

# PREDICTING SELECTIVITY AND THE FUNCTIONAL IMPACT OF CLINICAL CANCER MUTATIONS FOR SMALL MOLECULE KINASE INHIBITORS USING PHYSICAL MODELING

A Dissertation

Presented to the Faculty of the Louis V. Gerstner Jr.,

Graduate School of Biomedical Sciences,

Memorial Sloan Kettering Cancer Center

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Steven K. Albanese

February 2019

© 2019 Steven K. Albanese  
ALL RIGHTS RESERVED

## ABSTRACT

Small molecule kinase inhibitors have become a major focus of drug development for treating cancer, which accounted for 610,000 deaths and 1.7 million diagnoses in the United States in 2018 alone. Currently, there are 43 FDA approved small-molecule kinase inhibitors. The dominant paradigm for designing such inhibitors has been to optimize maximally selective ligands for a single target. Unfortunately, many such inhibitors fail in clinical trials due to a lack of efficacy and clinical safety. Tumors can evade inhibitors through multiple routes of resistance, including upregulation of a second kinase, mutations in the target kinase, or amplification of the target kinase. On the other hand, toxicity arises from on-target inhibition of the wild type kinase or off-target effects of promiscuous small molecules or their metabolites. ATP-competitive kinase inhibitors have great potential for promiscuity, as there are over 520 kinases in the human kinome that each bind a common substrate, ATP. Further, advances in sequencing technology have enabled the generation of datasets of disease associated alterations rich in missense mutations in kinases. While this technology has been particularly transformative in the field of oncology, where many patients are treated with kinase targeted therapies, most kinase missense mutations are rare, making it difficult to assess their functional impact. Physical modeling can provide a route for predicting small molecule kinase inhibitor selectivity, and the impact that missense mutations have on kinase structure and inhibitor binding. To assess the utility of free energy calculations for predicting selectivity, we performed relative free energy calculations on publicly available congeneric series of ligands on multiple kinase targets. We built a Bayesian graphical model to quantitate the correlation of errors for a given ligand on both target, to interrogate whether any fortuitous

cancellation of errors makes selectivity predictions more accurate than expected. To understand the functional impact of mutations on protein kinase structure and inhibitor binding, we performed massively parallel molecular dynamics simulations of clinically observed hyperactivating mTOR mutations and alchemical free energy calculations on clinical Abl mutations. To rigorously test our predictions, we have developed a panel of kinase expression constructs (now available through AddGene) appropriate for automated, high-throughput expression protocol in *E. coli*. We have demonstrated the utility of these constructs for engineering and expressing clinically observed missense mutations, testing a panel of 96 mutations in Src and Abl kinases gathered from publicly available cancer genomics datasets as well as the MSK-IMPACT clinical sequencing panel, and further expressed a separate panel of 95 clinically relevant Abl mutations. We measured the binding free energies for these clinically relevant Abl mutations for a panel of FDA-approved small molecule kinase inhibitors. Using this dataset as a benchmark, we tested the sensitivity and accuracy of absolute free energy calculations to predict the impact of mutations on inhibitor binding. Taken together, this work provides an assessment of physical modeling for predicting selectivity and resistance in drug design of small molecule kinase inhibitors.

## **BIOGRAPHICAL SKETCH**

Your biosketch goes here. Make sure it sits inside the brackets.

This document is dedicated to all gsk graduate students.

## **ACKNOWLEDGEMENTS**

Your acknowledgements go here. Make sure it sits inside the brackets.

# Table of Contents

Biographical Sketch . . . . .	v
Dedication . . . . .	vi
Acknowledgements . . . . .	vii
Table of Contents . . . . .	viii
List of Tables . . . . .	x
List of Figures . . . . .	xi
<b>1 Introduction</b>	<b>1</b>
<b>2 Predicting the selectivity of small molecule kinase inhibitors</b>	<b>2</b>
2.1 Introduction . . . . .	2
2.1.1 Selectivity is an important consideration in drug design . . . . .	2
2.1.2 Physical modeling can complement high throughput screens and machine learning to predict selectivity . . . . .	3
2.1.3 Alchemical free energy methods can accurately predict single target potencies and the impact of mutations . . . . .	4
2.1.4 Assessing the ability of alchemical free energy methods to predict selectivity . . . . .	5
2.2 Methods . . . . .	6
2.2.1 Numerical model of selectivity . . . . .	6
2.2.2 Structure Preparation . . . . .	7
2.2.3 Ligand Pose Generation . . . . .	8
2.2.4 Free Energy Calculations . . . . .	8
2.2.5 Charge Change Free Energy Calculations . . . . .	9
2.2.6 Statistical Analysis of FEP+ calculations . . . . .	9
2.2.7 Quantification of the correlation coefficient $\rho$ . . . . .	10
2.3 Results . . . . .	12
2.3.1 Correlation of errors can make selectivity predictions more accurate and speed up ligand optimization . . . . .	12
2.3.2 The CDK2 and CDK9 experimental dataset demonstrates the difficulty in achieving selectivity for closely related kinases . . . . .	15
2.3.3 The CDK2 and ERK2 dataset achieves higher levels of selectivity for more distantly related kinases . . . . .	18
2.3.4 FEP+ calculations show accurate potency predictions for ERK2/CDK2 and larger errors for CDK2/CDK9 . . . . .	21



2.3.5	Free energy calculation errors are correlated, accelerating selectivity optimization . . . . .	23
2.4	Discussion and Conclusions . . . . .	26
<b>3</b>	<b>Understanding the functional impact of mTOR clinical kinase mutations using physical modeling</b>	<b>31</b>
3.1	Introduction . . . . .	31
3.1.1	Section on mTOR biology and mutations . . . . .	31
3.1.2	Section on MD simulations and Folding@Home . . . . .	31
3.2	Methods . . . . .	31
<b>4</b>	<b>Predicting the impact of clinically-observed Abl inhibitors using physical modeling and biophysical experiments</b>	<b>32</b>
4.1	Introduction . . . . .	32
4.1.1	The long tail of rare kinase mutations frustrates prediction of drug resistance . . . . .	33
4.1.2	Alchemical free-energy methods can predict inhibitor binding affinities . . . . .	35
4.1.3	Alchemical approaches can predict the impact of protein mutations on free energy . . . . .	35
4.1.4	Assessing the potential for physical modeling to predict resistance to FDA-approved TKIs . . . . .	36
4.1.5	A benchmark of $\Delta pIC_{50}$ s for predicting mutational resistance .	37
4.1.6	Summary of experimental data that is available and what needs to be done to make better information available . . . . .	42
4.1.7	Open-source YANK/Amber forcefield stuff . . . . .	42

# List of Tables

4.1	Public $\Delta pIC_{50}$ datasets for 144 Abl kinase mutations and eight tyrosine kinase inhibitors (TKIs) with corresponding wild-type co-crystal structures . . . . .	39
-----	---	----

# List of Figures

2.1	Free energy calculations speed up selectivity optimization . . . . .	14
2.2	A CDK2/CDK9 selectivity dataset from Shao et <i>al.</i> , 2013 . . . . .	17
2.3	CDK2 and ERK2 selectivity dataset from Blake et <i>al.</i> , 2016 . . . . .	20
2.4	Relative free energy calculations can accurately predict potency, but show larger errors for selectivity predictions. . . . .	22
2.5	Correlation in selectivity prediction errors can be used to accelerate selectivity optimization . . . . .	25
2.6	CDK2 adopts an inactive conformation in the crystal structure used for the CDK2/ERK2 calculations . . . . .	28
2.7	Correlation coefficient $\rho$ controls the shape of the joint marginal dis- tribution of errors . . . . .	29
2.8	Correlation reduces the expected error for selectivity predictions . . .	30
4.1	Relative alchemical free-energy calculations can be used to predict affinity changes of FDA-approved selective kinase inhibitors arising from clinically-identified mutations in their targets of therapy. . . . .	34
4.2	Cross-comparison of the experimentally measured effects that mu- tations in Abl kinase have on ligand binding, performed by different labs. . . . .	41

CHAPTER 1  
**INTRODUCTION**

## CHAPTER 2

# PREDICTING THE SELECTIVITY OF SMALL MOLECULE KINASE INHIBITORS

## 2.1 Introduction

With the FDA approval of imatinib in 2001, targeted small molecule kinase inhibitors (SMKIs) have become a major class of therapeutics in treating cancer and other diseases. Currently, there are 43[2] approved SMKIs, and it is estimated that kinase targeted therapies account for as much as 50% of current drug development [3], with many more compounds currently in clinical trials. While there have been a number of successes, design of selective kinase inhibitors remains a challenge. Selectivity design is particularly difficult for kinases: there are more than 518 protein kinases [4, 5] with a highly conserved ATP binding site that is targeted by the majority of small molecule kinase inhibitors [6]. While these inhibitors do adopt several different binding modes [7–10], previous work has shown that both Type I (binding to the active, DFG-in conformation) and Type II (binding to the inactive, DFG-out conformation) inhibitors display a wide range of selectivities [11, 12], often binding a number of other targets unselectively in addition to a primary target. Even inhibitors approved by the FDA, often the result of extensive drug development programs, can bind to a large number of off-target kinases [13].

### 2.1.1 Selectivity is an important consideration in drug design

Selectivity is an important property to consider in drug development, either in the pursuit of a maximally selective inhibitor [14, 15] or in pursuit a polypharma-

cological agent [16–20], to avoid on-target toxicity (arising from inhibition of the intended target) [21] and off-target toxicity (arising from inhibition of unintended targets) [22, 23]. In either paradigm, considering the selectivity of a compound is complicated by the biology of kinases, which exist as nodes in complex signaling networks [24, 25] with feedback inhibition and cross-talk between pathways. Careful consideration of which kinase off-targets are being inhibited can avoid off-target toxicity due to alleviating feedback inhibition and inadvertently reactivating the targeted pathway [24, 25], or the upregulation of a secondary pathway by alleviation of cross-talk inhibition [26, 27]. Off-target toxicity can also be caused by inhibition outside of the kinome, such as gefitinib inhibiting CYP2D6 [22] and causing hepatotoxicity in lung cancer patients. In a cancer setting, on-target toxicity can be avoided by considering the selectivity of the SMKI for the oncogenic mutant form of the kinase over the wild type form of the kinase [28–30], demonstrated by number of first generation EGFR inhibitors. Selectivity considerations can also lead to beneficial effects: Imatinib, initially developed to target BCR-Abl fusion proteins, is also approved for treating gastrointestinal stromal tumors (GIST) [31] due to its activity against receptor tyrosine kinase KIT.

### **2.1.2 Physical modeling can compliment high throughput screens and machine learning to predict selectivity**

High-throughput screening platforms [12, 13, 32–35] are able to produce a wealth of experimental data, covering many SMKIs and a large percentage of the human kinome. This affinity data can enable various machine learning approaches geared towards predicting selectivity [36–39]. However, such screening platforms can be expensive to run for a development program considering a large number of

molecules, and requires the synthesis of compounds to be tested. Machine learning algorithms can be hampered by the unreliability of publicly available affinity data [40, 41] or by poor coverage of a novel chemotype in the model’s training set [32]. While there have been efforts to improve the experimental tractability for kinases [42–45] and generate high-quality data in a less high-throughput manner [8, 46], these methodologies are limited by a number of experimental concerns, such as ligand synthesis, ligand solubility and difficulty in expressing the proteins of interest without extensive tailoring of expression protocols.

### **2.1.3 Alchemical free energy methods can accurately predict single target potencies and the impact of mutations**

The selectivity of imatinib for Abl kinase over Src [47, 48] and within a family of non-receptor tyrosine kinases [49] has been studied extensively using molecular dynamics and free energy calculations. Advances in atomistic molecular mechanics forcefields [50–53] have reached a level of accuracy sufficient for predicting ligand potencies. Alchemical free energy calculations allow for prediction of ligand binding free energies, including all enthalpic and entropic contributions [54]. In kinases, free energy calculations have previously been shown to achieve mean unsigned errors of less than 1.0 kcal/mol on a number of systems [55–57] when predicting single target potencies. Beyond kinases, alchemical free energy calculations achieved a mean unsigned error of 0.6 kcal/mol in predicting BRD4 inhibitor potencies [58]. Protein mutation free energy calculations have also showed promise in predicting the  $\Delta\Delta G$  of mutation for a benchmark set of Abl mutations and FDA-approved kinase inhibitors with an accuracy of about 1 kcal/mol [59]. While the ability to predict ligand potencies and, to a lesser extent, the impact of mutations

has been studied previously, not much work has been done on assessing free energy methodology for predicting selectivity. Work on predicting the selectivity of three bromodomain inhibitors for a number of different bromodomains achieved promising accuracy for single target potencies of roughly 1 kcal/mol, but does not evaluate into the accuracy of  $\Delta\Delta G_{selectivity}$  predictions in depth [60].

#### **2.1.4 Assessing the ability of alchemical free energy methods to predict selectivity**

While we expect that the errors in both experiment and calculation will be larger when considering two targets instead of one, as seen in the Hauser paper, we were interested to see to what degree these errors in predicted binding free energies on the two targets were correlated and to what extent the current alchemical free energy methods can predict selectivity. Due to the narrow dynamic range of selectivities and larger experimental errors from the propagation of errors for multiple targets, we anticipate difficulty in predicting selectivity if the errors in the calculated free energies on the two targets are largely uncorrelated, or even anticorrelated. However, some level of correlation in the errors of the calculated free energies on the two targets could lead to a fortuitous cancellation of errors in predicting the selectivity between targets, enabling more accurate than expected calculations. Here, we investigate the utility of alchemical free energy calculations for the prediction of selectivity, hereafter taken to mean the  $\Delta\Delta G$  in binding free energies of the same compound for two targets. We employed state of the art relative free energy calculations [55, 61] to predict the selectivities of two different congeneric ligand series [62, 63], as well as present a simple numerical model to quantify the expected speed up in selectivity optimization expected for different combinations



of per target errors and correlation coefficient values. We built Bayesian graphical models of the relative free energy calculations to quantify the correlation of errors in our example calculations.

Insert summary  
of results

## 2.2 Methods

### 2.2.1 Numerical model of selectivity

To model the impact correlation would have on the expected uncertainty for selectivity predictions,  $\sigma_{selectivity}$  was calculated using Equation 2.1 for 1000 evenly spaced values of the correlation coefficient ( $\rho$ ) from 0 to 1, for a number of combinations of per target errors ( $\sigma_{target1}$  and  $\sigma_{target2}$ )

$$\sigma_{selectivity} = \sqrt{\sigma_{target1}^2 + \sigma_{target2}^2 - 2\rho\sigma_{target1}\sigma_{target2}} \quad (2.1)$$

The speed up in selectivity optimization that could be expected from using free energy calculations of a particular per target error ( $\sigma_{selectivity}$ ) was quantify as follows using NumPy (v 1.14.2). An original, true distribution for the changed in selectivity of 200000000 new compounds proposed with respected to a reference compound was modeled as a normal distribution centered around 0 with a standard deviation of 1 kcal/mol. This assumption was made on the basis that the majority of selectivity is driven by the scaffold, and R group modifications will do little to drive changes in selectivity. The 1 kcal/mol distribution is supported by the standard deviations of the selectivity in the experimental datasets referenced in this work, which are all less than, but close, to 1 kcal/mol.

Each of these proposed compounds were "screened" by a free energy calculation technique with a per target error ( $\sigma_{target}$ ) of 1 kcal/mol [55] and a specified correlation coefficient  $\rho$ . A  $\sigma_{selectivity}$  was calculated according to Equation 2.1. The noise of the computational method was modeled as a normal distribution centered around 0 with a standard deviation of  $\sigma_{selectivity}$  and added to the "true" change in selectivity. Any compound predicted to have an improvement in selectivity of 1.4 kcal/mol (1 log unit) would then be made and have its selectivity experimental measured. The speedup value for each value of  $\rho$  is calculated as the proportion of compounds made with a true selectivity gain of 1.4 kcal/mol divided by the proportion of compounds with a 1.4 kcal/mol improvement in the original distribution, where all of the compounds were made.

Finally, this process was repeated for a 100x (2.8 kcal/mol, 2 log unit) selectivity optimization and 50 linearly spaced values of the correlation coefficient ( $\rho$ ) between 0 and 1, for four values of  $\sigma_{selectivity}$  and 40000000 compounds in the original distribution.

### 2.2.2 Structure Preparation

Structures from the Shao [62] and Hole [64], and Blake [63] papers were downloaded from the PDB [65], selecting structures with the same co-ligand crystallized. For the Shao dataset, 4BCK (CDK2) and 4BCI (CDK9) were selected, which have ligand 12c cocrystallized. For the Blake dataset, 5K4J (CDK2) and 5K4I (ERK2) were selected, cocrystallized with ligand 21. The structures were prepared using Schrodinger's Protein Preparation Wizard [66] (release 2017-3). This pipeline modeled in internal loops and missing atoms, added hydrogens at the reported experimental pH (7.0 for the Shao dataset, 7.3 for the Blake dataset) for both the

protein and the ligand. All crystal waters were retained. The ligand was assigned protonation and tautomer states using Epik at the experimental  $\text{pH} \pm 2$ , and hydrogen bonding was optimized using PROPKA at the experimental  $\text{pH} \pm 2$ . Finally, the entire structure was minimized using OPLS3 with an RMSD cutoff of 0.3.

### 2.2.3 Ligand Pose Generation

Ligands were extracted from the publication entries in the BindingDB as 2D SMILES strings. 3D conformations were generated using LigPrep with OPLS3 [67]. Ionization state was assigned using Epik at experimental  $\text{pH} \pm 2$ . Stereoisomers were computed by retaining any specified chiralities and varying the rest. The tautomer and ionization state with the lowest epik state penalty was selected for use in the calculation. Ligand poses were generated by first aligning to the co-crystal ligand using the Largest Common Bemis-Murcko scaffold with fuzzy matching (Schrodinger 2017-4). Ligands that were poorly aligned or failed to align were then aligned using Maximum Common Substructure (MCSS). Finally, large R-groups were allowed to sample different conformations using MM-GBSA with a common core restrained. VSGB solvation model was used with the OPLS3 forcefield. No flexible residues were defined for the ligand.

### 2.2.4 Free Energy Calculations

The FEP+ panel (Maestro release 2017-4) was used to generate perturbation maps. Neutral perturbations were run for 15ns per replica, using an NPT ensemble and water buffer size of 5. A GCMC solvation protocol was used to sample buried water molecules in the binding pocket prior to the calculation, which discards any

retained crystal waters.

### 2.2.5 Charge Change Free Energy Calculations

For ligands where a protonation state change was expected to be relevant to binding based on a small state penalty, Jaguar pKa prediction calculations [68] were run to identify protonation state changes with pKas within 1 log unit of the experimental pH. The predicted pKas for one ligand (Shao 12b, 7.84) was within this range. To account for this, a pKa correction was performed. For this ligand, a separate perturbation map containing ligands 12a, 12c, 12b (neutral) and 12b (charged) was run for 30ns per replica using a post-calculation Coulombic charge correction. **Add in SALT concentration used for these.** The pKa correction was performed using Equation 2.2:

$$\Delta\Delta G_{corrected} = \Delta\Delta G_{uncorrected} - RT \log \left( \frac{10^{pK_a - pH} + 1}{e^{\frac{\Delta G_{neutral} - \Delta G_{charged}}{RT}} * (10^{pK_a - pH} + 1)} \right) \quad (2.2)$$

$\Delta\Delta G$  for each edge in perturbation map with 12a, 12c and 12b (neutral) was updated using the correction above and merged into the final map.

### 2.2.6 Statistical Analysis of FEP+ calculations

Each FEP+ calculation has a reported mean unsigned error (MUE) and root mean squared error (RMSE) with a bootstrapped 95% confidence interval. The MUE was calculated according Equation 2.3, while the RMSE was calculated according to Equation 2.4.

$$MUE = \frac{\sum_0^n |\Delta G_{calc} - \Delta G_{exp}|}{n} \quad (2.3)$$

$$RMSE = \frac{\sum_0^n \sqrt{\Delta G_{calc}^2 - \Delta G_{exp}^2}}{n} \quad (2.4)$$

Each RMSE and MUE is reported with a 95% confidence interval calculated from 10000 replicates of a choose-one-replace bootstrap protocol on the  $\Delta G$  values reported to account for the finite sample size of the ligands. The code used to bootstrap these values is available on github: <https://github.com/choderalab/selectivity>

### 2.2.7 Quantification of the correlation coefficient $\rho$

To quantify  $\rho$ , we built a Bayesian graphical model using pymc3 (v. 3.5) [69] and theano (v 1.0.3) [70], which has been made available on Github. For each phase (complex and solvent), the absolute free energy ( $G$ ) of ligand  $i$  was treated as a normal distribution (Equation 2.5). For each set of calculations, one ligand was chosen as the reference, and pinned to 0, with a standard deviation of 1 kcal/mol in order to improve the efficiency of sampling from the model.

$$G_{i,target}^{phase} = \mathcal{N}(\mu = 0, sd = 25.0 \text{ kcal/mol}) \quad (2.5)$$

For each edge of the FEP map (ligand  $i$  – ligand  $j$ ), there is a contribution from dummy atoms, that was modeled as in Equation 2.6.

$$c_{i,j} = \mathcal{N}(\mu = 0, sd = 25.0 \text{ kcal/mol}) \quad (2.6)$$

The model was restrained by including data from the FEP+ calculation.

$$\Delta G_{phase, ij, target}^{BAR} = \mathcal{N}(G_{j, target}^{phase} - G_{i, target}^{phase}, \delta^2 \Delta G_{phase, ij, target}^{BAR}, observed = \Delta G_{phase, ij, target}^{calc}) \quad (2.7)$$

Where  $\delta^2 \Delta G_{phase, ij, target}^{BAR}$  is the reported BAR uncertainty from the calculation, and  $\Delta G_{phase, ij, target}^{calc}$  is the BAR estimate of the free energy for the perturbation between ligands  $i$  and  $j$  in a given phase.

From this, we can calculate the  $\Delta \Delta G^{FEP}$  for each edge as in Equation 2.8:

$$\Delta \Delta G_{target, ij}^{FEP} = \Delta G_{complex, ij, target}^{BAR} - \Delta G_{solvent, ij, target}^{BAR} \quad (2.8)$$

To model the way an offset is calculated for the  $\Delta G$  reported by the FEP+ panel in Maestro:

$$offset = \frac{\sum^n G_{i, target}^{complex} - G_{i, target}^{solvent}}{n} - \frac{\sum^n \Delta G_i^{exp}}{n} \quad (2.9)$$

The offset was added to each  $\Delta G_i^{BAR}$  to calculate  $\Delta G_i^{sch}$ .

The experimental binding affinity was treated as a true value ( $\Delta G_{i, target}^{true}$ ) corrupted by experimental uncertainty, which is assumed to be 0.3 kcal/mol [41], with the values reported in the papers ( $\Delta G_{i, target}^{obs}$ ) treated as observations from this distribution (Equation 2.10)

$$\Delta G_{i, target}^{exp} = \mathcal{N}(\Delta G_{i, target}^{true}, 0.3 \text{ kcal/mol}, observed = \Delta G_{i, target}^{obs}) \quad (2.10)$$

$\Delta G_{i,target}^{true}$  was assigned a weak normal prior, as in equation 2.11.

$$\Delta G_{i,target}^{true} = \mathcal{N}(0, 50 \text{ kcal/mol}) \quad (2.11)$$

The error for a given ligand was calculated as in Equation 2.12.

$$\epsilon_i = \Delta G_i^{sch} - \Delta G_i^{true} \quad (2.12)$$

From these  $\epsilon$  values, we calculated the correlation coefficient,  $\rho$  as in Equation 2.13.

$$\rho = \frac{cov(\epsilon_{target1}, \epsilon_{target2})}{\sigma_{target1}\sigma_{target2}} \quad (2.13)$$

Where  $\sigma$  is the standard deviation of  $\epsilon$ . To quantify  $\rho$  for the CDK2/ERK2 calculations, the default NUTS sampler with jitter+adapt\_diag initialization, 1000 tuning steps, and a target accept probability of 0.8 was used to draw 10000 samples from the model. The CDK2/CDK9 model was sampled 20000 times using default NUTS sampler with jitter+adapt\_diag initialization and 3000 tuning steps.

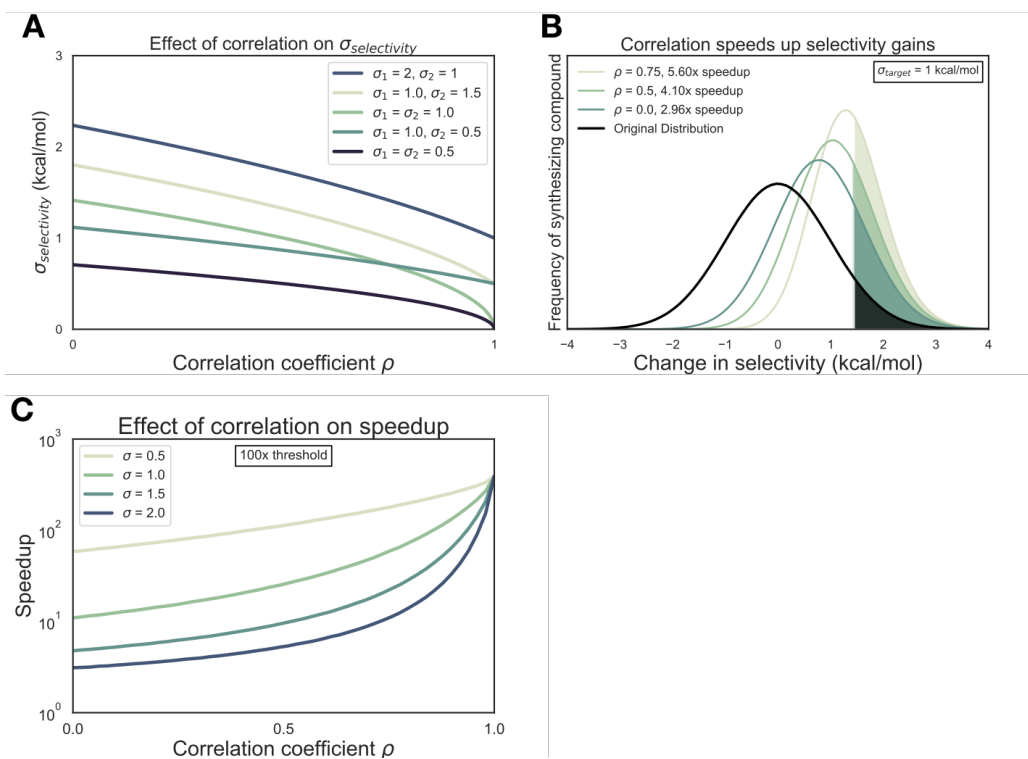
## 2.3 Results

### 2.3.1 Correlation of errors can make selectivity predictions more accurate and speed up ligand optimization

To demonstrate the potential impact correlation has on the uncertainty of selectivity predictions ( $\sigma_{selectivity}$ ) using alchemical free energy techniques, we created

a simple numerical model following equation 2.1, which takes into account each of the per target errors expected from the methodology as well as the correlation in those errors. As seen in Figure 2.1A, if the per target errors ( $\sigma_1$  and  $\sigma_2$ ) are the same,  $\sigma_{selectivity}$  approaches 0 as the correlation coefficient ( $\rho$ ) approaches 1. If the error for the free energy method is not the same,  $\sigma_{selectivity}$  gets smaller but approaches a non-zero value as  $\rho$  approaches 1. To quantify the expected speedup in selectivity optimization, we modeled the change in selectivity with respect to a reference compound for a number of compounds a medicinal chemist might suggest as a normal distribution centered around 0 with a standard deviation of 1 kcal/mol (Figure 2.1B, black curve), reflecting that most proposed changes would not drive large changes in selectivity. Then, suppose that each compound is screened computationally with a method free energy methodology with a per target ( $\sigma_{target}$ ) error of 1 kcal/mol, and all compounds predicted to have a 1.4 kcal/mol improvement in selectivity are synthesized and experimentally tested (Figure 2.1B, colored curves). The fold-change in the proportion of compounds that are made that have a true 1.4 kcal/mol improvement in selectivity compared to the original distribution can be calculated as a surrogate for the expected speedup. For a 1.4 kcal/mol selectivity improvement threshold (1 log unit), a correlation of 0.5 gives an expected speed up of 4.1x, which can be interpreted as 4.1x fewer compounds needing to be made before achieving a 1 log unit improvement in selectivity. This process can be extended for the even more difficult proposition of achieving a 2 log unit improvement in selectivity (Figure 2.1C), where 200-300x speedups can be expected, depending on  $\sigma_{target}$  for the free energy methodology.





**Figure 2.1: Free energy calculations speed up selectivity optimization** (A) The effect of correlation on expected errors for predicting selectivity ( $\sigma_{selectivity}$ ) in kcal/mol. Each curve represents a different combination of target errors ( $\sigma_1$  and  $\sigma_2$ ). (B) The change in selectivity for molecules proposed by medicinal chemists optimizing a lead candidate can be modeled by a normal distribution centered on 0 with a standard deviation of 1 kcal/mol (black curve). Each green curve corresponds to the distribution of compounds made after screening for a 1 log unit (1.4 kcal/mol) improvement in selectivity with a free energy methodology with a 1 kcal/mol per target error and a particular correlation. The shade region of each curve corresponds to the compounds with a real 1 log unit improvement in selectivity. The speed up is calculated as the ratio of the percentage of compounds made with a real 1 log unit improvement to the percentage of compounds that would be expected in the original distribution. (C) The speedup (y-axis, log scale) expected for 100x (2 log units, 2.8 kcal/mol) selectivity optimization as a function of correlation coefficient  $\rho$ . Each curve corresponds to a different  $\sigma_{target}$  value.

### 2.3.2 The CDK2 and CDK9 experimental dataset demonstrates the difficulty in achieving selectivity for closely related kinases

To begin quantifying the correlation of errors in free energy predictions for selectivity, we set out to gather datasets that met a number of criteria. We looked for datasets that contained binding affinity data for a number of kinase targets and ligands, as well as having crystal structures for each target with the same cocrystallized ligand. For the CDK2/CDK9 dataset [62], ligand 12c was cocrystallized with CDK2/cyclin A (Figure 2.2A, left) and CDK9/cyclin T (Figure 2.2B, left), work that was published in a companion paper [64]. In both CDK2 and CDK9, ligand 12c forms relatively few hydrogen bond interactions with the kinase. Each kinase forms a set of hydrogen bonds between the ligand scaffold and a hinge residue (C106 in CDK9 and L83 in CDK2) that is conserved across all of the ligands in this series. CDK9, which has slightly lower affinity for ligand 12c (Figure 2.2C, right), forms a lone interaction between the sulfonamide of ligand 12c and residue E107. On the other hand, CDK2 forms interactions between the sulfonamide of ligand 12c and residues K89 and H84. The congeneric series of ligands contains a number of challenging perturbations, particularly at substituent point R3 (Figure 2.2C, left). Ligand 12i also presented a challenging perturbation, moving the 1-(piperazine-1-yl)ethanone from the *meta* to *para* location.

This congeneric series of ligands also highlights two of the challenges of working from publicly available data. First, the dynamic range of selectivity is incredibly narrow, with a mean  $\Delta\Delta G_{selectivity}$  (CDK9 - CDK2) of only -0.65 kcal/mol, and a standard deviation of 0.88 kcal/mol. Additionally, experimental uncertainties are not reported for the experimental measurements. Thus, for this and subsequent

sets of ligands, the experimental uncertainty is assumed to be 0.3 kcal/mol based on previous work done to summarize uncertainty in experimental data [41, 59].

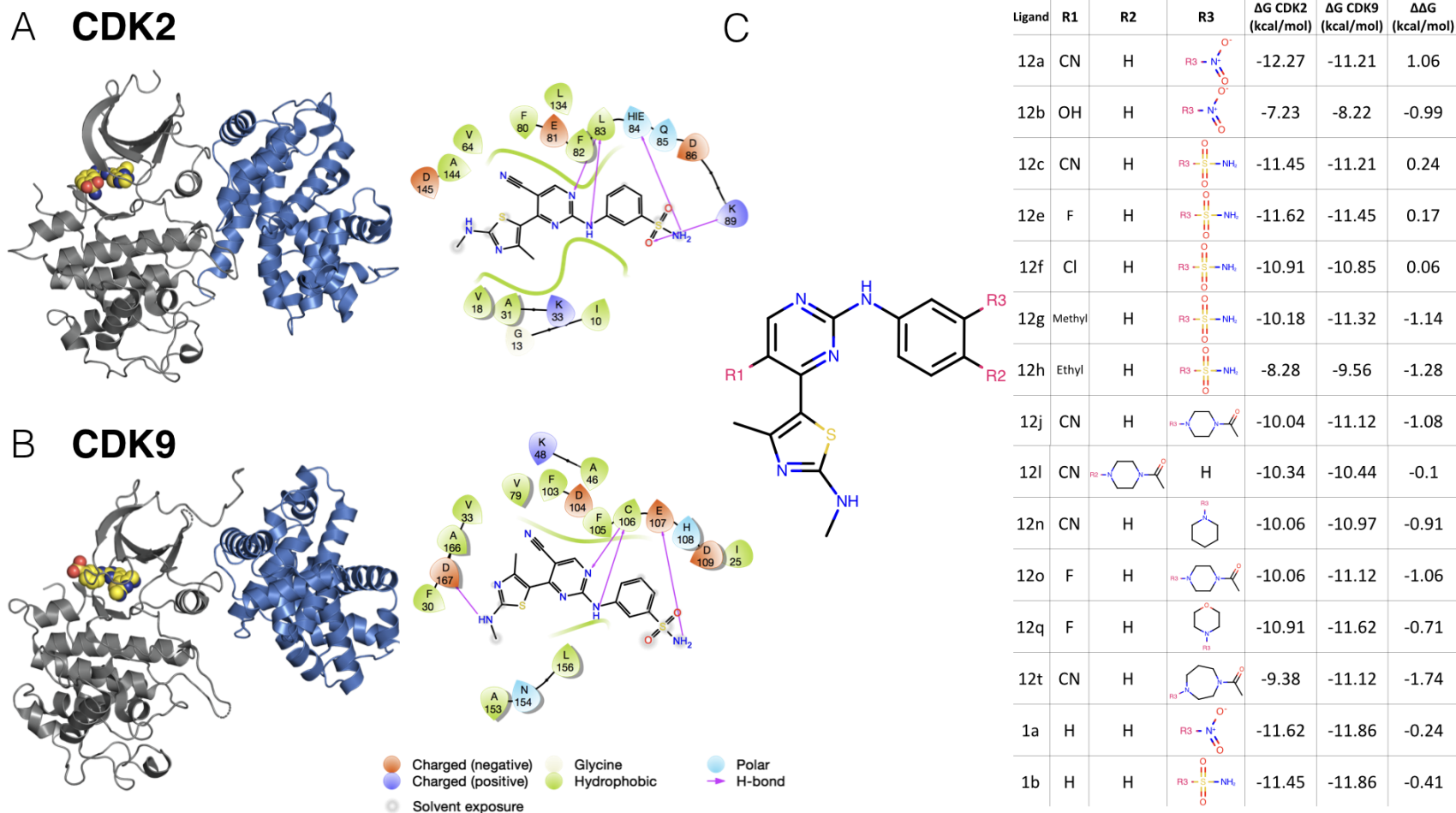


Figure 2.2: **A CDK2/CDK9 selectivity dataset from Shao et al., 2013** (A) (left) Crystal Structure (4BCK)[64] of CDK2 (gray ribbon) bound to ligand 12c (yellow spheres). Cyclin A is shown in blue ribbon (right) 2D ligand interaction map of ligand 12c in the CDK2 binding site. (B) (left) Crystal structure of CDK9 (4BCI)[64] (gray ribbon) bound to ligand 12c (yellow spheres). Cyclin T is shown in blue ribbon. (right) 2D ligand interaction map of ligand 12c in the CDK9 binding site. (C) (left) 2D structure of the common scaffold for all ligands in congeneric ligand series 12 from the publication (right) A table summarizing all R group substitutions as well as the published experimental binding affinities and selectivities[62].

### 2.3.3 The CDK2 and ERK2 dataset achieves higher levels of selectivity for more distantly related kinases

The CDK2/ERK2 dataset from Blake *et al.*, 2016 also met the criteria described above. Crystal structures for both CDK2 (Figure 2.3A, top) and ERK2 (Figure 2.3B, top) were available with ligand 22 co-crystallized. Of note, CDK2 was not crystallized with cyclin A, despite cyclin A being included in the affinity assay reported in the paper [63]. CDK2 adopts a DFG-in conformation with the  $\alpha$ C helix rotated out, away from the ATP binding site and breaking the conserved salt bridge between K33 and E51 (Supp. Figure 2.6A), indicative of an inactive kinase [9, 71]. By comparison, the CDK2 structure from the CDK2/CDK9 dataset adopts a DFG-in conformation with the  $\alpha$ C helix rotated in, forming the ionic bond between K33 and E51 indicative of an active kinase, due to allosteric activation by cyclin A. While missing cyclins have caused problems for free energy calculations in prior work, it is possible that the fully active conformation contributes equally to binding affinity for all of the ligands in the series, and the high accuracy of the potency predictions (Figure 2.4, top left) is the result of fortuitous cancellation of errors. The binding mode for this series is similar between both kinases. There is a set of conserved hydrogen bonds between the scaffold of the ligand and the backbone of one of the hinge residues (L83 for CDK2 and M108 for ERK2). The conserved lysine (K33 for CDK2 and K54 for ERK2), normally involved in the formation of a ionic bond with the  $\alpha$ C helix, forms a hydrogen bond with the scaffold (Figure 2.4A and 2.4B, bottom) in both CDK2 and ERK2. However, in the ERK2 structure, the hydroxyl engages a crystallographic water as well as N154 in a hydrogen bond network that is not present in the CDK2 structure. The congeneric ligand series features a single substituent point, with the R groups exposed to the solvent. This helps explain the extremely narrow distribution of selectivities, with a mean selectivity of -1.74

I probably need to either remove the reference to how closely related the kinases are, or come up with a way to quantify that - SKA

is there a good citation for this?

kcal/mol (ERK2 - CDK2) and standard deviation of 0.56 kcal/mol. This suggests that the selectivity is largely driven by the scaffold and unaffected by the R group substitutions.

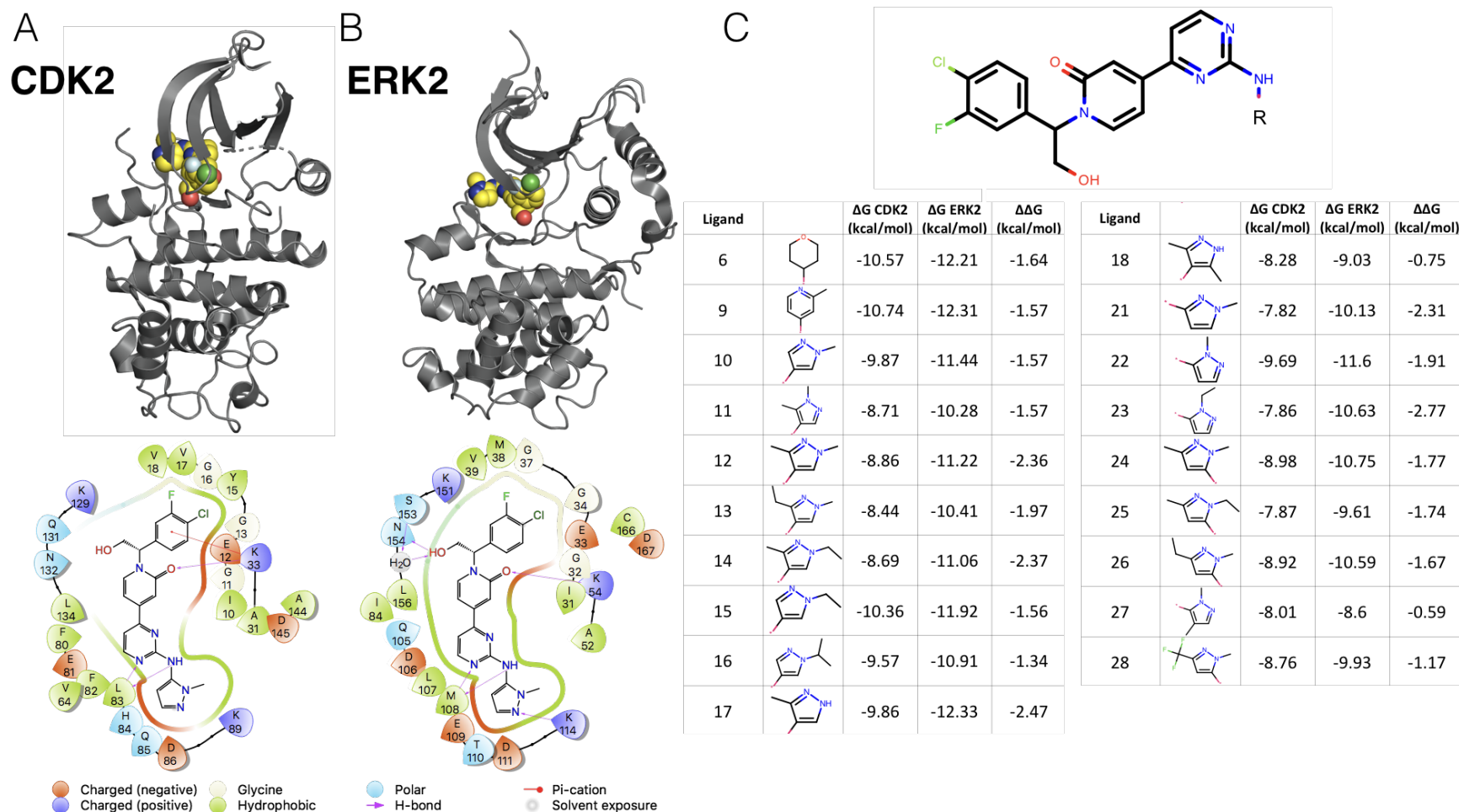


Figure 2.3: CDK2 and ERK2 selectivity dataset from Blake *et al.*, 2016

(A) (*top*) Crystal structure of CDK2 (5K4J) shown in gray cartoon and ligand 22 shown in yellow spheres. (*bot*) 2D interaction map of ligand 22 in the binding pocket of CDK2 (B) (*top*) Crystal structure of ERK2 (5K4I) shown in gray cartoon with ligand 22 shown in yellow spheres. (*bot*) 2D interaction map of ligand 22 in the binding pocket of ERK2. (C) (*top*) Common scaffold for all of the ligands in the Blake dataset, with R denoting attachment side for substitutions. (*bot*) Table showing R group substitutions and experimentally measured binding affinities and selectivities. Ligand numbers correspond to those used in publication.

### 2.3.4 FEP+ calculations show accurate potency predictions for ERK2/CDK2 and larger errors for CDK2/CDK9

The FEP+ predictions of single target potencies ( $\Delta G$ ) showed good accuracy for the CDK2 and ERK2 dataset (Figure 2.4, top), with an RMSE of  $0.37^{0.57}_{0.16}$  and  $0.53^{0.81}_{0.22}$  kcal/mol, respectively. All of the CDK2 and ERK2 potencies were predicted within 1 log unit of the experimental value. Despite the high accuracy for the single target potency predictions, the selectivity ( $\Delta\Delta G_{selectivity}$ ) predictions show an RMSE of  $0.81^{1.26}_{0.37}$  kcal/mol, with all of the predictions falling within 1 log unit of the experimental values (Figure 2.4, top right panel). Despite the high accuracy of the predictions, the narrow dynamic range and high uncertainty from experiment and calculation obscures any signal in the data. The CDK2 and CDK9 datasets show higher errors in the potency predictions, with an RMSE of  $1.39^{2.05}_{0.58}$  and  $1.71^{2.61}_{0.61}$  kcal/mol respectively. There are a number of outliers that fall outside of 1 log unit from the experimental value. While the higher per target errors make predicting potency more difficult, the selectivity predictions show a much lower RMSE of  $0.74^{1.25}_{0.31}$  kcal/mol. This suggests that some correlation in the error is leading to fortuitous cancellation of systematic error, leading to more accurate than expected predictions of  $\Delta\Delta G_{selectivity}$ .

this will change  
if we take of the  
ff uncertainty  
error bars



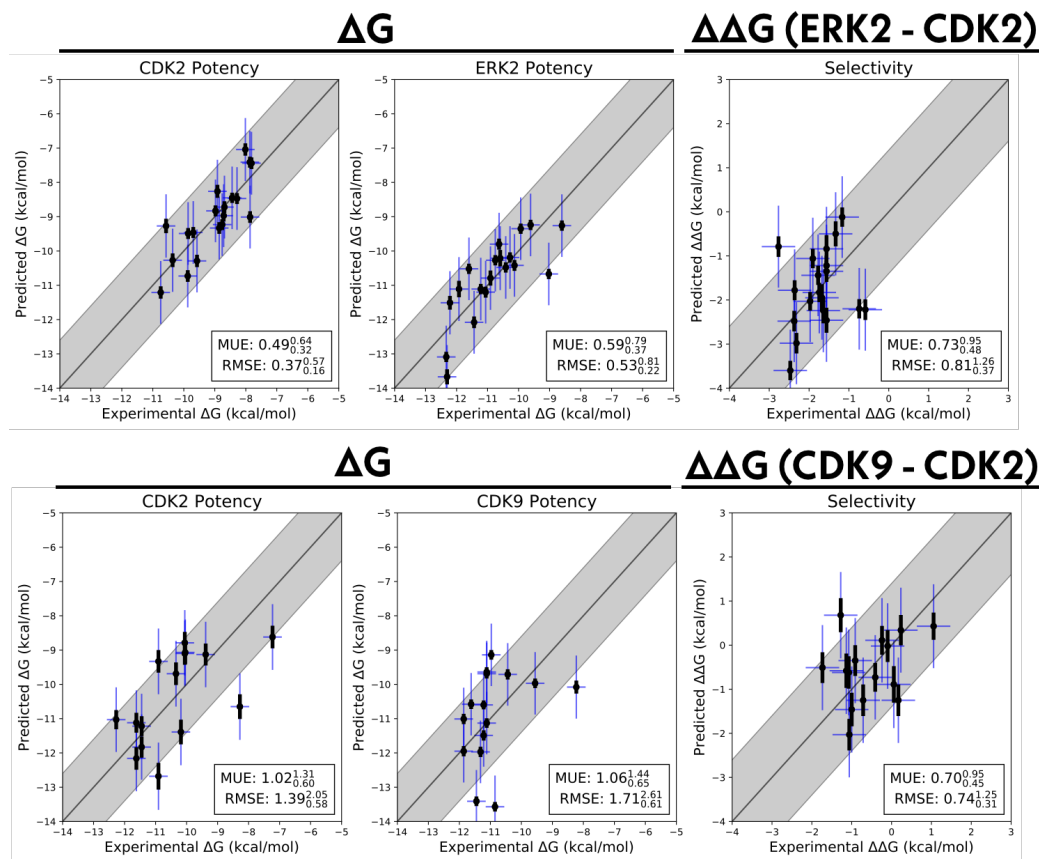


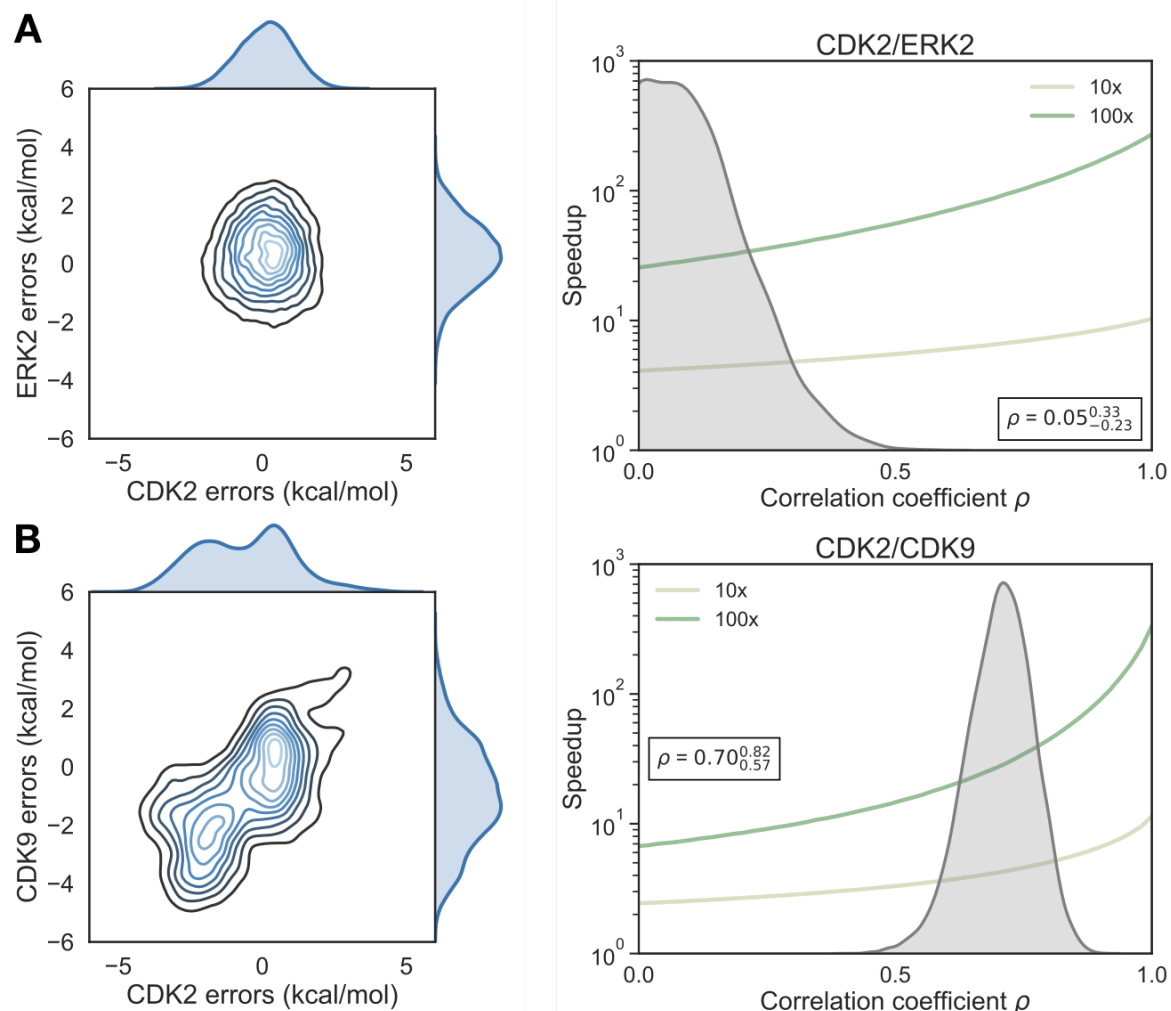
Figure 2.4: **Relative free energy calculations can accurately predict potency, but show larger errors for selectivity predictions.**

Single target potencies and selectivities for CDK2/ERK2 from the Blake datasets (*top*), and CDK2/CDK9 (*bottom*) from the Shao datasets. The experimental values are shown on the X-axis and calculated values on the Y-axis. Each data point corresponds to a ligand for a given target. All values are shown in units of kcal/mol. The horizontal error bars show the assumed experimental uncertainty of 0.3 kcal/mol[41]. To better highlight outliers that are unlikely due simply to forcefield errors, we presume the forcefield error ( $\sigma_{\text{FF}} \approx 0.9$  kcal mol<sup>-1</sup> [52]) also behaves as a random error. We show the total estimated statistical and forcefield error ( $\sqrt{\sigma_{\text{FF}}^2 + \sigma_{\text{calc}}^2}$ ) as vertical blue error bars. The black vertical error bars correspond to the statistical error ( $\sigma_{\text{calc}}$ ). The black line indicates agreement between calculation and experiment, while the gray shaded region represent 1.36 kcal/mol (or 1 log unit) error. The MUE and RMSE are shown on each plot with bootstrapped 95% confidence intervals.

### 2.3.5 Free energy calculation errors are correlated, accelerating selectivity optimization

To quantify the correlation coefficient ( $\rho$ ) of the errors in our calculations, we built a Bayesian graphical model, as described in the methods section. Briefly, we modeled the absolute free energy ( $G$ ) of each ligand in each phase (complex and solvent) as in equation 2.5. The model was chained to the FEP+ calculations by providing the  $\Delta G_{phase,i,j,target}^{calc}$  as observed data, as in equation 2.7. As in equation, the experimental data was modeled as a normal distribution centered around the true free energy of binding ( $\Delta G_{i,target}^{true}$ ) corrupted by experimental error, which is assumed to be 0.3 kcal/mol from previous work done to quantify the uncertainty in publicly available data [41]. The reported IC50 values from each dataset were treated as data observations (Equation 2.10) and the  $\Delta G_{i,target}^{true}$  was assigned a weak normal prior (Equation 2.11). The correlation coefficient was calculated for each sample according to equation 2.12. The correlation coefficient  $\rho$  for the CDK2/ERK2 calculations was quantified to be  $0.5^{0.33}_{-0.23}$ , indicating that the errors are largely uncorrelated between ERK2 and CDK2 (Figure 2.5A, right). The joint marginal distribution of the error ( $\epsilon$ ) for each target is symmetric, which is expected for cases in which  $\rho$  is 0 (Supp. Figure 2.7). Despite the weak correlation in errors, the high per target accuracy of these calculations should have a 2-3x speed up for 1 log unit selectivity optimization, and a 20-30x speed up for 2 log unit selectivity optimization (Figure 2.5A, right). The CDK2/CDK9 calculations show strong evidence of correlation, with a correlation coefficient of  $0.70^{0.82}_{0.57}$  (Figure 2.5B, right). The joint marginal distribution of errors is strongly diagonal, which is expected based on the value for  $\rho$  (Figure 2.5B, left). The high correlation in errors leads to a speed up of 4-5 for 1 log unit selectivity optimization and 30-40x for 2 log unit selectivity optimization (Figure 2.5B, right), despite the much higher per target

errors. Quantifying  $\rho$  for these calculations enables estimation of  $\sigma_{selectivity}$ , which is useful for estimating expected error for prospective studies, where the experimental values for  $\Delta\Delta G_{selectivity}$  are not yet known. Based on the distribution quantified for  $\rho$ , the expected  $\sigma_{selectivity}$  for the CDK2/CDK9 calculations is between 0.76 and 1.16 kcal/mol (Supp. Figure 2.8), which is in good agreement with the bootstrapped RMSE (Figure 2.4, bottom). For the CDK2/ERK2 calculations,  $\sigma_{selectivity}$  is expected to fall between 0.82 kcal/mol and 1.10 kcal/mol (Supp. Figure 2.8), which is also in good agreement with the bootstrapped RMSE (Figure 2.4, top).



**Figure 2.5: Correlation in selectivity prediction errors can be used to accelerate selectivity optimization**

(A) (left) The joint posterior distribution of the prediction errors for CDK2 (X-axis) and ERK2 (Y-axis) from the Bayesian graphical model. (right) Speedup in selectivity optimization (Y-axis) as a function of correlation coefficient (X-axis). The posterior marginal distribution of the correlation coefficient ( $\rho$ ) is shown in gray, while the expected speed up is shown for 100x (green curve) and 10x (yellow curve) selectivity optimization. The inserted box shows the mean and 95% confidence interval for the correlation coefficient. (B) (left) The same as above, with CDK2 (X-axis) and CDK9 (Y-axis). (right) As above, for the CDK2/CDK9 calculations.

## 2.4 Discussion and Conclusions

We have demonstrated, using a simple numerical model, the impact that free energy calculations with even weakly correlated errors can have on speeding up the optimization of selectivity in small molecule kinase inhibitors. While the expected speed up is dependent on the per target error of the method ( $\sigma_{target}$ ), the speedup is also highly dependent on the correlation of errors made for both targets. Unsurprisingly, free energy methods have greater impact as the threshold for selectivity optimization goes from 10x to 100x. While 100x selectivity optimization is difficult to achieve, the expected benefit from free energy calculations is also quite high, with 1 and 2 order of magnitude speedups possible. To quantify the correlation of errors in two example systems, we gathered experimental data for two congeneric ligand series with experimental data for CDK2 and ERK2, as well as CDK2 and CDK9. These datasets, which had crystal structures for both targets with the same ligand co-crystallized, exemplify the difficulty in predicting selectivity. The dynamic range of selectivity for both systems is incredibly narrow, with most of the perturbations not having a major impact on the overall selectivity achieved. Further, the data was reported with unreliable experimental uncertainties, which makes quantifying the errors made by the free energy calculations difficult. This issue is common when considering selectivity, as many kinase-oriented high throughput screens are carried out at a single concentration and not highly quantitative. Work is being done to increase the availability of publicly available, high quality biophysical data for selectivity measurements, which will benefit future work on predicting selectivity using physical models and machine learning techniques.

insert a discussion of outliers in the FEP+ calculations here?

discuss quantification of  $\rho$  (is there a way to know a priori what  $\rho$  might be?)

discussion of the expected speedup and sigma based on the quantification of  $\rho$

extensions include: separating statistical error (unless this gets added into the results section

Do we want to mention protein mutation FEP+ to enable cycle closure analysis for certain highly related targets?

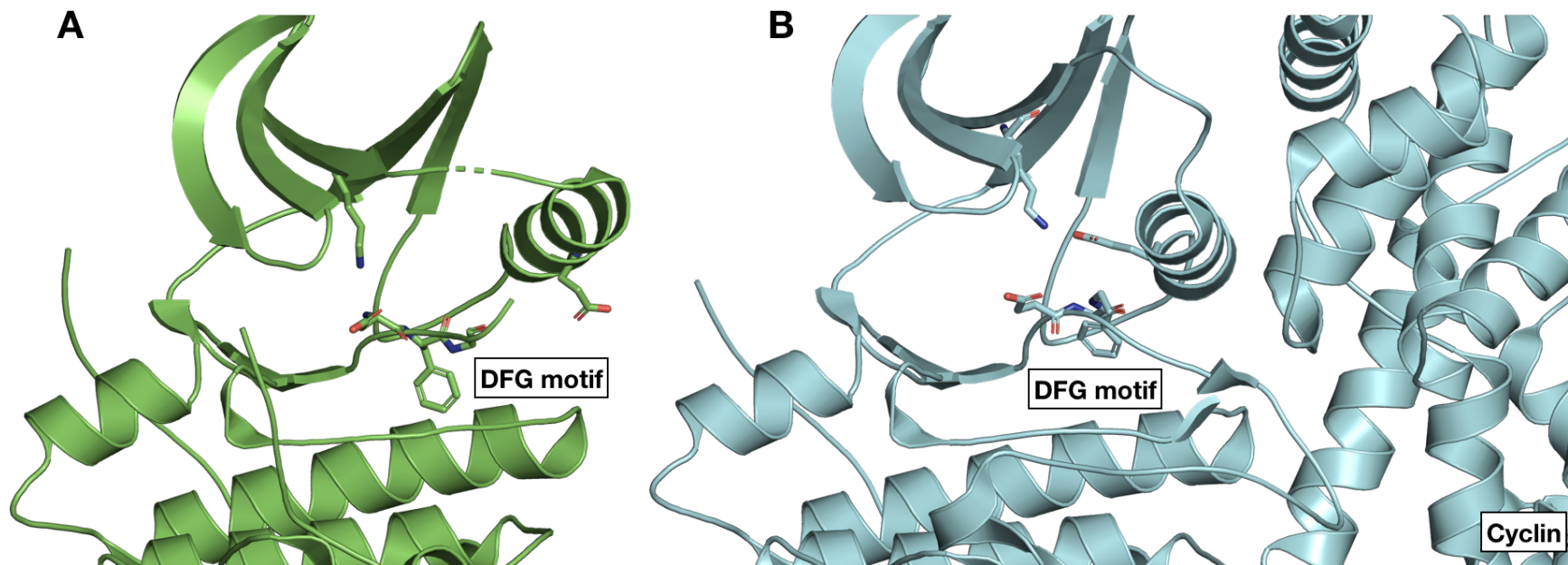


Figure 2.6: **CDK2 adopts an inactive conformation in the crystal structure used for the CDK2/ERK2 calculations** (A) CDK2 (5K4J) adopts an inactive conformation in the absence of its cyclin. The DFG motif is in a DFG-in conformation, with the  $\alpha$ C helix rotated outwards, breaking the salt bridge between K33 and E51 (Uniprot numbering) that is typically a marker of an active conformation. Notably, the Phe in the DFG motif does not completely form the hydrophobic spine due to the rotation of the  $\alpha$ C helix [? ] (B) The CDK2 structure used for the CDK2/CDK9 calculations (4BCK) contains cyclin A and adopts a DFG-in/ $\alpha$ C helix-in conformation that forms the salt bridge between K33 and E51. This is typically indicative of a fully active kinase [9, 71].

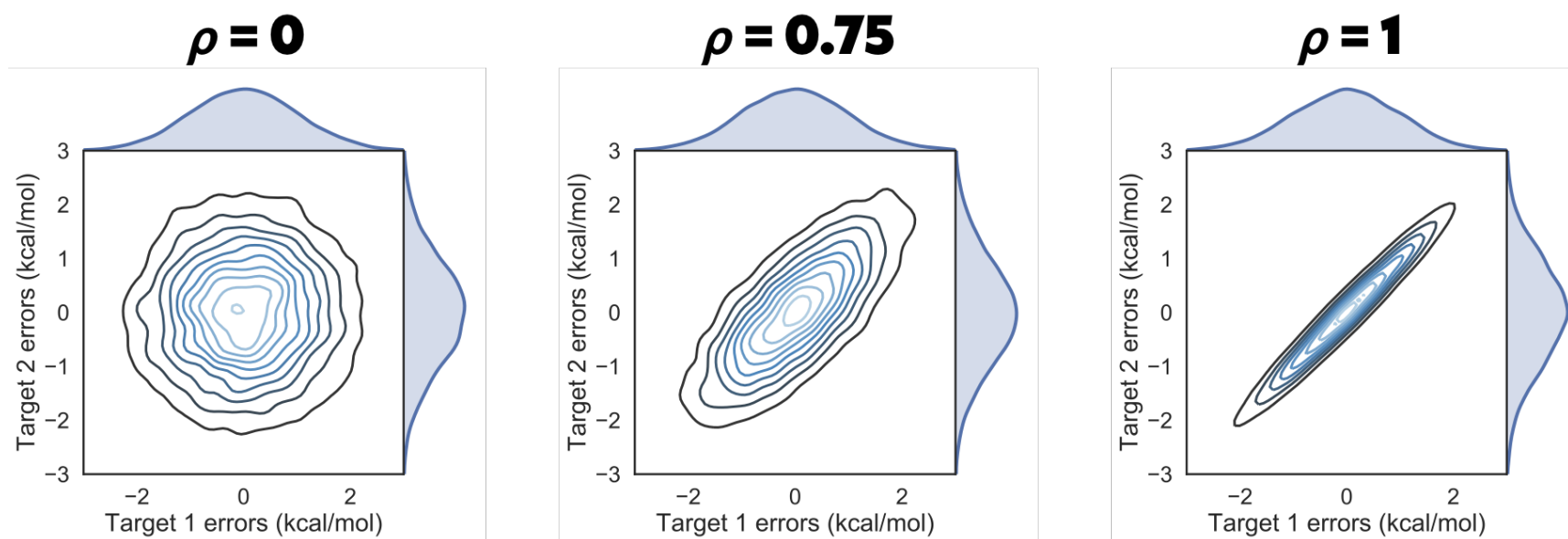


Figure 2.7: **Correlation coefficient  $\rho$  controls the shape of the joint marginal distribution of errors** As  $\rho$  increases, the joint marginal distribution of errors become more diagonal. Each panel shows 10000 samples drawn from a multivariate normal distribution centered around 0 kcal/mol, where the per target error was set to 1 kcal/mol and  $\rho$  to the value indicated in bold over the plot.



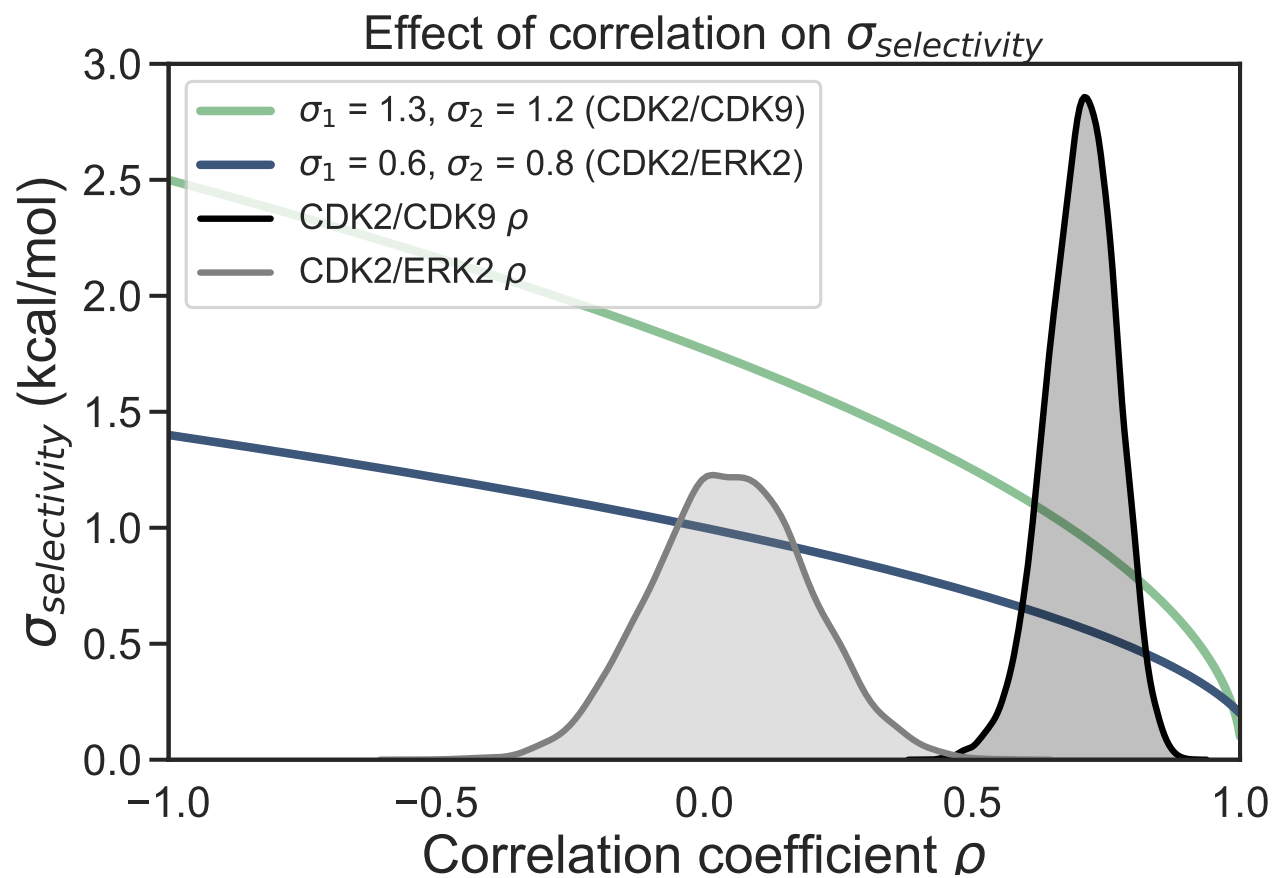


Figure 2.8: **Correlation reduces the expected error for selectivity predictions** As correlation coefficient  $\rho$  increases,  $\sigma_{selectivity}$  decreases. The intersection between CDK2/CDK9  $\sigma_{selectivity}$  (green curve) and  $\rho$  (black distribution) indicates the range of expected  $\sigma_{selectivity}$  values. The intersection for CDK2/ERK  $\sigma_{selectivity}$  (blue curve) and  $\rho$  (gray distribution) suggests the expected  $\sigma_{selectivity}$  range for that set of calculations.

CHAPTER 3

UNDERSTANDING THE FUNCTIONAL IMPACT OF MTOR CLINICAL  
KINASE MUTATIONS USING PHYSICAL MODELING

---

This will include  
the mTOR and  
Abl resistance  
section

## 3.1 Introduction

### 3.1.1 Section on mTOR biology and mutations

### 3.1.2 Section on MD simulations and Folding@Home

## 3.2 Methods

## CHAPTER 4

# PREDICTING THE IMPACT OF CLINICALLY-OBSERVED ABL INHIBITORS USING PHYSICAL MODELING AND BIOPHYSICAL EXPERIMENTS

### 4.1 Introduction

Targeted kinase inhibitors are a major therapeutic class in the treatment of cancer. A total of 43 selective small molecule kinase inhibitors have now been approved by the FDA [2], including 34 approved to treat cancer, and perhaps 50% of all current drugs in development target kinases [3]. Despite the success of selective inhibitors, the emergence of drug resistance remains a challenge in the treatment of cancer [72–79] and has motivated the development of second- and then third-generation inhibitors aimed at overcoming recurrent resistance mutations [30, 80–83].

While a number of drug resistance mechanisms have been identified in cancer (e.g., induction of splice variants [84], or alleviation of feedback [27]), inherent or acquired missense mutations in the kinase domain of the target of therapy are a major form of resistance to tyrosine kinase inhibitors (TKI) [18, 79, 85]. Oncology is entering a new era with major cancer centers now deep sequencing tumors to reveal genetic alterations that may render subclonal populations susceptible or resistant to targeted inhibitors [1], but the use of this information in precision medicine has lagged behind. It would be of enormous value in clinical practice if an oncologist could reliably ascertain whether these mutations render the target of therapy resistant or susceptible to available inhibitors; such tools would facilitate the enrollment of patients in mechanism-based basket trials [86, 87], help prioritize candidate compounds for clinical trials, and aid the development of next-generation

inhibitors.

#### **4.1.1 The long tail of rare kinase mutations frustrates prediction of drug resistance**

While some cancer missense mutations are highly recurrent and have been characterized clinically or biochemically, a long tail of rare mutations collectively accounts for the majority of clinically observed missense mutations (Figure ??a), leaving clinicians and researchers without knowledge of whether these uncharacterized mutations might lead to resistance. While rules-based and machine learning schemes are still being assessed in oncology contexts, work in predicting drug response to microbial resistance has shown that rare mutations present a significant challenge to approaches that seek to predict resistance to therapy [88]. Clinical cancer mutations may impact drug response through a variety of mechanisms by altering kinase activity, ATP affinity, substrate specificities, and the ability to participate in regulatory interactions, compounding the difficulties associated with limited datasets that machine learning approaches face. In parallel with computational approaches, high-throughput experimental techniques such as MITE-Seq [89] have been developed to assess the impact of point mutations on drug response. However, the complexity of defining selection schemes that reliably correlate with *in vivo* drug effectiveness and long turn-around times might limit their ability to rapidly and reliably impact clinical decision-making.

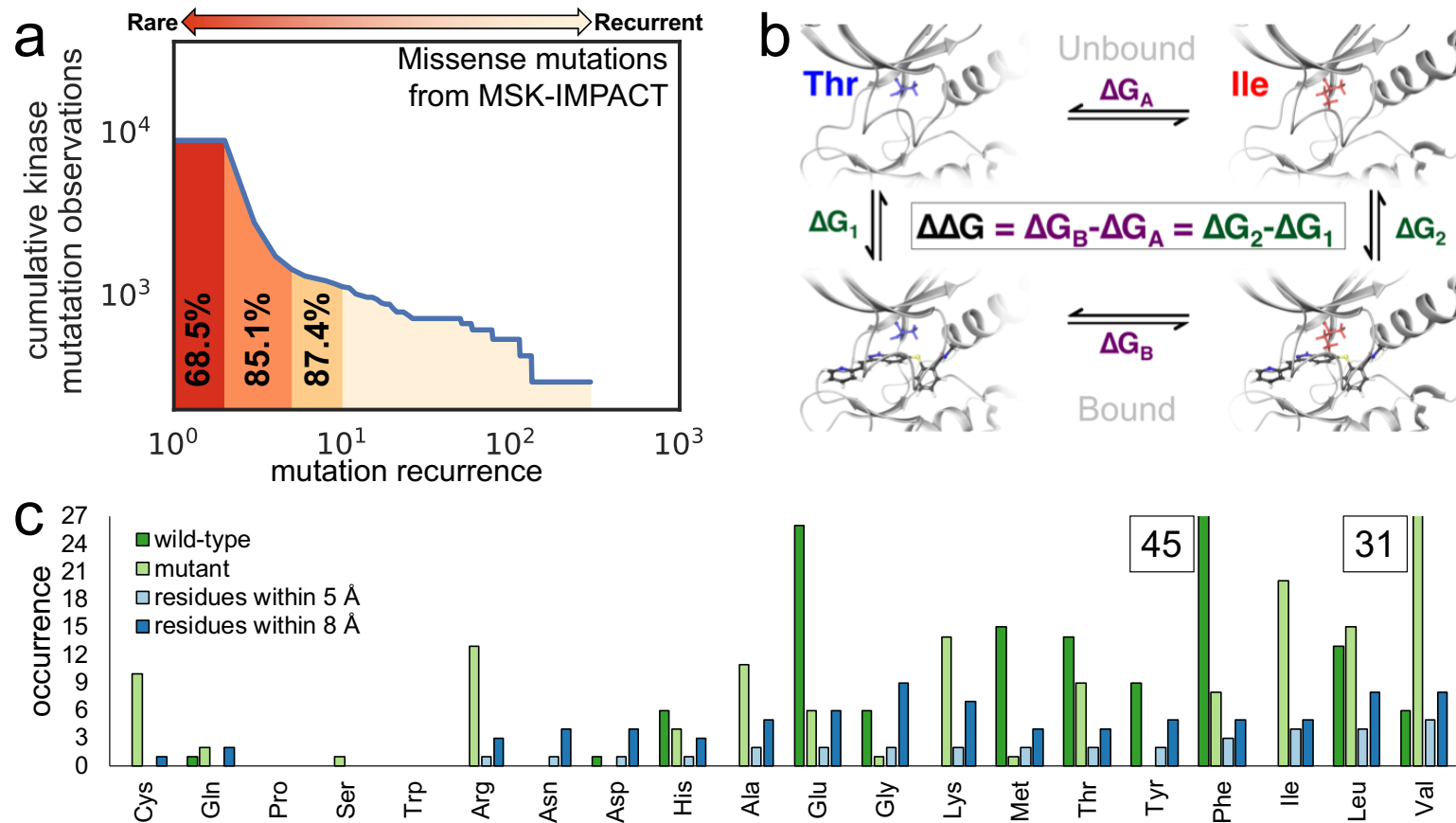


Figure 4.1: **Relative alchemical free-energy calculations can be used to predict affinity changes of FDA-approved selective kinase inhibitors arising from clinically-identified mutations in their targets of therapy.** (a) Missense mutation statistics derived from 10,336 patient samples subjected to Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT) deep sequencing panel [1] show that 68.5% of missense kinase mutations in cancer patients have never been observed previously, while 87.4% have been observed no more than ten times; the vast majority of clinically observed missense kinase mutations are unique to each patient. (b) To compute the impact of a clinical point mutation on inhibitor binding free energy, a thermodynamic cycle can be used to relate the free energy of the wild-type and mutant kinase in the absence (top) and presence (bottom) of the inhibitor. (c) Summary of mutations studied in this work. Frequency of the wild-type (dark green) and mutant (green) residues for the 144 clinically-identified Abl mutations used in this study (see [Table 4.1](#) for data sources). Also shown is the frequency of residues within 5 (light blue) and 8 (blue) of the binding pocket. The ordering of residues along the x-axis corresponds to the increasing occurrence of residues within 5 of the binding pocket. The number of wild-type Phe residues (n=45) and mutant Val residues (n=31) exceeded the limits of the y-axis.

### **4.1.2 Alchemical free-energy methods can predict inhibitor binding affinities**

Physics-based approaches could be complementary to machine-learning and experimental techniques in predicting changes in TKI affinity due to mutations with few or no prior clinical observations. Modern atomistic molecular mechanics forcefields such as OPLS3 [52], CHARMM [50], and AMBER FF14SB [51] have reached a sufficient level of maturity to enable the accurate and reliable prediction of receptor-ligand binding free energy. Alchemical free-energy methods permit receptor-ligand binding energies to be computed rigorously, including all relevant entropic and enthalpic contributions [90]. Encouragingly, kinase:inhibitor binding affinities have been predicted using alchemical free-energy methods with mean unsigned errors of 1.0 kcal mol<sup>-1</sup> for CDK2, JNK1, p38, and Tyk2 [55, 57]. Beyond kinases, alchemical approaches have predicted the binding affinity of BRD4 inhibitors with mean absolute errors of 0.6 kcal mol<sup>-1</sup> [91]. Alchemical methods have also been observed to have good accuracy (0.6 kcal mol<sup>-1</sup> mean unsigned error for Tyk2 tyrosine kinase) in the prediction of relative free energies for ligand transformations within a complex whose receptor geometry was generated using a homology model [92].

### **4.1.3 Alchemical approaches can predict the impact of protein mutations on free energy**

Alchemical free-energy calculations have also been used to predict the impact of mutations on protein-protein binding [93] and protein thermostabilities [94]. Recent work has found that protein mutations can be predicted to be stabilizing or

destabilizing with a classification accuracy of 71% across ten proteins and 62 mutations [95]. The impact of Gly to D-Ala mutations on protein stability was predicted using an alchemical approach with a similar level of accuracy [96]. Recently, one study has hinted at the potential utility of alchemical free-energy calculations in oncology by predicting the impact of a single clinical mutation on the binding free energies of the TKIs dasatinib and RL45 [97].

#### **4.1.4 Assessing the potential for physical modeling to predict resistance to FDA-approved TKIs**

Here, we ask whether physical modeling techniques may be useful in predicting whether clinically-identified kinase mutations lead to drug resistance or drug sensitivity. We perform state-of-the-art relative alchemical free-energy calculations using FEP+ [55], recently demonstrated to achieve sufficiently good accuracy to drive the design of small-molecule inhibitors for a broad range of targets during lead optimization [55, 57, 90, 98], to calculate the effect of point mutation on the binding free energy between the inhibitor and the kinase receptor (Figure 4.1b). We compare this approach against a fast but approximate physical modeling method implemented in Prime [99] (an MM-GBSA approach) in which an implicit solvent model is used to assess the change in minimized interaction energy of the ligand with the mutant and wild-type kinase. We consider whether these methods can predict a ten-fold reduction in inhibitor affinity (corresponding to a binding free energy change of  $1.36 \text{ kcal mol}^{-1}$ ) to assess baseline utility. As a benchmark, we compile a set of reliable inhibitor  $\Delta\text{pIC}_{50}$  data for 144 clinically-identified mutants of the human kinase Abl, an important oncology target dysregulated in cancers like chronic myelogenous leukemia (CML), for which six [2] FDA-approved TKIs

are available. While  $\Delta\text{pIC}_{50}$  can approximate a dissociation constant  $\Delta K_D$ , other processes contributing to changes in cell viability might affect  $\text{IC}_{50}$  in ways that are not accounted for by a traditional binding experiment, motivating a quantitative comparison between  $\Delta\text{pIC}_{50}$  and  $\Delta K_D$ . The results of this benchmark demonstrate the potential for FEP+ to predict the impact that mutations in Abl kinase have on drug binding, and a classification accuracy of  $88^{93}_{82}\%$  (for all statistical metrics reported in this paper, the 95% confidence intervals (CI) is shown in the form of  $(x^{upper}_{lower})$ ), an RMSE of  $1.07^{1.26}_{0.89}$  kcal mol<sup>-1</sup>, and an MUE of  $0.79^{0.92}_{0.67}$  kcal mol<sup>-1</sup> was achieved.

#### 4.1.5 A benchmark of $\Delta\text{pIC}_{50}$ s for predicting mutational resistance

To construct a benchmark evaluation dataset, we compiled a total of 144  $\Delta\text{pIC}_{50}$  measurements of Abl:TKI affinities, summarized in Table 4.1, taking care to ensure all measurements for an individual TKI were reported in the same study from experiments run under identical conditions. 131  $\Delta\text{pIC}_{50}$  measurements were available across the six TKIs with available co-crystal structures with wild-type Abl—26 for axitinib and 21 for bosutinib, dasatinib, imatinib, nilotinib, and ponatinib. 13  $\Delta\text{pIC}_{50}$  measurements were available for the two TKIs for which docking was necessary to generate Abl:TKI structures—7 for erlotinib and 6 for gefitinib. For added diversity, this set includes TKIs for which Abl is not the primary target—axitinib, erlotinib, and gefitinib. All mutations in this benchmark dataset have been clinically-observed (Table 4.1). Due to the change in bond topology required by mutations involving proline, which is not currently supported by the FEP+ technology for protein residue mutations, the three mutations H396P (axitinib,



gefitinib, erlotinib) were excluded from our assessment. As single point mutations were highly represented in the Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT) study analyzed in Figure 4.1a, we excluded double mutations from this work. However, the impact of mutations from multiple sites can potentially be modeled by sequentially mutating each site and this will be addressed in future work.

Table 4.1: Public  $\Delta pIC_{50}$  datasets for 144 Abl kinase mutations and eight tyrosine kinase inhibitors (TKIs) with corresponding wild-type co-crystal structures used in this study

TKI	$N_{mut}$	R	S	PDB	(kcal mol <sup>-1</sup> ) $ \Delta G_{max} - \Delta G_{min} $	Source	(kcal mol <sup>-1</sup> ) $\Delta G_{WT}$
axitinib	26	0	26	4wa9	2.05	[100]	-8.35
bosutinib	21	4	17	3ue4	2.79	[101]	-9.81
dasatinib	21	5	16	4xey	5.08	[101]	-11.94
imatinib	21	5	16	1opj	2.16	[101]	-9.19
nilotinib	21	4	17	3cs9	3.88	[101]	-10.74
ponatinib	21	0	21	3oxz	1.00	[101]	-11.70
<b>subtotal</b>	<b>131</b>	<b>18</b>	<b>113</b>				
erlotinib	7	1	6	<i>Dock to 3ue4</i>	1.73	[12]	-9.77
gefitinib	6	0	6	<i>Dock to 3ue4</i>	1.79	[12]	-8.84
<b>total</b>	<b>144</b>	<b>19</b>	<b>125</b>				

$N_{mut}$ : Total number of mutants for which  $\Delta pIC_{50}$  data was available.

Number of **R**esistant, **S**usceptible mutants using 10-fold affinity change threshold.

PDB: Source PDB ID, or *Dock to 3ue4*, which used 3ue4 as the receptor for Glide-SP docking inhibitors without co-crystal structure.

$\Delta G_{WT}$ : Binding free energy of inhibitor to wild-type Abl, as estimated from  $IC_{50}$  data.

Experimental  $\Delta pIC_{50}$  measurements for wild-type and mutant Abl were converted to  $\Delta\Delta G$  in order to make direct comparisons between physics-based models and experiment. However, computation of experimental uncertainties were required to understand the degree to which differences between predictions and experimental data were significant. Since experimental error estimates for measured  $IC_{50}$ s were not available for the data in *Table 4.1*, we compared that data to other sources that have published  $IC_{50}$ s for the same mutations in the presence of the same TKIs (*Figure 2.2a,b,c*). Cross-comparison of 97 experimentally measured  $\Delta\Delta G$ s derived from cell viability assay  $IC_{50}$  data led to an estimate of experimental variability of  $0.32^{0.36}_{0.28}$  kcal mol<sup>-1</sup> root-mean square error (RMSE) that described the expected repeatability of the measurements. Because multiple factors influence the  $IC_{50}$  aside from direct effects on the binding affinity—the focus of this study—we also compared  $\Delta\Delta G$ s derived from  $\Delta pIC_{50}$ s with those derived from binding affinity measurements ( $\Delta K_d$ ) for which data for a limited set of 27 mutations was available (*Figure ??d*); the larger computed RMSE of  $0.81^{1.04}_{0.59}$  kcal mol<sup>-1</sup> represents an estimate of the lower bound of the RMSE to the  $IC_{50}$ -derived  $\Delta\Delta G$ s that we might hope to achieve with FEP+ or Prime, which were performed using non-phosphorylated models, when comparing sample statistics directly. In comparing 31 mutations for which phosphorylated and non-phosphorylated  $\Delta K_d$ s were available, we found a strong correlation between the  $\Delta\Delta G$ s derived from those data ( $r=0.94$ , Supplementary Figure ??); the statistics of that comparison are similar to those of the inter-lab variability comparison.

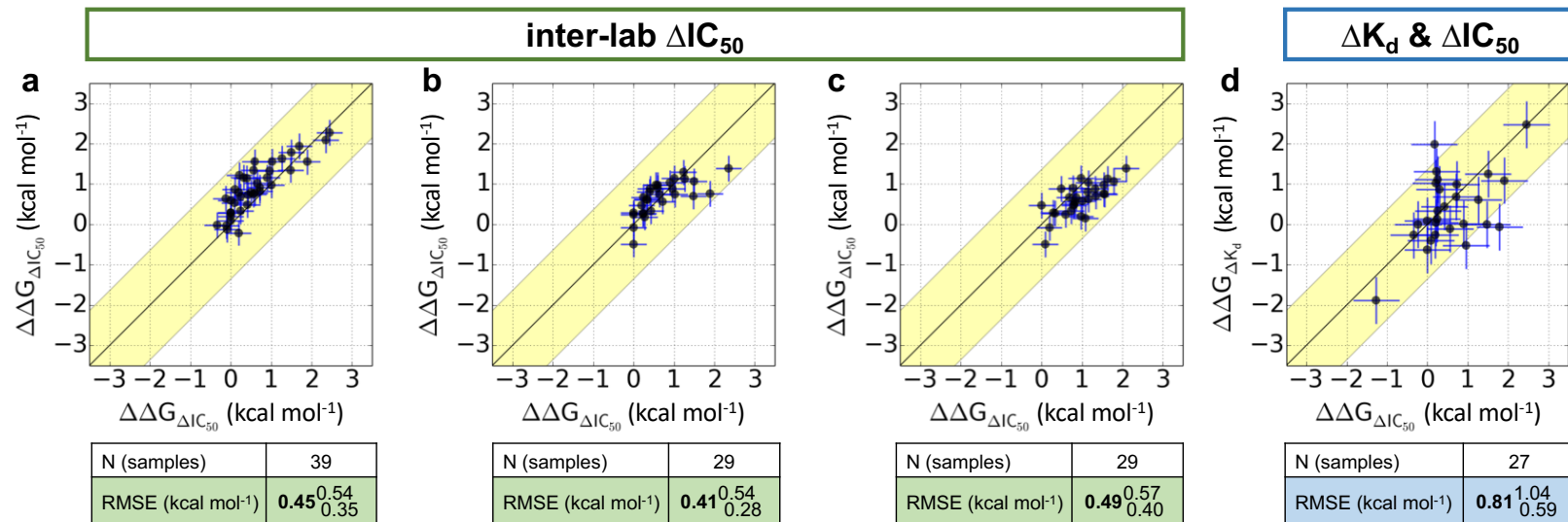


Figure 4.2: **Cross-comparison of the experimentally measured effects that mutations in Abl kinase have on ligand binding, performed by different labs.**  $\Delta\Delta G$  was computed from publicly available  $\Delta pIC_{50}$  or  $\Delta pK_d$  measurements and these values of  $\Delta\Delta G$  were then plotted and the RMSE between them reported. (a)  $\Delta pIC_{50}$  measurements (X-axis) from [101] compared with  $\Delta pIC_{50}$  measurements (Y-axis) from [? ]. (b)  $\Delta pIC_{50}$  measurements (X-axis) from [101] compared with  $\Delta pIC_{50}$  measurements (Y-axis) from [? ]. (c)  $\Delta pIC_{50}$  measurements (X-axis) from [? ] compared with  $\Delta pIC_{50}$  measurements (Y-axis) from [? ]. (d)  $\Delta pIC_{50}$  measurements (X-axis) from [101] compared with  $\Delta pK_d$  measurements (Y-axis) from [12] using non-phosphorylated Abl kinase. Scatter plot error bars in a,b,and c are  $\pm$ standard error (SE) taken from the combined 97 inter-lab  $\Delta\Delta G$ s derived from the  $\Delta pIC_{50}$  measurements, which was  $0.32^{0.36}_{0.28}$ ; the RMSE was  $0.45^{0.51}_{0.39}$  kcal mol<sup>-1</sup>. Scatter plot error bars in d are the  $\pm$ standard error (SE) of  $\Delta\Delta G$ s derived from  $\Delta pIC_{50}$  and  $\Delta pK_d$  from a set of 27 mutations, which is  $0.58^{0.74}_{0.42}$  kcal mol<sup>-1</sup>; the RMSE was  $0.81^{1.04}_{0.59}$  kcal mol<sup>-1</sup>.

**4.1.6 Summary of experimental data that is available and what needs to be done to make better information available**

**4.1.7 Open-source YANK/Amber forcefield stuff**

## BIBLIOGRAPHY

- [1] **Zehir A**, Benayed R, Shah RH, Syed A, Middha S, Kim HR, Srinivasan P, Gao J, Chakravarty D, Devlin SM, Hellmann MD, Barron DA, Schram AM, Hameed M, Dogan S, Ross DS, Hechtman JF, DeLair DF, Yao J, Mandelker DL, et al. Mutational Landscape of Metastatic Cancer Revealed from Prospective Clinical Sequencing of 10,000 Patients. *Nat Med*. 2017 May; 23(6):703–713. doi: [10.1038/nm.4333](https://doi.org/10.1038/nm.4333).
- [2] **Robert Roskoski Jr.** USFDA Approved Protein Kinase Inhibitors. . 2017; <http://www.brimr.org/PKI/PKIs.htm>, updated 3 May 2017.
- [3] **Santos R**, Ursu O, Gaulton A, Bento AP, Donadi RS, Bologa CG, Karlsson A, Al-Lazikani B, Hersey A, Oprea TI, Overington JP. A Comprehensive Map of Molecular Drug Targets. *Nat Rev Drug Discov*. 2016 Dec; 16(1):19–34. doi: [10.1038/nrd.2016.230](https://doi.org/10.1038/nrd.2016.230).
- [4] **Volkamer A**, Eid S, Turk S, Jaeger S, Rippmann F, Fulle S. Pocketome of human kinases: prioritizing the ATP binding sites of (yet) untapped protein kinases for drug discovery. *J Chem Inf Model*. 2015 Mar; 55(3):538–549.
- [5] **Manning G**, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The Protein Kinase Complement of the Human Genome. *Science*. 2002 Dec; 298(5600):1912–1934.
- [6] **Wu P**, Nielsen TE, Clausen MH. FDA-approved small-molecule kinase inhibitors. *Trends Pharmacol Sci*. 2015 Jul; 36(7):422–439.
- [7] **Cowan-Jacob SW**, Fendrich G, Floersheimer A, Furet P, Liebetanz J, Rummel G, Rheinberger P, Centeleghe M, Fabbro D, Manley PW, IUCr. Structural biology contributions to the discovery of drugs to treat chronic myelogenous leukaemia. *Acta Crystallogr D Biol Crystallogr*. 2007 Jan; 63(1):80–93.
- [8] **Seeliger MA**, Nagar B, Frank F, Cao X, Henderson MN, Kuriyan J. c-Src Binds to the Cancer Drug Imatinib with an Inactive Abl/c-Kit Conformation and a Distributed Thermodynamic Penalty. *Structure*. 2007 Mar; 15(3):299–311.

- [9] **Huse M**, Kuriyan J. The conformational plasticity of protein kinases. *Cell*. 2002 Jan; 109(3):275–282.
- [10] **Harrison SC**. Variation on an Src-like theme. *Cell*. 2003 Mar; 112(6):737–740.
- [11] **Anastassiadis T**, Deacon SW, Devarajan K, Ma H, Peterson JR. Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity. *Nat Biotechnol*. 2011 Nov; 29(11):1039–1045.
- [12] **Davis MI**, Hunt JP, Herrgard S, Ciceri P, Wodicka LM, Pallares G, Hocker M, Treiber DK, Zarrinkar PP. Comprehensive Analysis of Kinase Inhibitor Selectivity. *Nat Biotechnol*. 2011 Oct; 29(11):1046–1051. doi: [10.1038/nbt.1990](https://doi.org/10.1038/nbt.1990).
- [13] **Klaeger S**, Heinzlmeir S, Wilhelm M, Polzer H, Vick B, Koenig PA, Reinecke M, Ruprecht B, Petzoldt S, Meng C, Zecha J, Reiter K, Qiao H, Helm D, Koch H, Schoof M, Canevari G, Casale E, Depaolini SR, Feuchtinger A, et al. The target landscape of clinical kinase drugs. *Science*. 2017 Dec; 358(6367).
- [14] **Zhang J**, Yang PL, Gray NS. Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer*. 2009 Jan; 9(1):28–39.
- [15] **Huggins DJ**, Sherman W, Tidor B. Rational approaches to improving selectivity in drug design. *J Med Chem*. 2012 Feb; 55(4):1424–1444.
- [16] **Fan QW**, Cheng CK, Nicolaides TP, Hackett CS, Knight ZA, Shokat KM, Weiss WA. A dual phosphoinositide-3-kinase  $\alpha$ /mTOR inhibitor cooperates with blockade of epidermal growth factor receptor in PTEN-mutant glioma. *Cancer Res*. 2007 Sep; 67(17):7960–7965.
- [17] **Apsel B**, Blair JA, Gonzalez B, Nazif TM, Feldman ME, Aizenstein B, Hoffman R, Williams RL, Shokat KM, Knight ZA. Targeted polypharmacology: discovery of dual inhibitors of tyrosine and phosphoinositide kinases. *Nat Chem Biol*. 2008 Nov; 4(11):691–699.

- [18] **Knight ZA**, Lin H, Shokat KM. Targeting the Cancer Kinome through Polypharmacology. *Nat Rev Cancer*. 2010; 10(2):130.
- [19] **Hopkins AL**, Mason JS, Overington JP. Can we rationally design promiscuous drugs? *Curr Opin Struct Biol*. 2006 Feb; 16(1):127–136.
- [20] **Hopkins AL**. Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol*. 2008 Nov; 4(11):682–690.
- [21] **Rudmann DG**. On-target and off-target-based toxicologic effects. *Toxicol Pathol*. 2013 Feb; 41(2):310–314.
- [22] **Kijima T**, Shimizu T, Nonen S, Furukawa M, Otani Y, Minami T, Takahashi R, Hirata H, Nagatomo I, Takeda Y, Kida H, Goya S, Fujio Y, Azuma J, Tachibana I, Kawase I. Safe and successful treatment with erlotinib after gefitinib-induced hepatotoxicity: difference in metabolism as a possible mechanism. *J Clin Oncol*. 2011 Jul; 29(19):e588–90.
- [23] **Liu S**, Kurzrock R. Toxicity of targeted therapy: Implications for response and impact of genetic polymorphisms. *Cancer Treat Rev*. 2014 Aug; 40(7):883–891.
- [24] **Mendoza MC**, Er EE, Blenis J. The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. *Trends Biochem Sci*. 2011 Jun; 36(6):320–328.
- [25] **Tricker EM**, Xu C, Uddin S, Capelletti M, Ercan D, Ogino A, Pratilas CA, Rosen N, Gray NS, Wong KK, Jänne PA. Combined EGFR/MEK Inhibition Prevents the Emergence of Resistance in EGFR-Mutant Lung Cancer. *Cancer Discov*. 2015 Sep; 5(9):960–971.
- [26] **Bailey ST**, Zhou B, Damrauer JS, Krishnan B, Wilson HL, Smith AM, Li M, Yeh JJ, Kim WY. mTOR Inhibition Induces Compensatory, Therapeutically Targetable MEK Activation in Renal Cell Carcinoma. *PLoS One*. 2014 Sep; 9(9):e104413.
- [27] **Chandarlapaty S**, Sawai A, Scaltriti M, Rodrik-Outmezguine V, Grbovic-Huezo O, Serra V, Majumder PK, Baselga J, Rosen N. AKT Inhibition Relieves Feedback Sup-



pression of Receptor Tyrosine Kinase Expression and Activity. *Cancer Cell*. 2011 Jan; 19(1):58–71. doi: [10.1016/j.ccr.2010.10.031](https://doi.org/10.1016/j.ccr.2010.10.031).

- [28] **Pao W**, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, Singh B, Heelan R, Rusch V, Fulton L, Mardis E, Kupfer D, Wilson R, Kris M, Varmus H. EGF receptor gene mutations are common in lung cancers from “never smokers” and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proceedings of the National Academy of Sciences*. 2004 Sep; 101(36):13306–13311.
- [29] **Kim Y**, Li Z, Apetri M, Luo B, Settleman JE, Anderson KS. Temporal resolution of autophosphorylation for normal and oncogenic forms of EGFR and differential effects of gefitinib. *Biochemistry*. 2012 Jun; 51(25):5212–5222.
- [30] **Juchum M**, Günther M, Laufer SA. Fighting Cancer Drug Resistance: Opportunities and Challenges for Mutation-Specific EGFR Inhibitors. *Drug Resist Updat*. 2015 May; 20:12–28. doi: [10.1016/j.drug.2015.05.002](https://doi.org/10.1016/j.drug.2015.05.002).
- [31] **Din OS**, Woll PJ. Treatment of gastrointestinal stromal tumor: focus on imatinib mesylate. *Ther Clin Risk Manag*. 2008 Feb; 4(1):149–162.
- [32] **Drewry DH**, Wells CI, Andrews DM, Angell R, Al-Ali H, Axtman AD, Capuzzi SJ, Elkins JM, Ettmayer P, Frederiksen M, Gileadi O, Gray N, Hooper A, Knapp S, Laufer S, Luecking U, Michaelides M, Müller S, Muratov E, Aldrin Denny R, et al. Progress towards a public chemogenomic set for protein kinases and a call for contributions. *PLoS One*. 2017 Aug; 12(8):e0181585.
- [33] **Cheng AC**, Eksterowicz J, Geuns-Meyer S, Sun Y. Analysis of kinase inhibitor selectivity using a thermodynamics-based partition index. *J Med Chem*. 2010 Jun; 53(11):4502–4510.
- [34] **Uitdehaag JCM**, Verkaar F, Alwan H, de Man J, Buijsman RC, Zaman GJR. A guide to picking the most selective kinase inhibitor tool compounds for pharmacological validation of drug targets. *Br J Pharmacol*. 2012 Jun; 166(3):858–876.

- [35] **Vasta JD**, Corona CR, Wilkinson J, Zimprich CA, Hartnett JR, Ingold MR, Zimmerman K, Machleidt T, Kirkland TA, Huwiler KG, Ohana RF, Slater M, Otto P, Cong M, Wells CI, Berger BT, Hanke T, Glas C, Ding K, Drewry DH, et al. Quantitative, Wide-Spectrum Kinase Profiling in Live Cells for Assessing the Effect of Cellular ATP on Target Engagement. *Cell chemical biology*. 2018 Feb; 25(2):206–214.e11.
- [36] **Merget B**, Turk S, Eid S, Rippmann F, Fulle S. Profiling Prediction of Kinase Inhibitors: Toward the Virtual Assay. *J Med Chem*. 2017 Jan; 60(1):474–485.
- [37] **Volkamer A**, Eid S, Turk S, Rippmann F, Fulle S. Identification and Visualization of Kinase-Specific Subpockets. *J Chem Inf Model*. 2016 Feb; 56(2):335–346.
- [38] **Christmann-Franck S**, van Westen GJP, Papadatos G, Beltran Escudie F, Roberts A, Overington JP, Domine D. Unprecedentedly Large-Scale Kinase Inhibitor Set Enabling the Accurate Prediction of Compound–Kinase Activities: A Way toward Selective Promiscuity by Design? *Journal of chemical information and modeling*. 2016 Sep; 56(9):1654–1675.
- [39] **Gao Y**, Davies SP, Augustin M, Woodward A, Patel UA, Kovelman R, Harvey KJ. A broad activity screen in support of a chemogenomic map for kinase signalling research and drug discovery. *The Biochemical journal*. 2013 Apr; 451(2):313–328.
- [40] **Kramer C**, Kalliokoski T, Gedeck P, Vulpetti A. The Experimental Uncertainty of Heterogeneous Public K<sub>i</sub> Data. *J Med Chem*. 2012 Jun; 55(11):5165–5173. doi: 10.1021/jm300131x.
- [41] **Brown SP**, Muchmore SW, Hajduk PJ. Healthy Skepticism: Assessing Realistic Model Performance. *Drug Discov Today*. 2009; 14(7):420 – 427. doi: <http://dx.doi.org/10.1016/j.drudis.2009.01.012>.
- [42] **Albanese SK**, Parton DL, Işık M, Rodríguez-Laureano L, Hanson SM, Behr JM, Gradia S, Jeans C, Levinson NM, Seeliger MA, Chodera JD. An Open Library of Human Kinase Domain Constructs for Automated Bacterial Expression. *Biochemistry*. 2018 Aug; 57(31):4675–4689.

- [43] **Seeliger MA**, Young M, Henderson MN, Pellicena P, King DS, Falick AM, Kuriyan J. High yield bacterial expression of active c-Abl and c-Src tyrosine kinases. *Protein science : a publication of the Protein Society*. 2005 Dec; 14(12):3135–3139.
- [44] **The Structural Genomics Consortium**, SGC: Progress in protein kinase structural biology; 2012. <http://www.thesgc.org/scientists/resources/kinases>, updated 18 Jul 2012.
- [45] **Chambers SP**, Austen DA, Fulghum JR, Kim WM. High-throughput screening for soluble recombinant expressed kinases in *Escherichia coli* and insect cells. *Protein expression and purification*. 2004 Jul; 36(1):40–47.
- [46] **Levinson NM**, Kuchment O, Shen K, Young MA, Koldobskiy M, Karplus M, Cole PA, Kuriyan J. A Src-Like Inactive Conformation in the Abl Tyrosine Kinase Domain. *PLoS Biol*. 2006 May; 4(5):e144. doi: 10.1371/journal.pbio.0040144.
- [47] **Lin YL**, Meng Y, Jiang W, Roux B. Explaining why Gleevec is a specific and potent inhibitor of Abl kinase. *Proc Natl Acad Sci U S A*. 2013 Jan; 110(5):1664–1669.
- [48] **Lin YL**, Meng Y, Huang L, Roux B. Computational Study of Gleevec and G6G Reveals Molecular Determinants of Kinase Inhibitor Selectivity. *J Am Chem Soc*. 2014 Oct; 136(42):14753–14762.
- [49] **Lin YL**, Roux B. Computational Analysis of the Binding Specificity of Gleevec to Abl, c-Kit, Lck, and c-Src Tyrosine Kinases. *J Am Chem Soc*. 2013 Oct; 135(39):14741–14753.
- [50] **Huang J**, MacKerell AD. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J Comput Chem*. 2013 Sep; 34(25):2135–2145. doi: 10.1002/jcc.23354.
- [51] **Maier JA**, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput*. 2015 Aug; 11(8):3696–3713. doi: 10.1021/acs.jctc.5b00255.

- [52] **Harder E**, Damm W, Maple J, Wu C, Reboul M, Xiang JY, Wang L, Lupyan D, Dahlgren MK, Knight JL, Kaus JW, Cerutti DS, Krilov G, Jorgensen WL, Abel R, Friesner RA. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J Chem Theory Comput.* 2016 Jan; 12(1):281–296. doi: [10.1021/acs.jctc.5b00864](https://doi.org/10.1021/acs.jctc.5b00864).
- [53] **Cournia Z**, Allen B, Sherman W. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *Journal of chemical information and modeling.* 2017 Dec; 57(12):2911–2937.
- [54] **Chodera JD**, Mobley DL, Shirts MR, Dixon RW, Branson K, Pande VS. Alchemical free energy methods for drug discovery: progress and challenges. *Curr Opin Struct Biol.* 2011 Apr; 21(2):150–160.
- [55] **Wang L**, Wu Y, Deng Y, Kim B, Pierce L, Krilov G, Lupyan D, Robinson S, Dahlgren MK, Greenwood J, Romero DL, Masse C, Knight JL, Steinbrecher T, Beuming T, Damm W, Harder E, Sherman W, Brewer M, Wester R, et al. 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 Feb; 137(7):2695–2703. doi: [10.1021/ja512751q](https://doi.org/10.1021/ja512751q).
- [56] **Abel R**, Mondal S, Masse C, Greenwood J, Harriman G, Ashwell MA, Bhat S, Wester R, Frye L, Kapeller R, Friesner RA. Accelerating Drug Discovery through Tight Integration of Expert Molecular Design and Predictive Scoring. *Curr Opin Struct Biol.* 2017 Apr; 43:38–44. doi: [10.1016/j.sbi.2016.10.007](https://doi.org/10.1016/j.sbi.2016.10.007).
- [57] **Abel R**, Mondal S, Masse C, Greenwood J, Harriman G, Ashwell MA, Bhat S, Wester R, Frye L, Kapeller R, et al. Accelerating drug discovery through tight integration of expert molecular design and predictive scoring. *Current opinion in structural biology.* 2017; 43:38–44.
- [58] **Aldeghi M**, Heifetz A, Bodkin MJ, Knapp S, Biggin PC. Accurate calculation of the

absolute free energy of binding for drug molecules. *Chem Sci*. 2015 Dec; 7(1):207–218.

- [59] **Hauser K**, Negron C, Albanese SK, Ray S, Steinbrecher T, Abel R, Chodera JD, Wang L. Predicting resistance of clinical Abl mutations to targeted kinase inhibitors using alchemical free-energy calculations. *Communications Biology*. 2018 Jun; 1(1):70.
- [60] **Aldeghi M**, Heifetz A, Bodkin MJ, Knapp S, Biggin PC. Predictions of Ligand Selectivity from Absolute Binding Free Energy Calculations. *J Am Chem Soc*. 2017 Jan; 139(2):946–957.
- [61] **Abel R**, Wang L, Harder ED, Berne BJ, Friesner RA. Advancing Drug Discovery through Enhanced Free Energy Calculations. *Accounts of chemical research*. 2017 Jul; 50(7):1625–1632.
- [62] **Shao H**, Shi S, Huang S, Hole AJ, Abbas AY, Baumli S, Liu X, Lam F, Foley DW, Fischer PM, Noble M, Endicott JA, Pepper C, Wang S. Substituted 4-(Thiazol-5-yl)-2-(phenylamino)pyrimidines Are Highly Active CDK9 Inhibitors: Synthesis, X-ray Crystal Structures, Structure–Activity Relationship, and Anticancer Activities. *J Med Chem*. 2013 Feb; 56(3):640–659.
- [63] **Blake JF**, Burkard M, Chan J, Chen H, Chou KJ, Diaz D, Dudley DA, Gaudino JJ, Gould SE, Grina J, Hunsaker T, Liu L, Martinson M, Moreno D, Mueller L, Orr C, Pacheco P, Qin A, Rasor K, Ren L, et al. Discovery of (S)-1-(1-(4-Chloro-3-fluorophenyl)-2-hydroxyethyl)-4-(2-((1-methyl-1H-pyrazol-5-yl)amino)pyrimidin-4-yl)pyridin-2(1H)-one (GDC-0994), an Extracellular Signal-Regulated Kinase 1/2 (ERK1/2) Inhibitor in Early Clinical Development. *J Med Chem*. 2016 Jun; 59(12):5650–5660.
- [64] **Hole AJ**, Baumli S, Shao H, Shi S, Huang S, Pepper C, Fischer PM, Wang S, Endicott JA, Noble ME. Comparative Structural and Functional Studies of 4-(Thiazol-5-yl)-2-(phenylamino)pyrimidine-5-carbonitrile CDK9 Inhibitors Suggest the Basis for Isozyme Selectivity. *J Med Chem*. 2013 Feb; 56(3):660–670.

- [65] **Berman HM**, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C. The Protein Data Bank. *Acta Crystallogr D Biol Crystallogr*. 2002 Jun; 58(Pt 61):899–907.
- [66] **Sastry GM**, 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 Mar; 27(3):221–234.
- [67] **Harder E**, Damm W, Maple J, Wu C, Reboul M, Xiang JY, Wang L, Lupyan D, Dahlgren MK, Knight JL, Kaus JW, Cerutti DS, Krilov G, Jorgensen WL, Abel R, Friesner RA. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J Chem Theory Comput*. 2016 Jan; 12(1):281–296.
- [68] **Bochevarov AD**, Harder E, Hughes TF, Greenwood JR, Braden DA, Philipp DM, Rinaldo D, Halls MD, Zhang J, Friesner RA. Jaguar: a high-performance quantum chemistry software program with strengths in life and materials sciences. *Int J Quantum Chem*. 2013; 113(18):2110–2142.
- [69] **Salvatier J**, Wiecki TV, Fonnesbeck C. Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*. 2016; 2:e55.
- [70] **Al-Rfou R**, Alain G, Almahairi A, Angermueller C, Bahdanau D, Ballas N, Bastien F, Bayer J, Belikov A, Belopolsky A, Bengio Y, Bergeron A, Bergstra J, Bisson V, Blecher Snyder J, Bouchard N, Boulanger-Lewandowski N, Bouthillier X, de Brébisson A, Breuleux O, et al. Theano: A Python framework for fast computation of mathematical expressions. *arXiv e-prints*. 2016 May; abs/1605.02688. <http://arxiv.org/abs/1605.02688>.
- [71] **Hari SB**, Merritt EA, Maly DJ. Sequence determinants of a specific inactive protein kinase conformation. *Chemistry & biology*. 2013 Jun; 20(6):806–815.
- [72] **Shah NP**, Nicoll JM, Nagar B, Gorre ME, Paquette RL, Kuriyan J, Sawyers CL. Multiple BCR-ABL Kinase Domain Mutations Confer Polyclonal Resistance to the Ty-

rosine Kinase Inhibitor Imatinib (STI571) in Chronic Phase and Blast Crisis Chronic Myeloid Leukemia. *Cancer Cell*. 2002 Aug; 2(2):117–125.

- [73] **Buczek M**, Escudier B, Bartnik E, Szczyliik C, Czarnecka A. Resistance to tyrosine kinase inhibitors in clear cell renal cell carcinoma: From the patient's bed to molecular mechanisms. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*. 2014; 1845(1):31 – 41. <http://www.sciencedirect.com/science/article/pii/S0304419X13000437>, doi: <https://doi.org/10.1016/j.bbcan.2013.10.001>.
- [74] **Huang L**, Fu L. Mechanisms of Resistance to EGFR Tyrosine Kinase Inhibitors. *Acta Pharm Sin B*. 2015; 5(5):390–401.
- [75] **Meyer SC**, Levine RL. Molecular Pathways: Molecular Basis for Sensitivity and Resistance to JAK Kinase Inhibitors. *Clin Cancer Res*. 2014; 20(8):2051–2059. doi: [10.1158/1078-0432.CCR-13-0279](https://doi.org/10.1158/1078-0432.CCR-13-0279).
- [76] **Davare MA**, Vellore NA, Wagner JP, Eide CA, Goodman JR, Drilon A, Deininger MW, O'Hare T, Druker BJ. Structural Insight into Selectivity and Resistance Profiles of ROS1 Tyrosine Kinase Inhibitors. *Proc Natl Acad Sci*. 2015; 112(39):E5381–E5390. doi: [10.1073/pnas.1515281112](https://doi.org/10.1073/pnas.1515281112).
- [77] **Van Allen EM**, Wagle N, Sucker A, Treacy DJ, Johannessen CM, Goetz EM, Place CS, Taylor-Weiner A, Whittaker S, Kryukov GV, Hodis E, Rosenberg M, McKenna A, Cibulskis K, Farlow D, Zimmer L, Hillen U, Gutzmer R, Goldinger SM, Ugurel S, et al. The Genetic Landscape of Clinical Resistance to RAF Inhibition in Metastatic Melanoma. *Cancer Discov*. 2014; 4(1):94–109. doi: [10.1158/2159-8290.CD-13-0617](https://doi.org/10.1158/2159-8290.CD-13-0617).
- [78] **Rani S**, Corcoran C, Shiels L, Germano S, Breslin S, Madden S, McDermott MS, Browne BC, O'Connell N, Crown J, Gogarty M, Byrne AT, Driscoll L. Neuromedin U: A Candidate Biomarker and Therapeutic Target to Predict and Overcome Resistance to HER-Tyrosine Kinase Inhibitors. *Cancer Res*. 2014; 74(14):3821–3833. doi: [10.1158/0008-5472.CAN-13-2053](https://doi.org/10.1158/0008-5472.CAN-13-2053).

- [79] **Holohan C**, Van Schaeybroeck S, Longley DB, Johnston PG. Cancer Drug Resistance: An Evolving Paradigm. *Nat Rev Cancer*. 2013 Sep; 13(10):714–726. doi: 10.1038/nrc3599.
- [80] **Weisberg E**, Manley PW, Cowan-Jacob SW, Hochhaus A, Griffin JD. Second Generation Inhibitors of BCR-ABL for the Treatment of Imatinib-Resistant Chronic Myeloid Leukaemia. *Nat Rev Cancer*. 2007 May; 7(5):345–356. doi: 10.1038/nrc2126.
- [81] **Y Lu X**, Cai Q, Ding K. Recent Developments in the Third Generation Inhibitors of Bcr-Abl for Overriding T315I Mutation. *Curr Med Chem*. 2011 May; 18(14):2146–2157. doi: 10.2174/092986711795656135.
- [82] **Song Z**, Wang M, Zhang A. Alectinib: A Novel Second Generation Anaplastic Lymphoma Kinase (ALK) Inhibitor for Overcoming Clinically-Acquired Resistance. *Acta Pharm Sin B*. 2015 Jan; 5(1):34–37. doi: [10.1016/j.apsb.2014.12.007](https://doi.org/10.1016/j.apsb.2014.12.007).
- [83] **Neel DS**, Bivona TG. Resistance Is Futile: Overcoming Resistance to Targeted Therapies in Lung Adenocarcinoma. *Npj Precis Oncol*. 2017 Dec; 1(1). doi: 10.1038/s41698-017-0007-0.
- [84] **Gruber F**, Hjorth-Hansen H, Mikkola I, Stenke L, TA J. A Novel BCR-ABL Splice Isoform Is Associated with the L248V Mutation in CML Patients with Acquired Resistance to Imatinib. *Leuk Off J Leuk Soc Am Leuk Res Fund UK*. 2006 Dec; 20:2057–60.
- [85] **Housman G**, Byler S, Heerboth S, Lapinska K, Longacre M, Snyder N, Sarkar S. Drug Resistance in Cancer: An Overview. *Cancers*. 2014 Sep; 6(3):1769–1792. doi: 10.3390/cancers6031769.
- [86] **Redig AJ**, Jänne PA. Basket Trials and the Evolution of Clinical Trial Design in an Era of Genomic Medicine. *American Society of Clinical Oncology*; 2015.
- [87] **Hyman DM**, Taylor BS, Baselga J. Implementing Genome-Driven Oncology. *Cell*. 2017 Feb; 168(4):584–599. doi: [10.1016/j.cell.2016.12.015](https://doi.org/10.1016/j.cell.2016.12.015).



- [88] **Pesesky MW**, Hussain T, Wallace M, Patel S, Andleeb S, Burnham CAD, Dantas G. Evaluation of Machine Learning and Rules-Based Approaches for Predicting Antimicrobial Resistance Profiles in Gram-Negative Bacilli from Whole Genome Sequence Data. *Front Microbiol.* 2016 Nov; 7. doi: [10.3389/fmicb.2016.01887](https://doi.org/10.3389/fmicb.2016.01887).
- [89] **Melnikov A**, Rogov P, Wang L, Gnirke A, Mikkelsen TS. Comprehensive Mutational Scanning of a Kinase *in Vivo* Reveals Substrate-Dependent Fitness Landscapes. *Nucleic Acids Res.* 2014 Aug; 42(14):e112–e112. doi: [10.1093/nar/gku511](https://doi.org/10.1093/nar/gku511).
- [90] **Chodera JD**, Mobley DL, Shirts MR, Dixon RW, Branson K, Pande VS. Alchemical Free Energy Methods for Drug Discovery: Progress and Challenges. *Curr Opin Struct Biol.* 2011 Apr; 21(2):150–160. doi: [10.1016/j.sbi.2011.01.011](https://doi.org/10.1016/j.sbi.2011.01.011).
- [91] **Aldeghi M**, Heifetz A, Bodkin MJ, Knapp S, Biggin PC. Accurate Calculation of the Absolute Free Energy of Binding for Drug Molecules. *Chem Sci.* 2016; 7(1):207–218. doi: [10.1039/C5SC02678D](https://doi.org/10.1039/C5SC02678D).
- [92] **Cappel D**, Hall ML, Lenselink EB, Beuming T, Qi J, Bradner J, Sherman W. Relative Binding Free Energy Calculations Applied to Protein Homology Models. *J Chem Inf Model.* 2016; 56(12):2388–2400. doi: [10.1021/acs.jcim.6b00362](https://doi.org/10.1021/acs.jcim.6b00362).
- [93] **Clark AJ**, Gindin T, Zhang B, Wang L, Abel R, Murrett CS, Xu F, Bao A, Lu NJ, Zhou T, et al. Free Energy Perturbation Calculation of Relative Binding Free Energy between Broadly Neutralizing Antibodies and the gp120 Glycoprotein of HIV-1. *Journal of molecular biology.* 2017; 429(7):930–947.
- [94] **Steinbrecher T**, Zhu C, Wang L, Abel R, Negron C, Pearlman D, Feyfant E, Duan J, Sherman W. Predicting the Effect of Amino Acid Single-Point Mutations on Protein Stability—Large-Scale Validation of MD-Based Relative Free Energy Calculations. *Journal of molecular biology.* 2017; 429(7):948–963.
- [95] **Ford MC**, Babaoglu K. Examining the Feasibility of Using Free Energy Perturbation (FEP+) in Predicting Protein Stability. *J Chem Inf Model.* 2017 Jun; 57(6):1276–1285. doi: [10.1021/acs.jcim.7b00002](https://doi.org/10.1021/acs.jcim.7b00002).

- [96] **Zou J**, Song B, Simmerling C, Raleigh D. Experimental and Computational Analysis of Protein Stabilization by Gly-to- D -Ala Substitution: A Convolution of Native State and Unfolded State Effects. *J Am Chem Soc.* 2016 Dec; 138(48):15682–15689. doi: [10.1021/jacs.6b09511](https://doi.org/10.1021/jacs.6b09511).
- [97] **Mondal J**, Tiwary P, Berne BJ. How a Kinase Inhibitor Withstands Gatekeeper Residue Mutations. *J Am Chem Soc.* 2016; 138(13):4608–4615. doi: [10.1021/jacs.6b01232](https://doi.org/10.1021/jacs.6b01232).
- [98] **Lovering F**, Aevazelis C, Chang J, Dehnhardt C, Fitz L, Han S, Janz K, Lee J, Kaila N, McDonald J, Moore W, Moretto A, Papaioannou N, Richard D, Ryan MS, Wan ZK, Thorarensen A. Imidazotriazines: Spleen Tyrosine Kinase (Syk) Inhibitors Identified by Free-Energy Perturbation (FEP). *ChemMedChem.* 2016 Jan; 11(2):217–233. doi: [10.1002/cmdc.201500333](https://doi.org/10.1002/cmdc.201500333).
- [99] **Rapp C**, Kalyanaraman C, Schiffmiller A, Schoenbrun EL, Jacobson MP. A Molecular Mechanics Approach to Modeling Protein–Ligand Interactions: Relative Binding Affinities in Congeneric Series. *J Chem Inf Model.* 2011 Sep; 51(9):2082–2089. doi: [10.1021/ci200033n](https://doi.org/10.1021/ci200033n).
- [100] **Pemovska T**, Johnson E, Kontro M, Repasky GA, Chen J, Wells P, Cronin CN, McTigue M, Kallioniemi O, Porkka K, Murray BW, Wennerberg K. Axitinib Effectively Inhibits BCR-ABL1(T315I) with a Distinct Binding Conformation. *Nature.* 2015 Feb; 519(7541):102–105. doi: [10.1038/nature14119](https://doi.org/10.1038/nature14119).
- [101] **Schrock A**, Chen TH, Clackson T, Rivera VM. Comprehensive Analysis Of The In Vitro Potency Of Ponatinib, and All Other Approved BCR-ABL Tyrosine Kinase Inhibitors (TKIs), Against a Panel Of Single and Compound BCR-ABL Mutants. *Blood.* 2013; 122(21):3992–3992.