of binding free

energies were generated according to equation 1.2, and another 1,000 sets of

binding free energies were generated according to equation 1.3. The average

correlation coefficient and the standard deviation are reported as the expected

correlation coefficient between all pairs of these sets of FEP predictions and

experimental results and their uncertainty.

Details about the experimental binding affinity data

CDK2 inhibitors: The same set of 16 ligands that was used in a prior publication[6] was used here. The experimental binding free energies of the ligands were calculated from their reported IC50 values[13] using the following approximation:

$$\Delta G_{exp} \approx RT \ln(IC_{50}) \qquad\qquad (1.4)$$

where T=297 K, and R is the gas constant.

P38 inhibitors: All 34 ligands with experimental IC50 values reported in a prior publication[14] were included in the present study. The experimental binding free energies were calculated using the same formula 1.4.

Jnk1 inhibitors: All 21 ligands with a benzene core attached to the amide group reported in a prior publication[15] were included in the present study. The experimental binding free energies for the ligands were calculated from their reported IC50 values using formula 1.4.

BACE inhibitors: All the ligands reported in the prior publication[16] except for ligands 4e, 4f, 4g, 4h were included in the present study. For ligands 4e, 4f, 4g, 4h, due to the symmetry of the shape of the molecules, there might be two possible binding poses for these ligands, and they were not included in the FEP calculation. The

experimental binding free energies for the ligands were calculated using the following formula:

$$\Delta G_{\text{exp}} = RT \ln K_i \qquad (1.5)$$

For ligands where both the R and S enantiomers were used in the experimental binding assay, only one of the enantiomers is assumed to be bioactive, and the K_i for those ligands were calculated as one half the value reported in the paper.

Tyk2 inhibitors: 12 representative ligands in Table 4 of a prior publication[17] and 4 structurally similar ligands in another publication[18] from the same group were included in the present study. The experimental binding free energies for the ligands were calculated from their reported Ki values using formula 1.5.

MCL1 inhibitors: All 42 ligands in Table 3 of a prior publication[19] except for ligand 59 which can not be docked to the protein structure were included in the present study. The experimental binding free energies for these ligands were calculated from their reported Ki values using formula 1.5.

PTP1B inhibitors: 23 representative congeneric ligands from Table 1-3 in a prior publication[20] were included in the present study. The experimental binding free energies for these ligands were calculated from the reported Ki values using formula 1.5.

Thrombin inhibitors: All 11 ligands in Table 1 of a prior publication[21] except for ligand 4 which has a different net charge were included in the present study. The experimental binding free energies for these ligands were taken from their reported ITC data.

Input structures for the FEP calculations

The input structures for each series of the ligands and the corresponding crystal structure used in the FEP calculation are available for download.

https://drive.google.com/folderview?id=0BylmDElgu6QLTnJ2WGMzNXBENkk&usp=sharing

(Inputs_for_FEP.tgz)

The pkl file containing the perturbation graph for each series of ligands are also included.

Raw data for all the perturbations and the final predicted binding free energies for all the ligands are given in the attached excel file for download.

(Retrospective_FEP_results_summary)

This information is available free of charge via the internet at http://pubs.acs.org.

References:

(1)   Shivakumar, D.; Harder, E.; Damm, W.; Friesner, R. A.; Sherman, W. *J. Chem. Theory Comput.* **2012**, *8*, 2553.

(2)   Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; J., H.; Pullman, B., Ed.; Reidel: Dordrecht, 1981, p 331.

(3)   In *Schrodinger Suite 2012 Protein Preparation Wizard*; Schrodinger L. L. C.: New York, NY, 2012.

(4)   Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245.

(5)   Hahn, A. M.; Then, H. *Phys. Rev. E* **2009**, *80*, 031111.

(6)   Wang, L.; Deng, Y.; Knight, J. L.; Wu, Y.; Kim, B.; Sherman, W.; Shelley, J. C.; Lin, T.; Abel, R. *J. Chem. Theory Comput.* **2013**, *9*, 1282.

(7)   In *Jaguar, Schrodinger, Inc. version 5.5*; Schrodinger, Inc.: New York, NY, 2014.

(8)   In *Macromodel, Schrodinger, Inc. version 6.5*; Schrodinger, Inc.: New York, NY, 2014.

(9)   DuBay, K. H.; Hall, M. L.; Hughes, T. F.; Wu, C.; Reichman, D. R.; Friesner, R. A. *Journal of Chemical Theory and Computation* **2012**, *8*, 4556.

(10)   Bochevarov, A. D.; Harder, E.; Hughes, T. F.; Greenwood, J. R.; Braden, D. A.; Philipp, D. M.; Rinaldo, D.; Halls, M. D.; Zhang, J.; Friesner, R. A. *International Journal of Quantum Chemistry* **2013**, *113*, 2110.

(11)   Murphy, R. B.; Beachy, M. D.; Friesner, R. A.; Ringnalda, M. N. *The Journal of Chemical Physics* **1995**, *103*, 1481.

(12)   Brown, S. P.; Muchmore, S. W.; Hajduk, P. J. *Drug Discovery Today* **2009**, *14*, 420.

(13)   Hardcastle, I. R.; Arris, C. E.; Bentley, J.; Boyle, F. T.; Chen, Y.; Curtin, N. J.; Endicott, J. A.; Gibson, A. E.; Golding, B. T.; Griffin, R. J.; Jewsbury, P.; Menyerol, J.; Mesguiche, V.; Newell, D. R.; Noble, M. E.; Pratt, D. J.; Wang, L. Z.; Whitfield, H. J. *J. Med. Chem.* **2004**, *47*, 3710.

(14)   Goldstein, D. M.; Soth, M.; Gabriel, T.; Dewdney, N.; Kuglstatter, A.; Arzeno, H.; Chen, J.; Bingenheimer, W.; Dalrymple, S. A.; Dunn, J.; Farrell, R.; Frauchiger, S.; La Fargue, J.; Ghate, M.; Graves, B.; Hill, R. J.; Li, F.; Litman, R.; Loe, B.; McIntosh, J.; McWeeney, D.; Papp, E.; Park, J.; Reese, H. F.; Roberts, R. T.; Rotstein, D.; San Pablo, B.; Sarma, K.; Stahl, M.; Sung, M.-L.; Suttman, R. T.; Sjogren, E. B.; Tan, Y.; Trejo, A.; Welch, M.; Weller, P.; Wong, B. R.; Zecic, H. *J. Med. Chem.* **2011**, *54*, 2255.

(15)   Szczepankiewicz, B. G.; Kosogof, C.; Nelson, L. T. J.; Liu, G.; Liu, B.; Zhao, H.; Serby, M. D.; Xin, Z.; Liu, M.; Gum, R. J.; Haasch, D. L.; Wang, S.; Clampit, J. E.; Johnson, E. F.; Lubben, T. H.; Stashko, M. A.; Olejniczak, E. T.; Sun, C.; Dorwin, S. A.; Haskins, K.; Abad-Zapatero, C.; Fry, E. H.; Hutchins, C. W.; Sham, H. L.; Rondinone, C. M.; Trevillyan, J. M. *J. Med. Chem.* **2006**, *49*, 3563.

(16)   Cumming, J. N.; Smith, E. M.; Wang, L.; Misiaszek, J.; Durkin, J.; Pan, J.; Iserloh, U.; Wu, Y.; Zhu, Z.; Strickland, C.; Voigt, J.; Chen, X.; Kennedy, M. E.; Kuvelkar, R.; Hyde, L. A.; Cox, K.; Favreau, L.; Czarniecki, M. F.; Greenlee, W. J.; McKittrick, B. A.; Parker, E. M.; Stamford, A. W. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2444.

(17)   Liang, J.; Tsui, V.; Van Abbema, A.; Bao, L.; Barrett, K.; Beresini, M.; Berezhkovskiy, L.; Blair, W. S.; Chang, C.; Driscoll, J.; Eigenbrot, C.; Ghilardi, N.;

Gibbons, P.; Halladay, J.; Johnson, A.; Kohli, P. B.; Lai, Y.; Liimatta, M.; Mantik, P.; Menghrajani, K.; Murray, J.; Sambrone, A.; Xiao, Y.; Shia, S.; Shin, Y.; Smith, J.; Sohn, S.; Stanley, M.; Ultsch, M.; Zhang, B.; Wu, L. C.; Magnuson, S. *Eur. J. Med. Chem.* **2013**, *67*, 175.

(18)   Liang, J.; van Abbema, A.; Balazs, M.; Barrett, K.; Berezhkovsky, L.; Blair, W.; Chang, C.; Delarosa, D.; DeVoss, J.; Driscoll, J.; Eigenbrot, C.; Ghilardi, N.; Gibbons, P.; Halladay, J.; Johnson, A.; Kohli, P. B.; Lai, Y.; Liu, Y.; Lyssikatos, J.; Mantik, P.; Menghrajani, K.; Murray, J.; Peng, I.; Sambrone, A.; Shia, S.; Shin, Y.; Smith, J.; Sohn, S.; Tsui, V.; Ultsch, M.; Wu, L. C.; Xiao, Y.; Yang, W.; Young, J.; Zhang, B.; Zhu, B.-y.; Magnuson, S. *J. Med. Chem.* **2013**, *56*, 4521.

(19)   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. *J. Med. Chem.* **2012**, *56*, 15.

(20)   Wilson, D. P.; Wan, Z.-K.; Xu, W.-X.; Kirincich, S. J.; Follows, B. C.; Joseph-Mccarthy, D.; Foreman, K.; Moretto, A.; Wu, J.; Zhu, M.; Binnun, E.; Zhang, Y.-L.; Tam, M.; Erbe, D. V.; Tobin, J.; Xu, X.; Leung, L.; Shilling, A.; Tam, S. Y.; Mansour, T. S.; Lee, J. *J. Med. Chem.* **2007**, *50*, 4681.

(21)   Baum, B.; Mohamed, M.; Zayed, M.; Gerlach, C.; Heine, A.; Hangauer, D.; Klebe, G. *J. Mol. Biol.* **2009**, *390*, 56.

Tables and Figures

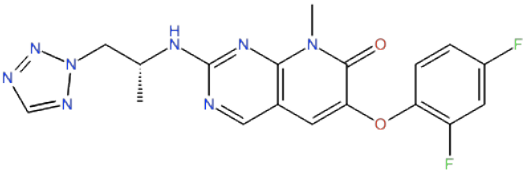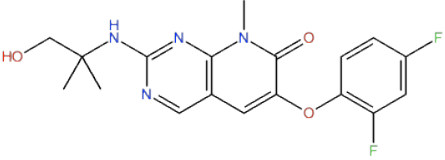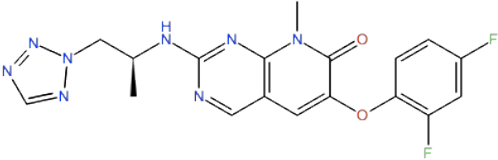**Table S1:** Representative p38 ligand pairs used in FEP/REST calculations
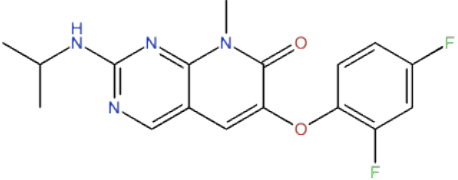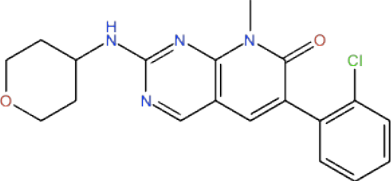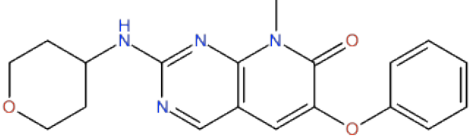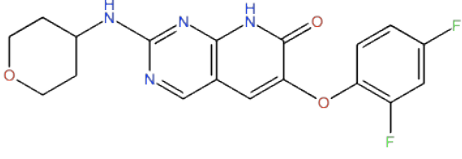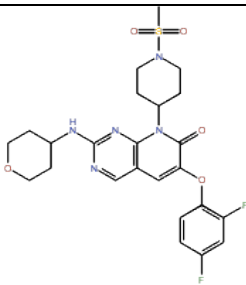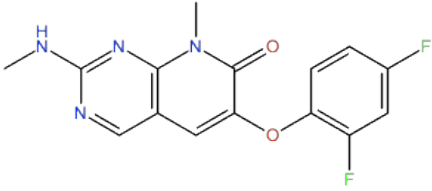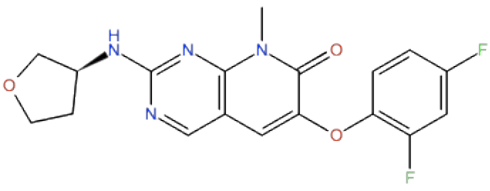
3fly	3fmh
2z	3fmk
2c	3flz
2k	2u
2v	2x

Table S2: Comparison of the RMSE for FEP predicted binding affinities for all 330 pairs of ligands and the RMSE assuming all the ligands are equally potent (guess 0 for all ΔΔG). (In units of kcal/mol)

ΔΔG (exp)	MUE(FEP)	MUE(guess 0)	RMSE(FEP)	RMSE(guess 0)
<1.0	0.82	0.46	1.05	0.54
1.0~1.5	0.90	1.26	1.11	1.27
1.5~2.0	1.06	1.75	1.24	1.75
>2.0	1.42	2.64	1.77	2.71

Table S3: The number of OPLS2.1 missing torsions identified for 3 ligand sets. The set of ligands studied in the present work, a set of approved drug molecules and the CACDB, a large database of drug-like compounds used in the development of OPLS2.1.

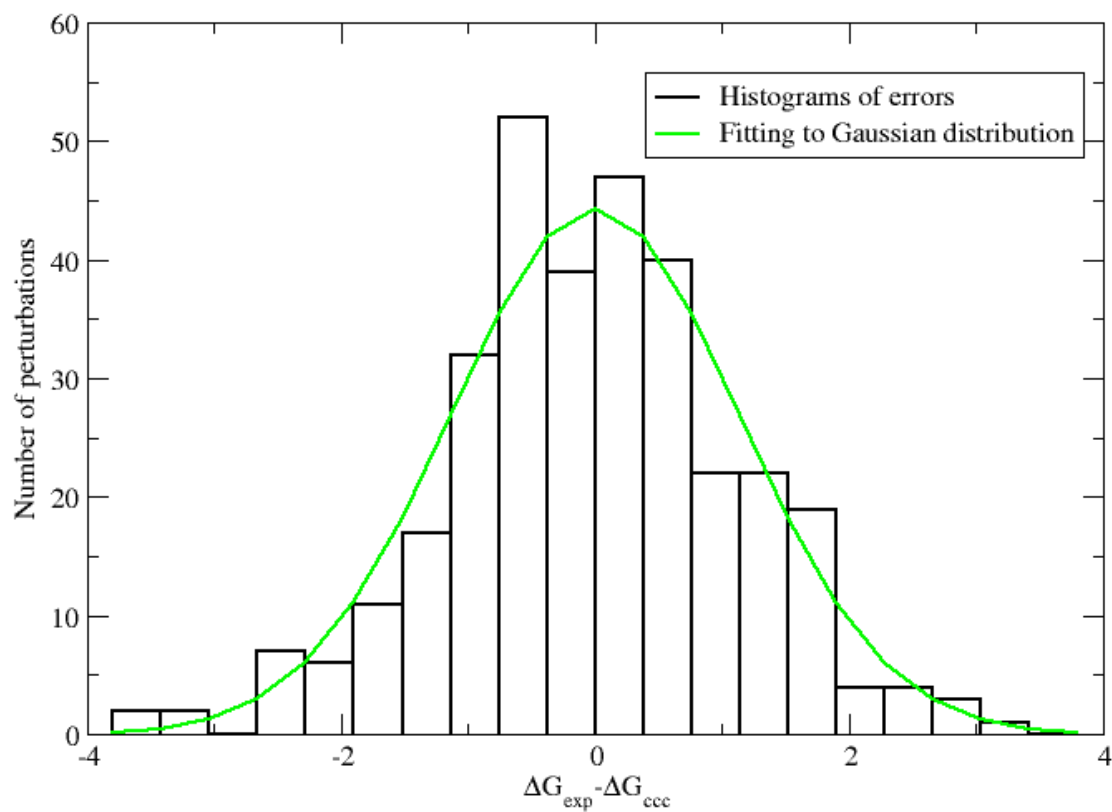| Ligand set	Number of Ligands	Fraction of Covered Rotatable Bonds	Fraction of Fully Covered Ligands
FEP Ligands	199	84%	57%
Approved Drug Molecules	1119	84%	53%
CACDB	>6M	93%	66%

Figure S1. The error distribution of the FEP predicted relative binding free energies compared to experimental data, and fitting of the distribution by a Gaussian function with standard deviation of 1.14. (1.14 kcal/mol is the RMSE of the FEP predicted binding free energies compared with experimental data.)
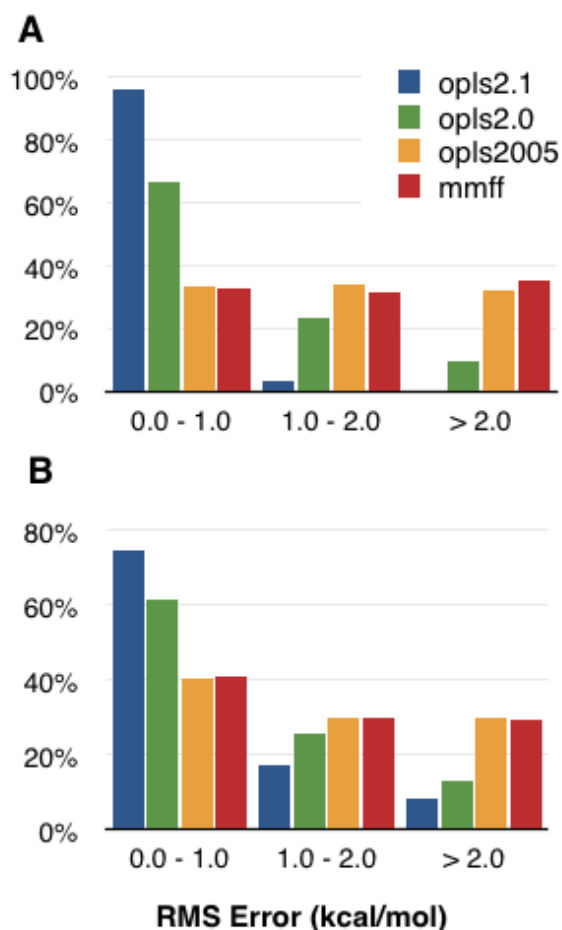
Figure S2. Histograms of the root mean square error in force-field relative energies evaluated over one-dimensional torsional angle scans (panel A) and between conformational minima (panel B) on a set of 8,365 compounds. Errors are established with respect to quantum mechanical LMP2/cc-pVTZ(-f) energies evaluated on B3LYP/6-31G* optimized structures.