

# Assessing the Performance of Docking, FEP, and MM/GBSA Methods on a Series of KLK6 Inhibitors

Wemenes José Lima Silva<sup>1</sup>, Renato Ferreira de Freitas<sup>1\*</sup>

<sup>1</sup>Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, São Bernardo do  
Campo, SP, Brazil; \*renato.f@ufabc.edu.br

## Abstract

Kallikrein 6 (KLK6) is an attractive drug target for the treatment of neurological diseases and for various cancers. Herein, we explore the accuracy and efficiency of different computational methods and protocols to predict the free energy of binding ( $\Delta G_{\text{bind}}$ ) of a series of KLK6 inhibitors. We found that the performance of the methods varied strongly with the tested system. For only one of the three KLK6 datasets, the docking scores were in good agreement ( $R^2 \geq 0.5$ ) with experimental values of  $\Delta G_{\text{bind}}$ . A similar result was obtained with MM/GBSA calculations based on single minimized structures. Improved binding affinity predictions were obtained with the free energy perturbation (FEP) method, with an overall MUE and RMSE of 0.53 and 0.68 kcal/mol, respectively. This result indicates that FEP can be a promising tool for the structure-based optimization of KLK6 inhibitors.

**Keywords:** Kallikrein 6, Free energy of binding, Docking, Molecular dynamics, MM/GBSA, Free energy perturbation

## Introduction

The serine protease kallikrein 6 (KLK6) is abundantly expressed in the central nervous system [1] and has been associated with neurodegenerative diseases. KLK6 cleaves Amyloid Precursor Protein (APP) in vitro [2] and, therefore, is able to generate amyloid beta peptides that accumulate in the form of amyloid plaques in the brain of patients with Alzheimer's disease. KLK6 also cleaves the  $\alpha$ -synuclein protein in vitro [3] which form abnormal aggregates of proteins known as Lewy bodies, which develop in the cytoplasm of nerve cells in Parkinson's disease and other neurodegenerative diseases. In addition, KLK6 hydrolyzes myelin basic protein (MBP) [4], and when its enzymatic activity is blocked with antibodies, there is a reduction in the severity of symptoms caused by central nervous system inflammation, a model for multiple sclerosis, in mice [5]. KLK6 also plays an important role in several cancers as component of the tumor microenvironment [6]. Recent studies have shown that inhibition of KLK6 reduced the invasiveness of pancreatic cancer cells [6, 7]. As a result, this enzyme may represent an attractive target for therapeutic development.

The KLK6 inhibitors reported so far were found through high-throughput screening (HTS) [8, 9], from natural products [10], and from yeast surface display (YSD) technology [11]. While some of these compounds are potent KLK6 inhibitors, usually they show low selectivity as a result of the structurally similar binding site to other KLKs and to other trypsin-like proteases [11].

Computational methods such as molecular docking, molecular dynamics (MD) simulations, and more recently the use of free energy perturbation methods are commonly used to identify and optimize serine protease inhibitors [12–16]. The common goal of all of these methods is to predict the free energy of binding ( $\Delta G_{\text{bind}}$ ) to find small molecules that bind strongly to the target receptor. This is especially important in lead optimization, which will

require the design, synthesis and biological evaluation of hundreds of compounds [17]. A computational method that can predict with reasonable accuracy the  $\Delta G_{\text{bind}}$  of congeneric ligands would enable the efficient selection of analogs for synthesis and testing. This would reduce significantly the time and costs of the lead optimization process [18].

Despite the wide application of in silico methods in the design of serine protease inhibitors, there are few reports where such methods were applied to discover new KLK inhibitors [19–21]. Here, we present a comparison of four different approaches to calculate absolute or relative binding affinities for four datasets of ligands of KLK6. We hope this work will showcase computational methods as valuable tools in the hit-to-lead optimization process of KLK6 inhibitors.

## **Material and methods**

### *Protein structure preparation*

Initially, we carried out a search in the literature for KLK6 inhibitors where at least one compound of the series was co-crystalized with the enzyme. After this search, two crystal structures deposited in the Protein Data Bank (PDB) [22] were selected: 3VFE and 4D8N. Then, Chimera [23] was used to prepare the structure of the complex for the in silico studies. This step involved the addition of hydrogen atoms, adjustment of the protonation state to pH 7.4 [24], calculation of Gasteiger charges, and minimization of the complex [25]. In the next step, the structures of compounds in complex with the enzyme were extracted and saved in SMILES format using Marvin Sketch (Marvin 19.26, 2022, ChemAxon). Then, OMEGA 3.0.0 (Santa Fe, NM, USA) was used to generate the 3D structure of these ligands.

### *Database preparation*

From the literature search, four KLK6 inhibitor datasets were selected (Table S1). The K61 inhibitor dataset contains 6 compounds with an amidinothiophene P1 group and a pyrrolidinone-sulphonamide scaffold linker [20]. The K62 and K64 inhibitor dataset contains 8 and 15 compounds, respectively, with a *p*-amidobenzylamine P1 group and a 2-hydroxybenzamide scaffold [19] [27]. The K63 inhibitor dataset contains 20 compounds with a N-(4-benzamidino)-oxazolidinone scaffold [8]. Their structures were saved in the SMILES format, and the potency ( $IC_{50}/pIC_{50}$ ) of these compounds were also extracted. Then, OMEGA 3.0.0 (Santa Fe, NM, USA) was used to generate the 3D structure of the ligands. In the next step, a flexible alignment was performed between the database compounds that share the same substructure of the co-crystallized ligand. For this, the script `tetheredMinimization.py` [29] was used. This RDKit script takes a reference 3D molecular structure and a set of ligands and does a tethered minimization based on the maximum common substructure (MCS) of the ligand and reference structure. Also, this script prepares an SDF file with a property containing all atoms that should be constrained, which can then be used by rDock [30] to do a tethered scaffold docking.

### *Docking Validation*

First, the co-crystallized compounds were redocked to their cognate receptors using the rDock software [30]. The second validation test was the cross-docking of the compounds. The accuracy of the docking was assessed based on the RMSD value of the best ranked pose. Pharmacophoric restrictions were applied to residues D189 and G193 in the two validation tests to ensure that the generated poses present the hydrogen bonds interactions observed in the crystallographic structures between inhibitor and enzyme.

### *Protocols for Molecular Docking*

Three docking protocols were applied to the database of inhibitors using rDock: scoring of the aligned pose (P1); minimization and scoring of the inhibitors (P2); tethered scaffold docking (150 runs) (P3).

### *Molecular Dynamics Simulation*

The enzyme-inhibitor complexes prepared previously were submitted to molecular dynamics simulations using the Amber20 [31]. For each complex, two different force fields were evaluated in the simulations: (1) ff19SB with OPC water model; and (2) ff14SB with SPCE water model. The complexes were solvated in an octahedral box and their charges were neutralized by Na<sup>+</sup> or Cl<sup>-</sup> ions using tleap [32, 33]. The solvated complexes were minimized in three stages: in the first stage, a force constant of 10.0 kcal/mol/Å<sup>2</sup> were applied on the heavy atoms of the complex, and the system were minimized for 5000 steps using the steepest descent method in the first 1000 steps and the conjugate gradient method for the rest of the steps; the second stage minimized the system for 2500 steps, without any restraints on the heavy atoms and applying the steepest descent method; in the third stage, the complex were minimized for 2500 steps, without restraints on the heavy atoms and applying the steepest descent method for the first 1000 steps and the conjugate gradient method for the rest of the steps. Then, the minimized complexes were equilibrated in six stages. First, the equilibration was carried out using the isothermal-isochoric (NVT) ensemble, and the temperature raised from 0 to 300K under the control of the Langevin Thermostat. The heavy atoms of the complex were restrained in this step using a force constant of 10.0 kcal/mol/Å<sup>2</sup>. The others equilibration stages were carried out using the isothermal-isobaric (NPT) ensemble with decreasing restraints on the

heavy atoms of the complexes of: 10.0, 5.0, 2.0, 0.5, and 0.1 kcal/mol/Å<sup>2</sup>. After the equilibration, the complexes were submitted to a production run for 20ns. At the end of the simulations, the free energy of binding between the inhibitor and enzyme was calculated using the MM/GBSA and MM/PBSA methods [34, 35] for different intervals (2-5ns, 2-10ns, 2-15ns, 2-20ns). These calculations were also performed for the minimized structure of the complex.

### *Free Energy Perturbation*

For FEP calculations the poses aligned previously with the python script were used as the initial binding poses. The perturbation maps were generated with the LOMAP mapping algorithm [36] for each dataset of KLK6 inhibitors. In the LOMAP algorithm, a maximum common substructure (MCS) is generated between pairs of compounds and the similarity is measured. Then, pairs of compounds with high similarities are connected by edges. Each edge represents a FEP calculation that will be performed between the pair of compounds.

Here, we use the FEP method combined with the improved sampling method, REST (Replica Exchange with Solute Tempering) [37]. The FEP/REST method presents an efficient  $\lambda$ -hopping protocol for sampling local structural rearrangements for the calculation of relative protein-ligand binding affinity within affordable simulation times.

All FEP/ REST calculations were conducted with the academic LigandFEP methodology of Desmond [38, 39] using the OPLS\_2005 force field [40]. The systems were solvated in an orthogonal box of SPC 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 protocol, consisting of: (i) a minimization using a Brownian dynamics NVT integrator for 100ps with the solute molecules restrained (50 kcal/mol/Å<sup>2</sup>), (ii) 12 ps simulation in the NVT ensemble, keeping the restraints and temperature at 10 K, (iii) 12 ps simulation in the NPT ensemble, keeping restraints and temperature at 10 K, (iv) 24 ps simulation in the NPT

simulation with solute heavy atom restraints at 300 K, and (v) 240 ps simulation in the NPT ensemble at 300 K without restraints. The FEP production runs were performed for 5ns in the NPT ensemble for both the complex and the solvent systems. The  $\lambda$ -hopping stage was split into 12 windows. The Bennett acceptance ratio method (BAR) [41] was used to calculate the free energy. Errors were estimated for each free energy calculation using both bootstrapping [42, 43] and the BAR analytical error prediction [41–43]. FEP/REST simulations were run on a single GPU Nvidia GeForce RTX 2080.

### *Statistical Correlation*

The statistical correlation between the experimental and the calculated free binding energy was evaluated using the coefficient of determination ( $R^2$ ), and Kendall rank correlation coefficient ( $\tau$ ) using the R function `cor()` implemented in the software R [44]. The mean absolute error (MAE) and root-mean-square error (RMSE) were calculated using the R functions `mae()` and `rmse()` [44].

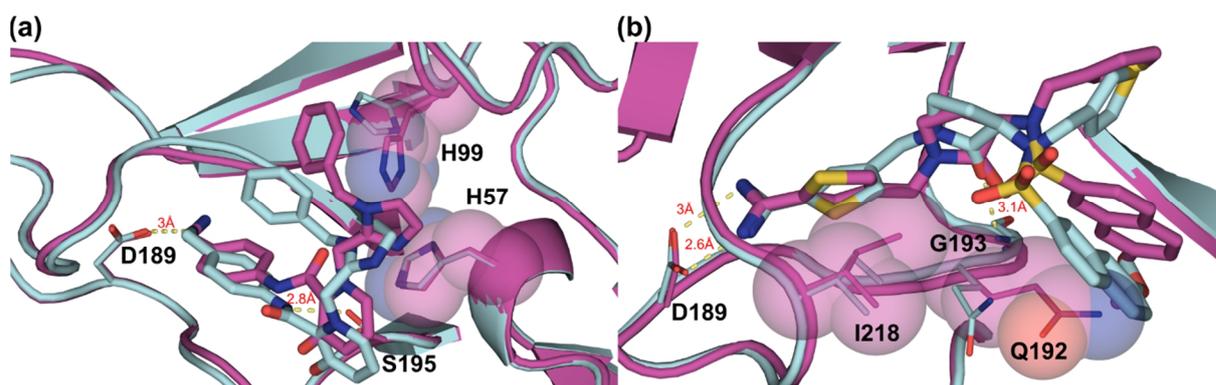
## Results

### *Docking validation*

rDock was able to redock the inhibitors with an RMSD  $< 2 \text{ \AA}$  for both inhibitors (Table S2). For the cross-docking validation, the RMSD values obtained with rDock were  $\leq 2.4 \text{ \AA}$ . Such results were already expected, since the side chains of the active site residues adjust to the structure of the inhibitors in order to optimize the interactions between the enzyme and the inhibitor, altering the interaction surface and preventing ligands with a different scaffold to bind with their observed conformation (Table S3).

In the cross-docking of the inhibitor 0HM in the 3VFE structure, the ligand moves towards the side chains of residues H57 and H99, losing its hydrogen bond with S195, but keeping the hydrogen bond with D189 (Figure 1a). The change in the conformation of the side chain of H99 in the 3VFE structure compared with 4D8N structure, favored the formation of a  $\pi$ -stacking interaction between the methyl-benzene and dimethylimidazole groups of the inhibitor 0HM and the imidazole ring of H99, and between the imidazole group of H57 and the dimethylimidazole group of the inhibitor.

For the cross-docking of the inhibitor 0HL in the 4D8N structure, the ligand assumed a conformation very close to the experimental structure with an RMSD of  $1.7 \text{ \AA}$  (Figure 1b). The minimal differences between the poses of the ligand can be explained by the side chain of Q192 that displaces the methoxy-naphthalene group and, consequently, the sulfonamide group of the inhibitor. Finally, the small difference in the conformation of thiophene is due to the displacement of the side chain of residue I218.



**Figure 1** - (a) Overlap between docked 0HM inhibitor (pink) in structure 3VFE (pink) and crystallized 0HM inhibitor (cyan) in structure 4D8N (cyan). (b) Overlap between docked 0HL inhibitor (pink) in structure 4D8N (pink) and crystallized 0HL inhibitor (cyan) in structure 3VFE (cyan).

### *Molecular docking*

Each dataset investigated here have similarities and differences in their structures. All of them have a basic group that binds the S1 pocket of KLK6. For the K62 dataset this group is the *p*-amidobenzylamine P1 group. The K61 dataset has an amidinothiophene group binding at the S1 pocket of KLK6, while the K63 dataset have a benzamidine, which is overrepresented in serine protease inhibitors. Aside from being similar in their P1 groups, the K61 and K63 datasets occupy similar sub pockets within the enzyme, despite having completely different structures. The diversity of these datasets leads to differences in the factors driving the free energy of binding such as hydrophobic interactions, hydrogen bonding interactions, solvation, and entropic effects. The computational methods evaluated here should, in principle, be able to capture these changes.

It is well known that docking scores almost never correlates with  $\Delta G_{\text{bind}}$ . This is a result of the several simplifications used in the overall docking process to allow the fast screen of huge libraries [45]. In the present study, for two datasets (K61 and K63), rDock scores failed to give any meaningful correlation with the experimental free energy. Surprisingly, for the K62 dataset, the docking scores resulted in a significant correlation ( $R^2 = 0.8$ ;  $\tau = 0.8$ ) with the

experimental  $\Delta G_{\text{bind}}$  (Table 1). This result was obtained with the P2 protocol, which minimized and calculated the score of a ligand at the binding site. The other two protocols (P1 and P3) tested resulted in poor correlations ( $R^2 < 0.2$ ). The poor performance of rDock were somewhat expected as docking scoring functions were developed primarily for virtual screening applications rather than lead optimization [46].

**Table 1.** Correlation statistics between  $\Delta G_{\text{exp}}$  and  $\Delta G_{\text{cal}}$  with different computational methods.

Dataset	Protocols	Docking	Minimization (MM/GBSA)	Production (MM/GBSA)
		$R^2/\tau$		
<b>K61 (n=6)</b>	P2	0.5*/0.5*	0.9/1.0	0.4*/0.5*
<b>K62 (n=8)</b>	P2	0.8/0.8	0.4*/0.1*	0.3*/-0.2*
<b>K63 (n=20)</b>	P1	0.2*/-0.1*	0.4/0.4	0.1*/0.2*

\*Statistically insignificant ( $p > 0.5$ )

### ***MM/GBSA binding free energy calculations***

We also assessed the performance of the MM/GBSA method to score the binding poses. Usually this method is considered more accurate since they have improved models for solvation and electrostatic interactions and conformational change compared to most docking programs [47]. However, in the present study the  $\Delta G_{\text{bind}}$  obtained after the minimization of the complex with the MM/GBSA method and the FF14SB force field resulted in a weak statistical correlation ( $R^2 \leq 0.4$ ) with the experimental values for two datasets (Table 1). The poor performance seems to be independent of the force field since similar results were obtained with the more recent FF19SB force field (Table S4). As observed with docking, the MM/GBSA method also failed to find a correlation with the  $\Delta G_{\text{bind}}$  for the K63 dataset. This disappointing result obtained with the two methods could be due to the incorrect pose of the ligands. This was the only dataset where the reference ligand was not crystallized with the enzyme. The reference

pose was obtained by docking and the other compounds in the dataset were aligned to this pose. However, the crystal structure of KLK6 in complex with compound **31** was released recently [48]. The docking pose was very similar (RMSD = 1.88 Å) to the binding mode shown by the cocrystal structure (data not shown), suggesting that an incorrect binding mode was not the cause of the poor performance of docking and MM/GBSA in predicting the  $\Delta G_{\text{bind}}$ .

The K61 dataset was the only dataset where calculated  $\Delta G_{\text{bind}}$  for the minimized complex showed a significant correlation ( $R^2 = 0.9$ ,  $\tau = 1.0$ ) with the experimental values. Also, it was also higher than the one obtained from docking. It has been suggested that neutral ligands are more amenable to MM-GBSA calculations [49]. According to Sun et al., ligand binding affinity prediction accuracy decreases with net charge of the ligand [50]. Since all the ligands studied here have a positive charge, this could partially explain the poor performance of the MM/GBSA method for two of the three datasets.

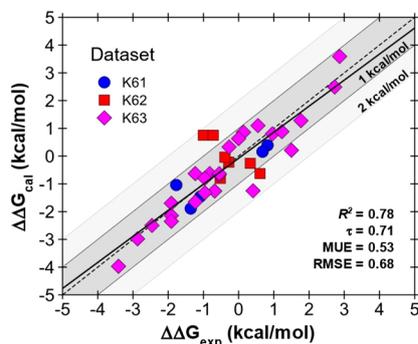
Our results showed that using the MM/GBSA method to estimate the free energy of binding based on molecular dynamics simulations of up to 20ns also failed to provide any correlation with the experimental values (Table 1). Regardless of the force field used, all datasets showed  $R^2$  values  $< 0.5$  and were statistically insignificant (Table S4). These results agree with previous work showing that the correlations provided by the MM/GBSA (and MM/PBSA) method are critically dependent on the tested receptor-ligand system [50, 51] and longer MD simulation is not always necessary to achieve better predictions [52]. A possible hypothesis for the poor correlations obtained with MD could be the instability of the complexes. However, all compounds in the three datasets were stable during the simulation with an average RMSD of 1.18 Å. Only 4 of 76 complexes showed an RMSD  $\geq 2.0$  Å, with the largest deviation being only 2.52 Å (Table S5).

### ***Binding affinities estimated by FEP***

Next, we evaluated if the FEP method would give a better correlation with the experimental values of  $\Delta G_{\text{bind}}$ . FEP is one of the most robust methods to predict the free energy of binding [53, 54]. Through a series of nonphysical intermediate states, FEP estimates the difference between two states' free energies by performing extended sampling at each intermediate state [55]. As this method is extremely time consuming, it is used more often in the lead optimization phase where less compounds are evaluated and the modifications are made around a conserved core. FEP has been applied with success, both retrospectively and prospectively, to predict the binding affinity of congeneric series of ligands of different systems [56–59].

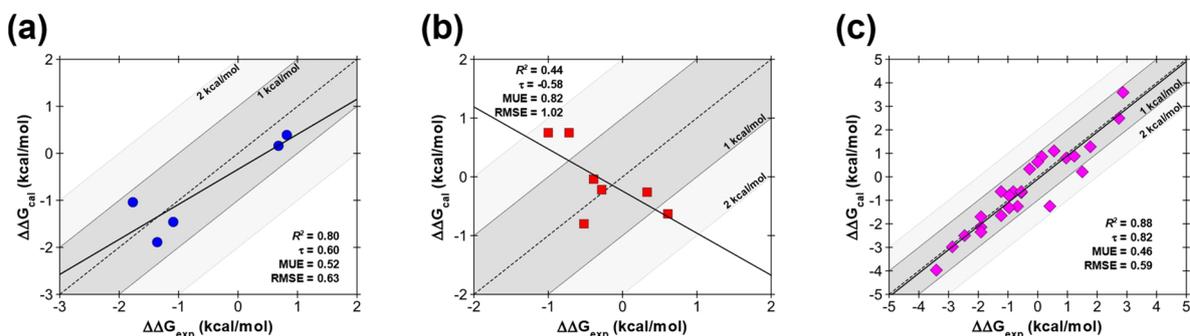
Overall, we have estimated the relative free energy of binding ( $\Delta\Delta G_{\text{bind}}$ ) of 37 compounds of KLK6 by FEP. LOMAP was used to generate the perturbation maps for each dataset. The maps involved 5, 7, and 25 transformations for the K61, K62, and K63 datasets, respectively. Structures of the individual ligands and the perturbation maps for the FEP calculations in each data set are given in Supporting Information. The scatter plot of predicted versus experimental  $\Delta\Delta G_{\text{bind}}$  for all three datasets is shown in Figure 2. FEP-predicted relative binding affinities for most of the ligands are below 1.0 kcal/mol of experimental values. Only five of the 37 transformations studied deviate from their experimental binding affinities by more than 1.0 kcal/mol. The average  $R^2$  and  $\tau$ , between FEP-predicted binding affinities and experimental results for all perturbations were 0.78 and 0.71, respectively. In addition, the MUE and RMSE were 0.53 kcal/mol and 0.68 kcal/mol, respectively. For 84.2% and 56.1% of the transformations, the MUE was  $\leq 1.0$  kcal/mol and  $\leq 0.5$  kcal/mol, respectively. It is promising that FEP can provide predictions with a MUE and RMSE values near 1 kcal/mol, which is a

cutoff for a computational method to be effectively used in the compound optimization process [60].

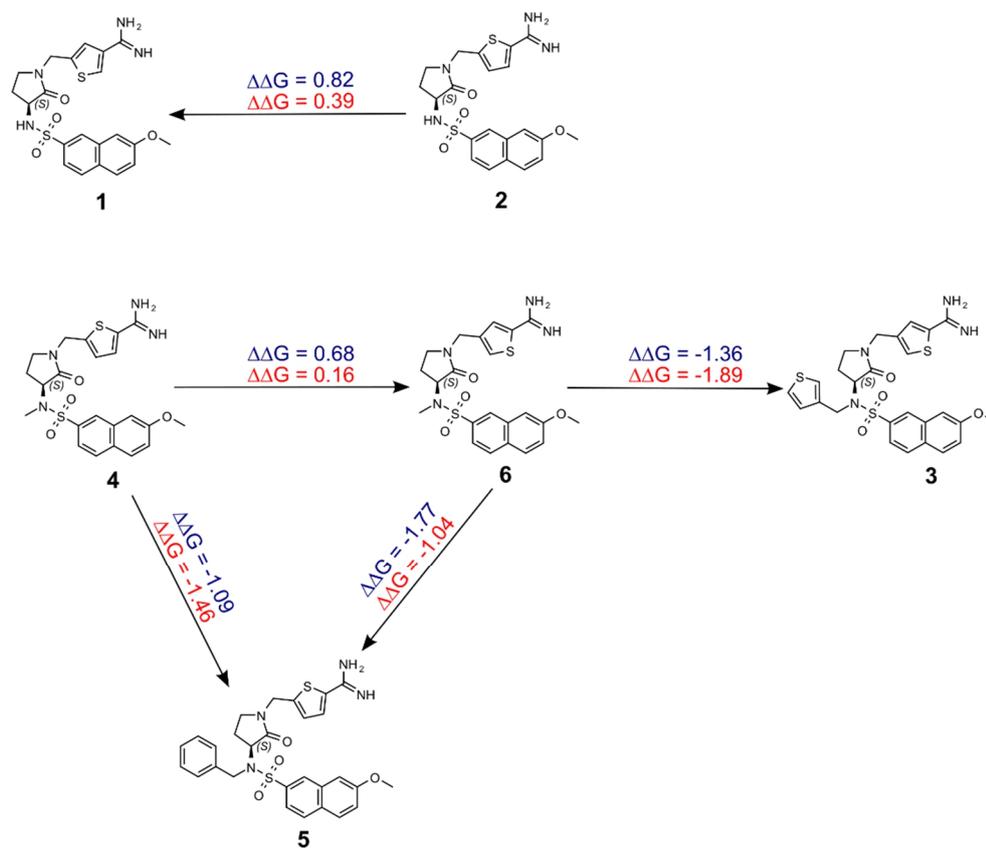


**Figure 2.** Correlation between FEP-predicted and experimental data for all four set.

For the K61 dataset, the values of  $R^2$ ,  $\tau$ , MUE and RMSE were 0.80, 0.60, 0.52 kcal/mol, and 0.63 kcal/mol respectively (Figure 3a). The K61 dataset has the smallest number of transformations ( $P = 5$ ) (Table S5). The transformation with the smallest absolute error was between **4**  $\rightarrow$  **5** (0.37 kcal/mol) in which a benzene group is added under a methyl group (Figure 4). On the other hand, the transformation between **6**  $\rightarrow$  **5** had the largest error (0.73 kcal/mol). This transformation involves a change in the attachment position the amidinothiophene group and the replacement of a benzene ring by a methyl group (Figure 4).

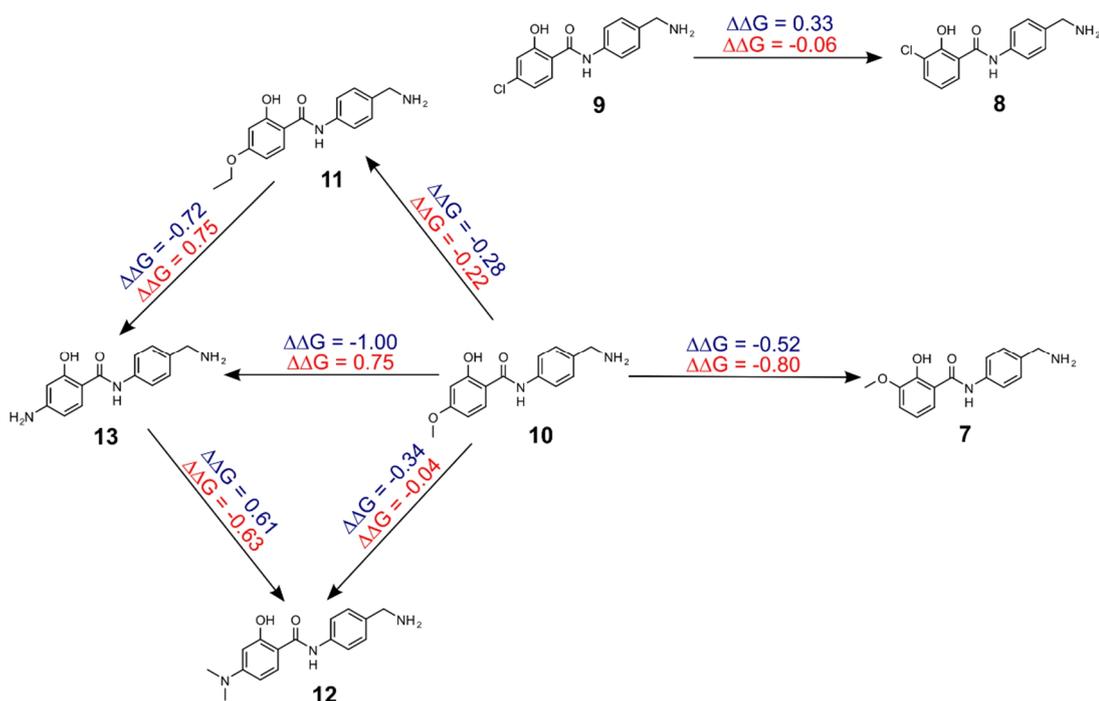


**Figure 3.** Correlation between FEP-predicted and experimental data for all KLK6 datasets. (a) K61 dataset; (b) K62 dataset; (c) K63 dataset.



**Figure 4.** Mutation graph used for the ligands of the K61 dataset. Arrows indicate the orientation of the chemical transformations. The experimental and calculated  $\Delta\Delta G$  values (in kcal/mol) are colored in blue and red, respectively.

With only 8 compounds, LOMAP built a map network with 7 perturbations for the K62 dataset. FEP results were actually anticorrelated with the experimental data for this dataset as evident from Figure 3b and the negative value of the Kendal coefficient ( $\tau = -0.59$ ). The lower correlation could be due to a change in the ionization state of the phenol oxygen, which was treated as neutral in all cases (except in the transformation **9** → **8**) but may vary based on the benzylic substituent (Figure 5).

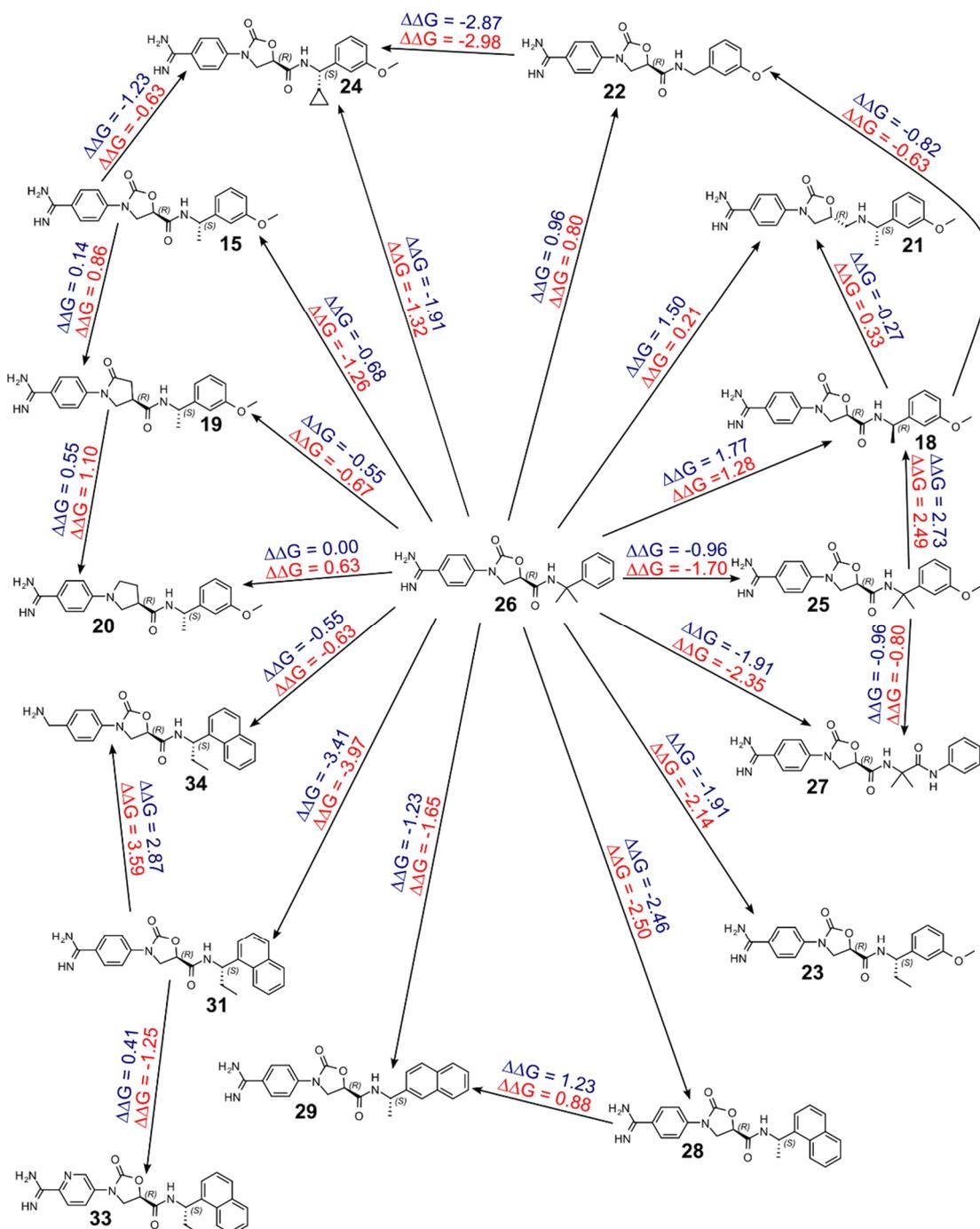


**Figure 5.** Mutation graph used for the ligands of the K62 dataset. Arrows indicate the orientation of the chemical transformations. The experimental and calculated  $\Delta\Delta G$  values (in kcal/mol) are colored in blue and red, respectively.

In addition, for only three transformations FEP predicted the correct sign with small absolute errors ( $\leq 0.3$  kcal/mol). However, the sign of the other four transformations were wrongly predicted with errors  $\geq 0.6$  kcal/mol. The correct prediction of the sign of the ligand transformations, i.e. whether one ligand will be more or less active, plays a critical role during the lead optimization process, especially for a limited number of ligands, and is a fundamental goal of FEP [61].

The best results were obtained for the K63 dataset. FEP predicted relative binding affinities were in a good agreement with experimental data: RMSE = 0.59 kcal/mol, MUE = 0.46 kcal/mol,  $R^2 = 0.88$ , and  $\tau = 0.82$ . This dataset is the one with more compounds (16) and perturbations (25) among all datasets studied here (Figure 6). Furthermore, for this dataset, 15 (out of 25) transformations had an error  $\leq 0.5$  kcal/mol and for 23 (out of 25) the

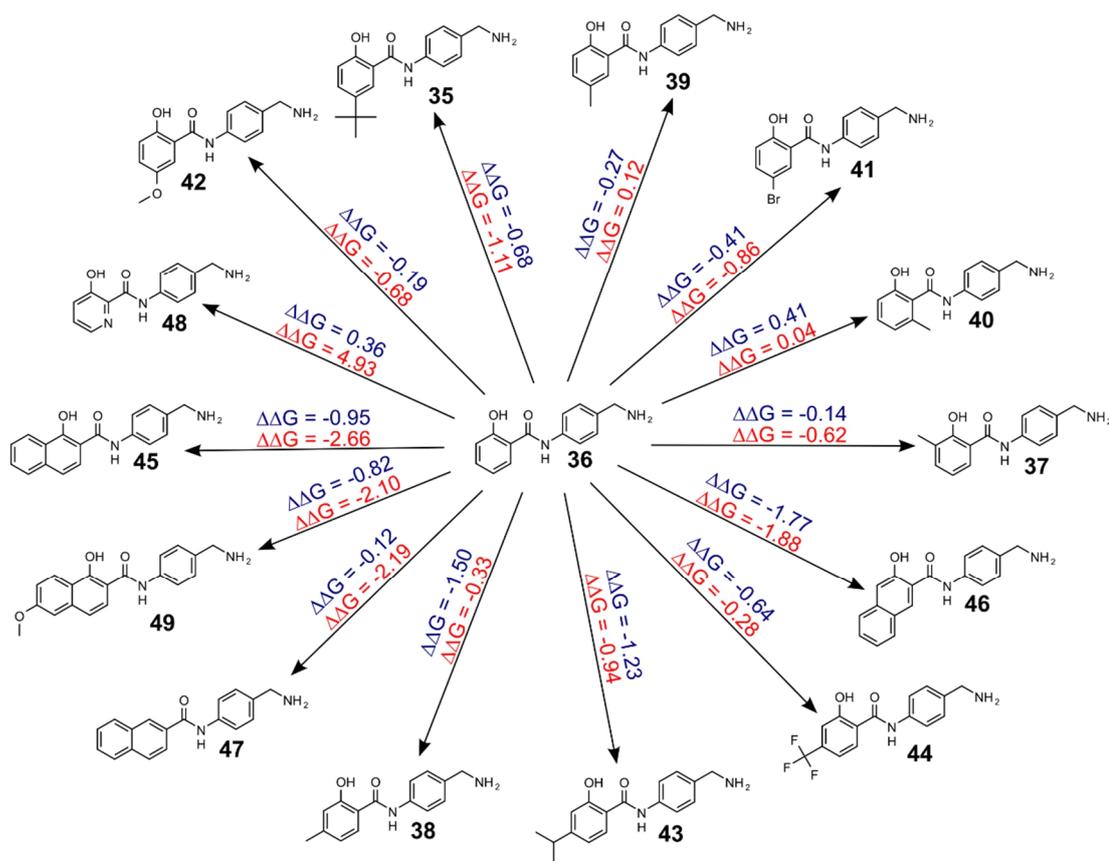
error was  $\leq 0.72$  kcal/mol. The FEP method also predicted the correct sign of the transformations for 23 pairs.



**Figure 6.** Mutation graph used for the ligands of the K63 dataset. Arrows indicate the orientation of the chemical transformations. The experimental and calculated  $\Delta\Delta G$  values (in kcal/mol) are colored in blue and red, respectively.

It is worthwhile to mention that at the time this work was conducted there was no crystal structure available with any of the compounds in this dataset. The predicted binding pose of compound **15** in the KLK6 active site was obtained with rDock, and all the other ligands were aligned to this pose. To our knowledge, there are just a few studies in which FEP was applied to predict the binding affinities of congeneric series of ligands using homology models or binding poses predicted by docking [62, 63]. The results were mixed. In some systems, the performance using homology models or the binding mode hypothesis obtained from docking was generally consistent with experimental data. However, in other test cases, unsatisfactory prediction accuracy has been obtained.

Finally, the last dataset evaluated here is a series of 15 compounds (K64) with the same scaffold as in the K62 dataset. For six of them (**35**, **38**, **43**, **45**, **46**, and **49**) the IC<sub>50</sub> is available, but they have a narrow range (1 kcal/mol) in the experimental free energy. Thus, a correlation between experiment and calculated  $\Delta G_{\text{bind}}$  should not be expected [64, 65]. The other nine compounds only have the percentage of inhibition at 10  $\mu\text{M}$ . So, we decided to simulate a real-world scenario where FEP would be used to prioritize the synthesis of compounds. We selected the smallest compound (**36**) as reference to which all other compounds were mutated (Figure 7). Encouragingly, all six compounds were identified as more potent than the reference, and four are among the top-5 predictions. FEP would have found all six actives if only 10 compounds were made. Moreover, four inactive compounds would not have been synthesized, saving valuable resources.



**Figure 7.** Mutation graph used for the ligands of the K64 dataset. Arrows indicate the orientation of the chemical transformations. The experimental and calculated  $\Delta\Delta G$  values (in kcal/mol) are colored in blue and red, respectively.

## Conclusion

The computational methods evaluated here were not accurate in predicting the binding free energy for all datasets. Docking and the MM/GBSA method (applied to a minimized pose), showed a good correlation with the experimental values for only one dataset each. Molecular dynamics simulations, usually considered more robust and more computationally expensive, gave poor results. Of the methods tested, FEP appeared to be the most accurate and we were able to obtain high accuracy predictions ( $RMSE < 1$  kcal/mol) for two datasets, including one where the experimental binding mode were not available. Finally,

in a simulation of a real-world drug discovery project, FEP was effective in ranking the most potent compounds at the top of the list. In summary, FEP has the potential to be successfully applied to the lead optimization stage of new KLK6 inhibitors.

### **Acknowledgements**

W.J.L.S. and R.F.F. gratefully acknowledges funding from São Paulo Research Foundation – FAPESP (2021/04450-7, 2018/11011-7 and 2019/08603-2). We are grateful to OpenEye Scientific Software, Inc. for providing us with an academic license for Omega. We thank D.E. Shaw Research for providing us with an academic license for Desmond.

### **Contributions**

W.J.L.S. and R.F.F. designed the study and analyzed the data. W.J.L.S. performed the simulations. W.J.L.S. and R.F.F. wrote the manuscript. R.F.F. was responsible for the project. All authors reviewed the manuscript.

### **Competing Interest**

The authors declare no competing interest.

### **References**

1. Yousef GM, Kishi T, Diamandis EP (2003) Role of kallikrein enzymes in the central nervous system. *Clin Chim Acta* 329:1–8. <https://doi.org/10.1016/S0009->

8981(03)00004-4

2. Magklara A, Mellati AA, Wasney GA, et al (2003) Characterization of the enzymatic activity of human kallikrein 6: Autoactivation, substrate specificity, and regulation by inhibitors. *Biochem Biophys Res Commun* 307:948–955. [https://doi.org/10.1016/S0006-291X\(03\)01271-3](https://doi.org/10.1016/S0006-291X(03)01271-3)
3. Pampalakis G, Sykioti VS, Ximerakis M, et al (2017) KLK6 proteolysis is implicated in the turnover and uptake of extracellular alpha-synuclein species. *Oncotarget* 8:14502–14515. <https://doi.org/10.18632/oncotarget.13264>
4. Blaber SI, Scarisbrick IA, Bennett MJ, et al (2002) Enzymatic properties of rat myelencephalon-specific protease. *Biochemistry* 41:1165–1173. <https://doi.org/10.1021/bi015781a>
5. Blaber SI, Ciric B, Christophi GP, et al (2004) Targeting kallikrein 6 proteolysis attenuates CNS inflammatory disease. *FASEB J* 18:920–922. <https://doi.org/10.1096/fj.03-1212fje>
6. Werner J, Bernhard P, Cosenza-Contreras M, et al (2023) Targeted and explorative profiling of kallikrein proteases and global proteome biology of pancreatic ductal adenocarcinoma, chronic pancreatitis, and normal pancreas highlights disease-specific proteome remodelling. *Neoplasia* 36:100871. <https://doi.org/10.1016/j.neo.2022.100871>
7. Zhang L, Lovell S, De Vita E, et al (2022) A KLK6 Activity-Based Probe Reveals a Role for KLK6 Activity in Pancreatic Cancer Cell Invasion. *J Am Chem Soc* 144:22493–22504. <https://doi.org/10.1021/jacs.2c07378>
8. De Vita E, Smits N, van den Hurk H, et al (2020) Synthesis and Structure-Activity Relationships of N-(4-Benzamidino)-Oxazolidinones: Potent and Selective Inhibitors of Kallikrein-Related Peptidase 6. *ChemMedChem* 15:79–95. <https://doi.org/10.1002/cmdc.201900536>

9. De Vita E, Schüler P, Lovell S, et al (2018) Depsipeptides Featuring a Neutral P1 Are Potent Inhibitors of Kallikrein-Related Peptidase 6 with On-Target Cellular Activity. *J Med Chem* 61:8859–8874. <https://doi.org/10.1021/acs.jmedchem.8b01106>
10. Soualmia F, Bosc E, Amiri SA, et al (2018) Insights into the activity control of the kallikrein-related peptidase 6: Small-molecule modulators and allostereism. *Biol Chem* 399:1073–1078. <https://doi.org/10.1515/hsz-2017-0336>
11. Sananes A, Cohen I, Shahar A, et al (2018) A potent, proteolysis-resistant inhibitor of kallikrein-related peptidase 6 (KLK6) for cancer therapy, developed by combinatorial engineering. *J Biol Chem* 293:12663–12680. <https://doi.org/10.1074/jbc.RA117.000871>
12. London N, Miller RM, Krishnan S, et al (2014) Covalent docking of large libraries for the discovery of chemical probes. *Nat Chem Biol* 10:1066–1072. <https://doi.org/10.1038/nchembio.1666>
13. Deng J, Li N, Liu H, et al (2012) Discovery of Novel Small Molecule Inhibitors of Dengue Viral NS2B-NS3 Protease Using Virtual Screening and Scaffold Hopping. *J Med Chem* 55:6278–6293. <https://doi.org/10.1021/jm300146f>
14. Yang T, Wu JC, Yan C, et al (2011) Virtual screening using molecular simulations. *Proteins Struct Funct Bioinforma* 79:1940–1951. <https://doi.org/10.1002/prot.23018>
15. Mondal D, Florian J, Warshel A (2019) Exploring the Effectiveness of Binding Free Energy Calculations. *J Phys Chem B* 123:8910–8915. <https://doi.org/10.1021/acs.jpcb.9b07593>
16. Fernández-Bachiller MI, Hwang S, Schembri ME, et al (2022) Probing Factor Xa Protein–Ligand Interactions: Accurate Free Energy Calculations and Experimental Validations of Two Series of High-Affinity Ligands. *J Med Chem* 65:13013–13028. <https://doi.org/10.1021/acs.jmedchem.2c00865>
17. Paul SM, Mytelka DS, Dunwiddie CT, et al (2010) How to improve RD productivity:

- The pharmaceutical industry's grand challenge. *Nat Rev Drug Discov* 9:203–214.  
<https://doi.org/10.1038/nrd3078>
18. Homeyer N, Stoll F, Hillisch A, Gohlke H (2014) Binding free energy calculations for lead optimization: Assessment of their accuracy in an industrial drug design context. *J Chem Theory Comput* 10:3331–3344. <https://doi.org/10.1021/ct5000296>
  19. Liang G, Chen X, Aldous S, et al (2012) Human kallikrein 6 inhibitors with a para-amidobenzylanmine P1 group identified through virtual screening. *Bioorganic Med Chem Lett* 22:2450–2455. <https://doi.org/10.1016/j.bmcl.2012.02.014>
  20. Liang G, Chen X, Aldous S, et al (2012) Virtual screening and x-ray crystallography for human kallikrein 6 inhibitors with an amidinothiophene p1 group. *ACS Med Chem Lett* 3:159–164. <https://doi.org/10.1021/ml200291e>
  21. Tan X, Bertonati C, Qin L, et al (2013) Identification by in silico and in vitro screenings of small organic molecules acting as reversible inhibitors of kallikreins. *Eur J Med Chem* 70:661–668. <https://doi.org/10.1016/j.ejmech.2013.10.040>
  22. Berman HM (2000) The Protein Data Bank. *Nucleic Acids Res* 28:235–242. <https://doi.org/10.1093/nar/28.1.235>
  23. Pettersen EF, Goddard TD, Huang CC, et al (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612
  24. Ahmed HU, Blakeley MP, Cianci M, et al (2007) The determination of protonation states in proteins. *Acta Crystallogr Sect D Biol Crystallogr* 63:906–922. <https://doi.org/10.1107/S0907444907029976>
  25. Shapovalov M V, Dunbrack Jr RL (2011) A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure* 19:844–858
  26. ChemAxon L (2013) MarvinSketch

27. Aït Amiri S, Deboux C, Soualmia F, et al (2021) Identification of First-in-Class Inhibitors of Kallikrein-Related Peptidase 6 That Promote Oligodendrocyte Differentiation. *J Med Chem* 64:5667–5688. <https://doi.org/10.1021/acs.jmedchem.0c02175>
28. OMEGA (2013) OpenEye Scientific Software,
29. Schmidtke P (2019) Tethered Minimization
30. Ruiz-Carmona S, Alvarez-Garcia D, Foloppe N, et al (2014) rDock: A Fast, Versatile and Open Source Program for Docking Ligands to Proteins and Nucleic Acids. *PLoS Comput Biol* 10:1–8. <https://doi.org/10.1371/journal.pcbi.1003571>
31. Salomon-Ferrer R, Case DA, Walker RC (2013) An overview of the Amber biomolecular simulation package. *Wiley Interdiscip Rev Comput Mol Sci* 3:198–210
32. Machado MR, Pantano S (2020) Split the Charge Difference in Two! A Rule of Thumb for Adding Proper Amounts of Ions in MD Simulations. *J Chem Theory Comput* 16:1367–1372. <https://doi.org/10.1021/acs.jctc.9b00953>
33. Schmit JD, Kariyawasam NL, Needham V, Smith PE (2018) SLTCAP: A Simple Method for Calculating the Number of Ions Needed for MD Simulation. *J Chem Theory Comput* 14:1823–1827. <https://doi.org/10.1021/acs.jctc.7b01254>
34. Kollman PA, Massova I, Reyes C, et al (2000) Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc Chem Res* 33:889–897. <https://doi.org/10.1021/ar000033j>
35. Tsui V, Case DA (2000) Theory and applications of the generalized born solvation model in macromolecular simulations. *Biopolymers* 56:275–291. [https://doi.org/10.1002/1097-0282\(2000\)56:4<275::AID-BIP10024>3.0.CO;2-E](https://doi.org/10.1002/1097-0282(2000)56:4<275::AID-BIP10024>3.0.CO;2-E)
36. Liu S, Wu Y, Lin T, et al (2013) Lead optimization mapper: Automating free energy calculations for lead optimization. *J Comput Aided Mol Des* 27:755–770.

- <https://doi.org/10.1007/s10822-013-9678-y>
37. Wang L, Berne BJ, Friesner RA (2012) On achieving high accuracy and reliability in the calculation of relative protein-ligand binding affinities. *Proc Natl Acad Sci U S A* 109:1937–1942. <https://doi.org/10.1073/pnas.1114017109>
  38. Bowers KJ, Chow DE, Xu H, et al (2006) Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In: *ACM/IEEE SC 2006 Conference (SC'06)*. IEEE, pp 43–43
  39. D. E. Shaw Research (2017) Desmond Molecular Dynamics System
  40. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J Phys Chem B* 105:6474–6487. <https://doi.org/10.1021/jp003919d>
  41. Bennett CH (1976) Efficient estimation of free energy differences from Monte Carlo data. *J Comput Phys* 22:245–268. [https://doi.org/10.1016/0021-9991\(76\)90078-4](https://doi.org/10.1016/0021-9991(76)90078-4)
  42. Pohorille A, Jarzynski C, Chipot C (2010) Good Practices in Free-Energy Calculations. *J Phys Chem B* 114:10235–10253. <https://doi.org/10.1021/jp102971x>
  43. Paliwal H, Shirts MR (2011) A Benchmark Test Set for Alchemical Free Energy Transformations and Its Use to Quantify Error in Common Free Energy Methods. *J Chem Theory Comput* 7:4115–4134. <https://doi.org/10.1021/ct2003995>
  44. R Core Team (2020) R: A Language and Environment for Statistical Computing
  45. Pantsar T, Poso A (2018) Binding Affinity via Docking: Fact and Fiction. *Molecules* 23:1899. <https://doi.org/10.3390/molecules23081899>
  46. Wang L, Wu Y, Deng Y, et al (2015) 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* 137:2695–2703.

- <https://doi.org/10.1021/ja512751q>
47. Graves AP, Shivakumar DM, Boyce SE, et al (2008) Rescoring Docking Hit Lists for Model Cavity Sites: Predictions and Experimental Testing. *J Mol Biol* 377:914–934. <https://doi.org/10.1016/j.jmb.2008.01.049>
  48. Baumann A, Isak D, Lohbeck J, et al (2022) Scalable synthesis and structural characterization of reversible KLK6 inhibitors. *RSC Adv* 12:26989–26993. <https://doi.org/10.1039/D2RA04670A>
  49. El Khoury L, Santos-Martins D, Sasmal S, et al (2019) Comparison of affinity ranking using AutoDock-GPU and MM-GBSA scores for BACE-1 inhibitors in the D3R Grand Challenge 4. *J Comput Aided Mol Des* 33:1011–1020. <https://doi.org/10.1007/s10822-019-00240-w>
  50. Sun H, Li Y, Tian S, et al (2014) Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set. *Phys Chem Chem Phys* 16:16719–16729. <https://doi.org/10.1039/C4CP01388C>
  51. Genheden S, Ryde U (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov* 10:449–461. <https://doi.org/10.1517/17460441.2015.1032936>
  52. Hou T, Wang J, Li Y, Wang W (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model* 51:69–82. <https://doi.org/10.1021/ci100275a>
  53. Wang L, Wu Y, Deng Y, et al (2015) 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* 137:2695–2703.

- <https://doi.org/10.1021/ja512751q>
54. Cournia Z, Allen B, Sherman W (2017) Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J Chem Inf Model* 57:2911–2937. <https://doi.org/10.1021/acs.jcim.7b00564>
  55. Wang L, Chambers J, Abel R (2019) Protein–Ligand Binding Free Energy Calculations with FEP+. pp 201–232
  56. O’ Donovan DH, Gregson C, Packer MJ, et al (2021) Free energy perturbation in the design of EED ligands as inhibitors of polycomb repressive complex 2 (PRC2) methyltransferase. *Bioorganic Med Chem Lett* 39:127904. <https://doi.org/10.1016/j.bmcl.2021.127904>
  57. Albanese SK, Chodera JD, Volkamer A, et al (2020) Is Structure-Based Drug Design Ready for Selectivity Optimization? *J Chem Inf Model* 60:6211–6227. <https://doi.org/10.1021/acs.jcim.0c00815>
  58. Deflorian F, Perez-Benito L, Lenselink EB, et al (2020) Accurate Prediction of GPCR Ligand Binding Affinity with Free Energy Perturbation. *J Chem Inf Model* 60:5563–5579. <https://doi.org/10.1021/acs.jcim.0c00449>
  59. Zara L, Moraca F, Van Muijlwijk-Koezen JE, et al (2022) Exploring the Activity Profile of TbrPDEB1 and hPDE4 Inhibitors Using Free Energy Perturbation. *ACS Med Chem Lett* 13:904–910. <https://doi.org/10.1021/acsmchemlett.1c00690>
  60. He X, Liu S, Lee T-S, et al (2020) Fast, Accurate, and Reliable Protocols for Routine Calculations of Protein–Ligand Binding Affinities in Drug Design Projects Using AMBER GPU-TI with ff14SB/GAFF. *ACS Omega* 5:4611–4619. <https://doi.org/10.1021/acsomega.9b04233>
  61. Fratev F, Sirimulla S (2019) An Improved Free Energy Perturbation FEP+ Sampling Protocol for Flexible Ligand-Binding Domains. *Sci Rep* 9:16829.

- <https://doi.org/10.1038/s41598-019-53133-1>
62. Schindler CEM, Baumann H, Blum A, et al (2020) Large-Scale Assessment of Binding Free Energy Calculations in Active Drug Discovery Projects. <https://doi.org/10.1021/acs.jcim.0c00900>
  63. Cappel D, Hall ML, Lenselink EB, et al (2016) Relative Binding Free Energy Calculations Applied to Protein Homology Models. <https://doi.org/10.1021/acs.jcim.6b00362>
  64. Athanasiou C, Vasilakaki S, Dellis D, Cournia Z (2018) Using physics-based pose predictions and free energy perturbation calculations to predict binding poses and relative binding affinities for FXR ligands in the D3R Grand Challenge 2. *J Comput Aided Mol Des* 32:21–44. <https://doi.org/10.1007/s10822-017-0075-9>
  65. Bhati AP, Coveney P V (2022) Large Scale Study of Ligand – Protein Relative Binding Free Energy Calculations : Actionable Predictions from Statistically Robust Protocols. <https://doi.org/10.1021/acs.jctc.1c01288>