Molecular Matchmaking: A Computational Study of the Electrostatic Interaction Between Chronic Myeloid Leukemia Drugs and Bcr-Abl Oncoprotein

Lucy Y. Liu

Advisor: Mala L. Radhakrishnan

Submitted in Partial Fulfillment of the Prerequisite for Honors in Chemistry

April 2013

© 2013 Lucy Liu
ABSTRACT

In this project, we systematically use several computational techniques such as charge optimization and component analysis to study molecular recognition and binding in the chronic myeloid leukemia (CML) drug systems. Using CML drugs and their biological target, the Bcr-Abl oncoprotein, we systematically conduct a comparative analysis on five CML drugs bound to both the wild-type (WT) and T315I mutant Abl kinase. While early generation drugs (imatinib, nilotinib, and dasatinib) interact with Thr315 via a hydrogen bond, novel drugs ponatinib and PPY-A bypass interacting with Thr315 altogether. With the mutation to Ile at position 315, early generation drugs may experience a significant loss in favorable binding due to loss of electrostatic interaction and introduction of steric hindrance. To investigate the differential binding of these drugs to the WT and mutant, we optimize each of the drugs to the Abl kinase, allowing us to study how each drug binds to the native form. We also optimize PPY-A and ponatinib to the mutant T315I, comparing this charge distribution with the one generated from optimizing to the native form. Using component analysis, we identify chemical moieties of each drug that contribute favorably or unfavorably to the electrostatic free energy of binding. Taken together, we hope that by studying CML drugs, we will gain some insight into the larger picture of electrostatic binding interaction and potentially provide future direction for rational drug design and battling drug resistance.
ACKNOWLEDGMENTS

First, I would like to thank my advisor, Professor Mala Radhakrishnan, for her unflagging support throughout my career at Wellesley. I am grateful for the three and a half years of guidance and encouragement. Her challenging, honest, and motivating evaluation of my research work has cultivated my research career and made this thesis, and a valuable experience at Wellesley, possible.

I am grateful to my lab mates, Amy Kreienkamp ’13, Connie Chen ’15, Ying Yi Zhang ’13, and Helena Qi ’14, for constantly encouraging me, believing in my endeavors to do this thesis, and providing me with laughter and fun over the summer and in lab events.

I would also like to thank my thesis committee members, Professors Don Elmore, Christopher Arumainayagam, and Martin Magid for their support in taking on this thesis project. I would especially like to acknowledge Professor Arumainayagam for showing faith in me both in my chemistry and medical school aspirations.

I would like to acknowledge my friends at Wellesley, who, throughout the four years we’ve been together, never wavered in their belief that I can aspire to do whatever I wanted. Thank you, Ji-Su Park, for suffering with me through the tough but rewarding thesis experience; for indulging in a large Upper Crust Pizza as a reward for our hard work; and for studying with me in Chicago over spring break. Thank you Ivy Wang, Sarah Pang, Irene Pang, Nian Liu, and Gina Seo for their kindness, patience, support, and encouragement throughout my years at Wellesley. Thank you especially to Ivy Wang who brought me meringues from Flour Bakery the night before my thesis was due. I’d also like to acknowledge Bryan Lee, who has been my rock this semester, helping me maintain a healthy balance between working hard and enjoying life.
I’d like to acknowledge my family – my mom, dad, twin sister, and brother – for understanding my deep dedication to research and accepting my varied interests and goals. I thank them for listening to me for hours on the phone, debating this choice and that throughout my career at Wellesley.

Finally, I’d like to thank Wellesley College for the past four years and providing me with such great opportunities to engage in rewarding research so early on in my undergraduate career. Without the funding of the Jerome A. Schiff Fellowship, National Science Foundation Research Experience for Undergraduates, and various Chemistry Department fellowships in previous years, I would not have had the extensive research experience I have today.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>6</td>
</tr>
<tr>
<td>Theory</td>
<td>16</td>
</tr>
<tr>
<td>Introduction</td>
<td>16</td>
</tr>
<tr>
<td>Electrostatics and Dielectric Theory</td>
<td>18</td>
</tr>
<tr>
<td>Charge Optimization</td>
<td>20</td>
</tr>
<tr>
<td>Component Analysis</td>
<td>24</td>
</tr>
<tr>
<td>Methods</td>
<td>25</td>
</tr>
<tr>
<td>Structure Preparation</td>
<td>25</td>
</tr>
<tr>
<td>Charge Optimization</td>
<td>27</td>
</tr>
<tr>
<td>Component Analysis</td>
<td>27</td>
</tr>
<tr>
<td>Sensitivity Analysis</td>
<td>28</td>
</tr>
<tr>
<td>Figure Generation</td>
<td>28</td>
</tr>
<tr>
<td>Results</td>
<td>28</td>
</tr>
<tr>
<td>First and Second Generation Drugs</td>
<td>29</td>
</tr>
<tr>
<td><em>Imatinib</em></td>
<td>29</td>
</tr>
<tr>
<td><em>Nilotinib</em></td>
<td>32</td>
</tr>
<tr>
<td><em>Dasatinib</em></td>
<td>35</td>
</tr>
<tr>
<td>Third Generation Drugs</td>
<td>38</td>
</tr>
<tr>
<td><em>Ponatinib</em></td>
<td>38</td>
</tr>
<tr>
<td><em>PPY-A</em></td>
<td>43</td>
</tr>
<tr>
<td>Discussion</td>
<td>49</td>
</tr>
<tr>
<td>References</td>
<td>59</td>
</tr>
</tbody>
</table>
INTRODUCTION

Currently representing over 15% of all adult leukemias, chronic myelogenous leukemia (CML) is a cancer affecting white blood cells with onset in adulthood.\textsuperscript{1} With a five-year survival rate of about 60%, CML is caused by a genetic defect in haematological stem cells (HSC).\textsuperscript{1,2} A chromosomal translocation between chromosomes 9 and 22 generates an abnormal “Philadelphia chromosome,” resulting in fusion of the \textit{breakpoint cluster (BCR)} gene and \textit{c-Abl} gene. The fused Bcr-Abl oncoprotein lacks an auto-regulatory domain that c-Abl kinase relies on, promoting unregulated tyrosine kinase activity. Abnormal expression of the oncoprotein forces the bone marrow to produce an uncontrolled proliferation of leukemic cells.

As the current frontline therapy against CML, imatinib (Gleevec or STI571; Novartis Pharma AG) is a tyrosine kinase inhibitor (TKI) and offers patients a long survival rate with potent efficacy.\textsuperscript{2,3} While treatment with imatinib successfully induces complete cytogenetic response in over 80% of patients in early stages of CML, the treatment is much less effective with patients in advanced stages.\textsuperscript{4} Unfortunately, the remaining 20% of patients may experience relapse and develop resistance to imatinib.\textsuperscript{3,5}

\textbf{Figure 1. Structure of A) Imatinib. B) Nilotinib. C) Dasatinib.} Chemical moieties are colored.
An understanding of the Abl kinase binding site and inhibitors’ binding mechanism is essential to the construction of novel CML drugs. Multiple crystallographic studies have contributed to the understanding of Abl kinase binding and drug resistance mechanisms. Bcr-Abl kinase is a tyrosine kinase, which binds to ATP and catalyzes the phosphorylation of a tyrosine hydroxy group. The Bcr-Abl kinase is activated upon interaction with a receptor on the extracellular domain of the kinase. Because the Bcr-Abl kinase phosphorylates many substrates and does not depend on ligand binding for activation, most TKIs available target the ATP binding site. Unfortunately, because Abl kinase is only one of many structurally similar tyrosine kinases, it is difficult to design a TKI that binds specifically and tightly to the correct kinase. Through high throughput screening, large compound databases, and structure-activity studies, Novartis successfully produced imatinib, a potent and competitive inhibitor of Bcr-Abl.

Imatinib is a DFG-out inhibitor, which binds only to the inactive conformation of the Abl kinase domain. In this conformation, the hydrophobic pocket in the Abl kinase domain differs from that of other tyrosine kinases. The Asp-Phe-Gly (DFG) loop in the active site changes upon inhibitor binding, inactivating the Abl kinase. For selectivity, kinase inhibitors take advantage of the different shapes and volumes of hydrophobic pockets of different kinases. Imatinib utilizes the enlarged, extensive hydrophobic pocket resulting from the inactive conformation of Abl kinase domain, making imatinib highly specific and potent against the Bcr-Abl kinase. Imatinib binds to the Abl kinase domain via multiple hydrogen bonds and van der Waals interactions, including interaction with residues Met 318, Thr 315, Glu 286, Ala 38, Ile 360, His 361, Val 299, Ala 269, Val 256, Phe 382, Leu 248, Leu370 and Phe 317.
Figure 2. Binding of imatinib (green) to Abl kinase (magenta). Potential hydrogen bonds are displayed in red dotted lines. Interacting residues are labeled and highlighted.

A major barrier to imatinib therapy is increasing drug resistance and relapse in patients.\textsuperscript{5} While the mechanisms of drug resistance can include gene amplification of the tyrosine kinase, most result from point mutations.\textsuperscript{5,11} These point mutations alter amino acid residues within the Abl domain, interfering with drug binding while maintaining ATP-binding and protein phosphorylation capabilities. Over twenty-five imatinib-resistant point mutations of the oncoprotein are known. Each mutation decreases the sensitivity of imatinib in treating CML.
patients. Figure 3 shows the most common point mutations in the Abl domain. Most of these point mutations occur in the binding interface, introducing steric hindrance or favoring the active conformation of the kinase.²

Figure 3. Common point mutations at the Abl kinase (magenta) binding interface. Binding interface residues are shown in gray and bound drug is shown in red. Commonly mutated amino acid residues are highlighted and labeled.
Several second generation TKIs have emerged based on knowledge derived from crystal structure information.\textsuperscript{11-12,13} Novartis chemists successfully modified imatinib through structural biology and rational drug design strategies, altering the metyipiperazinyl group of imatinib.\textsuperscript{14} The resulting drug, nilotinib (Tasigna or AMN107, Novartis Pharma AG), binds with higher affinity to Bcr-Abl kinase and also recognizes all mutants except the gatekeeper mutant T315I.\textsuperscript{15} Bristol-Myers Squibb Company developed another second generation TKI, dasatinib (Sprycel or BMS-354825, Bristol-Myers Squibb Company), from a different chemical scaffold based on another tyrosine kinase SRC.\textsuperscript{5} Like nilotinib, dasatinib displayed more potency and activity against imatinib-resistant Bcr-Abl mutants, but also was unable to treat patients with the T315I mutation. Unlike nilotinib, dasatinib is a DFG-in inhibitor, which binds to the active form of the Abl kinase, much like ATP or ATP-analogs.\textsuperscript{9} Figures 1b and 1c show nilotinib and dasatinib respectively.

While these second generation drugs show increased affinity to most mutations, they remain ineffective against a common mutation among patients: the binding site gatekeeper T315I and related T315A mutation.\textsuperscript{10,16} As DFG-out inhibitors, imatinib and nilotinib bind competitively at the ATP binding site. A major component of these inhibitors’ success is the key hydrogen bond formed with the side chain of threonine 315, the gatekeeper at the back of the ATP binding site. With the T315I mutation, isoleucine’s bulkier side chain blocks inhibitor entry into an adjacent hydrophobic pocket and forfeits the hydrogen bond, introducing steric hindrance and significantly reducing potency and selectivity.\textsuperscript{17}

As patients relapse and develop further resistance after treatment with these second generation TKIs, there remains a need for novel treatment that is not only potent and specific to Bcr-Abl but also active against mutant Bcr-Abl kinases including the T315I mutant. Currently,
multiple pharmaceutical companies are working to develop third generation drugs that will inhibit the Bcr-Abl kinase while maintaining the high specificity existing inhibitors have towards the tyrosine kinase.\textsuperscript{1} Among these third generation drugs, Ariad Pharmaceuticals’ ponatinib has proven potent against both native and mutant forms of Bcr-Abl and is currently undergoing phase 3 trials.\textsuperscript{1,9,18,19} In particular, to avoid the steric clash with the bulky side chain of gatekeeper residue isoleucine 315, ponatinib utilizes a linear triple bond linkage. Figure 4a shows the structure of ponatinib. By bypassing the original steric hindrance introduced by isoleucine, ponatinib is able to interact with the hydrophobic pocket while maintaining favorable hydrogen bonding interactions, including interactions with residues Glu 286, Met 318, His 361, and Ile 360 on the WT and mutant Bcr-Abl.\textsuperscript{17} Furthermore, ponatinib also distributes its binding interactions over a wide range of residues on Abl kinase’s binding interface. This ensures that single point mutations are unlikely to significantly affect affinity of the drug-target interaction.\textsuperscript{17}

\textbf{Figure 4. Structure of third generation drugs: A) Ponatinib. B) PPY-A.} Chemical moieties are colored.
Another third generation drug, a pyropyridine inhibitor PPY-A (SGX Pharmaceuticals), binds to T315I as well as the native Abl kinase.¹ Unlike ponatinib, PPY-A bypasses the steric clash of T315I gatekeeper residue by simply avoiding interaction with the hydrophobic pocket. Like dasatinib, PPY-A binds to the active conformation of the Abl kinase domain as a “DFG-in” inhibitor, rather than the inactive conformation. However, unlike dasatinib, PPY-A interacts with the glycine-rich loop in an important hydrogen bond, overcoming the T315I mutation.¹ Figure 4b shows the structure of PPY-A.

Rational drug design has aided the development of each of these CML tyrosine kinase inhibitors. Developed in the 1990s, imatinib is considered one of the first rationally designed drugs to be approved for cancer therapy.¹⁴ The discovery of oncogenes, or cancer-related genes, gave pharmaceutical companies a tangible biological target to consider when developing small molecule inhibitors.²⁰ In the case of CML, identification of the Abl kinase provided a clear target to which drugs could be optimized. Once a target molecule in a biologically relevant pathway is chosen, a lead compound is needed to begin the drug design process. Traditionally, lead compounds were the result of random screening using in vitro or in vivo experiments.¹⁹ However, recent developments in technology, including X-ray crystallography, NMR, docking tools, and computational chemistry, have increased the efficiency of drug design, enabling pharmaceutical companies to select and optimize lead compounds.¹⁹,²¹ With imatinib, Novartis was able to utilize high throughput screening to identify a derivative of phenylaminopyrimidine as a potential lead compound for structure-based design.¹⁴ Once a lead compound has been identified, various tests, both experimental and computational, can then determine the binding profile, reactivity, absorptivity, and other pharmacokinetic properties of a potential drug.
As illustrated with the various point mutations of Bcr-Abl and subsequent generations of TKI, one of the major challenges of drug design is to develop a therapy with both high affinity and a tailored specificity profile. To optimize a lead compound, an understanding of the binding mechanism and 3D structure of the target is crucial. Furthermore, understanding the determinants of binding is essential to rational drug design. For example, to treat CML, imatinib was initially developed to bind to only one out of many structurally similar cellular kinases – the Bcr-Abl kinase. Thus, imatinib must be designed to avoid off-target binding to minimize toxic side effects. Luckily, Novartis scientists were able to utilize the hydrophobic pocket unique to the Bcr-Abl kinase domain. However, other drugs, such as those used to treat HIV, must recognize multiple mutant variants to combat drug resistance. With emerging point mutations and resistance to imatinib, second and third generation inhibitors for CML not only must bind specifically to the Bcr-Abl tyrosine kinase rather than other tyrosine kinases, but they also must bind to multiple mutant variants of this Bcr-Abl tyrosine kinase. The development of such a balance between narrow and broad specificity requires successful understanding of the structural, physical, and chemical properties of both target and ligand. As such, with both second generation and third generation TKIs as excellent examples, understanding the principles of molecular binding and recognition significantly facilitates the design of novel drugs. Developed with computational aid and rational drug design strategies, ponatinib is currently in clinical trials but show the most promise. It is able to bind to both mutant and native variants of Bcr-Abl while maintaining high affinity and low toxicity.

The development of these drugs has shown that understanding the 3D structure and target binding can aid greatly in designing highly specific drugs for treatment of CML. In this study, we analyze the binding profile and structures of the CML drug systems, including all three
generations of tyrosine kinase inhibitors: imatinib, dasatinib, nilotinib, PPY-A, and ponatinib. CML drugs provide a unique framework with which we can explore the molecular determinants of binding and recognition in a biological environment. Each drug displays different levels of specificity, with imatinib being highly specific to only the native Bcr-Abl and ponatinib being promiscuous. All drugs display high affinity binding to the WT Abl kinase and, in the case of ponatinib and PPY-A, to the mutant Abl kinase.

Computation has recently proved to be a powerful tool in the design of many other drugs requiring a tailored specificity profile, including drugs targeting the rapidly mutating HIV. Among the computational methods, molecular dynamics simulations, docking studies, and studies of electrostatics have permeated the field of computational molecular design. So far, several computational studies have used molecular dynamics simulations to focus on the conformational binding of imatinib to the Bcr-Abl kinase fusion protein. As well, Lin et al. recently implemented an all-atom molecular dynamics simulation modeling solvent explicitly to calculate the absolute free energy of binding between imatinib and the oncoprotein. However, to my knowledge, no study has focused solely and consistently on electrostatic complementarity of CML drugs.

Thus, to provide further insight into the binding and specificity profile of imatinib and other CML drugs, this study uses computational methods to examine the electrostatic portion of binding free energy of CML drugs, focusing on the native and T315I gatekeeper mutant Bcr-Abl. While other components of binding, such as van der Waals interactions and hydrophobic effects, also play crucial roles in binding, electrostatics contribute greatly to specificity and affinity. In the context of drug design, electrostatics has proven to be a highly flexible handle to manipulate, improving a compound’s specificity or promiscuity. Single residues or alterations in the
charges on a functional group can significantly impact binding.\textsuperscript{34,35} Moreover, electrostatics can be easily manipulated and studied computationally. In this work, we use charge optimization, a theoretical framework developed by Lee, Kangas, and Tidor.\textsuperscript{36,37} Charge optimization is a computational technique that finds the optimal charge distribution on a drug or ligand to maximally favor electrostatic interactions. This method has been used to analyze multiple systems and to design novel therapeutics.\textsuperscript{38,39} In particular, Sims \textit{et al.} successfully utilized charge optimization to facilitate kinase inhibitor design for CDK2 and protein kinase A systems.\textsuperscript{38}

While CML drugs have been analyzed extensively using crystallography and, to an extent, molecular dynamics simulations, this work is the first to apply charge optimization techniques to CML tyrosine kinase inhibitors. Furthermore, to the best of our knowledge, no other work has examined these CML drugs comparatively with consistent computational methodology. The ability to computationally alter specific moieties of existing drugs to improve selectivity or stability can significantly facilitate the drug discovery process, potentially impacting whole populations. In our work, we optimize each of five drugs to the Bcr-Abl kinase, allowing us to study how each drug binds to the native form. We also optimize PPY-A and ponatinib to the mutant T315I, comparing this charge distribution with the one generated from optimizing to the native form. Additionally, we apply component analysis to the drug, identifying chemical moieties of the drug that contribute favorably or unfavorably to the electrostatic free energy of binding. Taken together, we hope that by studying CML drugs, we will gain some insight into the larger picture of electrostatic binding interaction and potentially provide future direction for rational drug design and battling drug resistance.
THEORY

Introduction

To conduct a comparative study of CML drugs with the native and mutant Bcr-Abl kinase, we use a metric that quantifies a drug’s electrostatic interaction with its receptor(s). In particular, we calculate the electrostatic component of the free energy of binding between two molecules in a biological environment. Consider a reversible system in which a ligand (in our case, the drug) and receptor come together to form a complex:

\[
\text{Ligand (drug) + receptor } \rightarrow \text{ complex}
\]

The Gibbs free energy between the bound and unbound states of the system gives us a handle to study the electrostatics of molecular interaction. Multiple factors contribute to the change in Gibbs free energy (\(\Delta G_{\text{total}}\)):

\[
\Delta G_{\text{total}} = \Delta G_{\text{elec}} + \Delta G_{\text{SASA}} + \Delta G_{\text{vdW}} + \Delta G_{\text{int}}
\]

\(\Delta G_{\text{elec}}\) measures the electrostatic component of Gibbs free energy. \(\Delta G_{\text{SASA}}\) is a function of the change in the system’s solvent accessible surface area. It takes into account both the loss of dispersion interactions between solute and solvent, and the solvent entropy gain upon binding. \(\Delta G_{\text{vdW}}\) measures the changes in van der Waals energy upon complex formation. Finally, \(\Delta G_{\text{int}}\) measures the change in internal conformational energy of the system and is a function of geometry of the system. Because we assume rigid binding, \(\Delta G_{\text{int}} = 0\) in our system.

In this study, we focus on the electrostatic component of Gibbs free energy, \(\Delta G_{\text{elec}}\), using a continuum electrostatic model. With previously established framework, we can computationally calculate and quantify electrostatic interactions, which play a crucial role in binding. While the most accurate method to study electrostatic interactions is to solve the Schrödinger Equation for a molecule’s electronic wave function and integrate Coulomb’s law
over the electron density as a function of time, it is impossible to solve for a system with more than one electron exactly. Furthermore, it is currently computationally infeasible to solve the Schrödinger Equation dynamically with time.

As our systems are considerably larger than a one-electron system, we approximate electrostatic interactions by assigning partial atomic charges to atoms within our molecules. Rigid binding is assumed. Rather than calculating the absolute properties of a molecule’s electron clouds, especially of those in large molecular systems, partial atomic charges model the electron distribution as single, rigid point charges that reflect their physical properties such as electronegativity. For example, as a highly electronegative atom, fluorine would be assigned a more negative partial atomic charge as it tends to pull in more electron density in a molecule. With assigned partial atomic charges, we can then obtain charge distributions for each entity in our system and model the electrostatic interactions of binding. To determine partial atomic charges for drug molecules, we use a restrained electrostatic potential (RESP) two-stage fitting method.40 This method finds the molecular electrostatic potentials at the molecule’s Merz-Kollman surface using ab initio quantum mechanical calculations. Charges are then fitted to the system using a least-squares method to reproduce the electrostatic potentials at each point.41 To account for the conformational-dependence of such fitting, RESP-fitting method introduces restraints in the form of penalty functions. The goal of such restraints is to minimize the magnitude of charges. In the two-stage fitting method, the first stage has the most degrees of freedom and weak restraints, to obtain best fits for polar areas of the molecule. The second stage holds the polar fixed charges and varies the nonpolar charges to readjust to a forced symmetry. This latter stage has strong restraints.
Electrostatics and Dielectric Theory

Even with partial atomic charges obtained, modeling electrostatic interactions between biological molecules is difficult. Consider our system of a receptor and ligand in the unbound state coming together in water to form a bound complex state. In vacuum, if our receptor possesses a positive charge, a negatively charged ligand would bind favorably to the receptor. In water, however, the ligand must overcome favorable interactions with the partially positive hydrogens of water to reach the receptor and bind to it. Similar, the receptor must overcome favorable interactions with the partially negative oxygens of water. Thus, the electrostatic interaction between the receptor and ligand is screened by the polar water.

While it is possible to use molecular dynamics simulations to account for the polarization and reorganization of each polar water molecule upon complex formation, it is extremely time consuming and computationally costly. Instead, we approximate this effect by modeling water implicitly as a dielectric continuum.

Figure 5. Schematic of our model, showing a receptor, R, and ligand, L, in water (A) and in a continuum framework (B). Receptor and ligand show partial atomic charges.
Here, we present the theory of continuum electrostatics. First, we consider systems in vacuum. With the charge distribution of each molecule obtained by RESP fitting or other parameterization, we can use the Poisson equation, the fundamental equation of electrostatics, to study electrostatic interactions.\(^{42}\)

\[ \nabla^2 \phi(r) = \frac{-\rho(r)}{\varepsilon_0} \]  

(1)

The Poisson equation relates potential \( \phi(r) \) to the charge distribution \( \rho(r) \). Both are functions of position, \( r \). The solution to the Poisson equation (1) yields Coulomb’s law when \( \rho(r) \) is a superposition of point charges:

\[ \phi(r) = \sum_i \frac{kq_i}{|r - r_i|} \]  

(2)

\( q \) is the value of the \( j^{th} \) charge. The electrostatic energy required to bring a charge to a point is then simply the potential at that point multiplied by the charge:

\[ E = q\phi \]  

(3)

Thus, the electrostatic energy of interaction between a pair of charges can be calculated as such:

\[ E = q\phi = \frac{kq_1q_2}{|r - r_i|} \]  

(4)

where \( k \) is the constant \( \frac{1}{4\pi\varepsilon_0} \), and \( q_1 \) and \( q_2 \) are respective charge values. \( \varepsilon_0 \) is the permittivity of free space constant.

However, this expression for electrostatic interactions in vacuo is impractical. Rather than existing in free vacuum, all systems of biological interest exist in an aqueous solution, which we model as a dielectric continuum. Once a medium is introduced to the system, the
electrostatic interactions are weakened, or screened, by a factor of \( D \), the dielectric constant. Thus, the Poisson equation becomes:

\[
\nabla^2 \phi(r) = \frac{-\rho(r)}{\varepsilon_0 D}
\]

The value of dielectric constants is related to the polarizability of a medium and can be found in various textbook references.

Water is assigned a dielectric constant of 80. The high dielectric constant accounts for the polarizability of water. While literature uses a range of values from 2 – 40 for proteins and small molecules, a dielectric of 4 can reproduce experimental data.\(^{43,44}\) In general, the more polarizable a medium is, the higher the dielectric constant and the stronger the solvent screening is relative to a vacuum. Thus, we can think of our system as low dielectric cavities within a high dielectric medium. With partial atomic charges assigned throughout the low dielectric molecules in our system, we can use the following form of the Poisson Equation to study electrostatic interactions in varying dielectric space:

\[
\nabla \cdot D(r) \nabla \phi(r) = \frac{-\rho(r)}{\varepsilon_0}
\]

Note that in this study, we do not take into account salt. However, mobile ions exist in our body alongside proteins and drugs. Thus, future studies can use Debye-Hückel theory to implicitly account for the additional force field sources generated by the dissolved ions.

**Charge Optimization**

In this section, we will briefly outline a previously established continuum electrostatics framework developed by Tidor et al. that allows us to study CML drugs computationally.\(^{36,37}\) Due to the polar nature of water, biological molecules interact extensively with the solvent. In the unbound state, a ligand is surrounded completely by water, which polarizes to interact
favorably with the ligand. In order to bind with the receptor, the water molecules at the binding interface sacrifice their interactions with the receptor so the ligand can interact with the receptor. Complex formation strips away favorable interactions water has with the ligand, thus entailing a ligand desolvation penalty. The receptor desolvation penalty similarly accounts for the loss of favorable interactions with water upon binding. The interaction term captures the energy of complex formation between the ligand and receptor. Figure 6 captures the idea behind these terms.

![Figure 6](image)

**Figure 6. Schematic demonstrating the ligand desolvation penalty and interaction upon complex formation.** From the unbound to bound state, the ligand’s favorable interaction with the dielectric medium (solvent) is replaced with a low dielectric cavity in the shape of the binding partner. This yields a desolvation cost. From there, the ligand binds to the receptor, replacing the low-dielectric cavity with the receptor containing partial atomic charges. The bound state (usually favorable) interaction completes this thermodynamic cycle.

Given these solvent interactions, charge optimization allows us to calculate the optimal hypothetical charge distribution of each drug to minimize the electrostatic binding free energy toward a given target molecule (recall that the more negative the free energy of binding is, the more favorable that interaction is). This technique allows us to find an appropriate balance between the desolvation penalties paid by the receptor and ligand upon binding with the favorable interaction of the complex. We can then compare the hypothetical optima to the natural
charge distribution to identify portions of the drug that could potentially be altered to improve binding. Several studies have utilized this computational technique to study small molecules and proteins\textsuperscript{36,39} and aid in the design and analysis of novel drugs\textsuperscript{38,45}.

Using the continuum electrostatics model and Poisson equation, we can express the electrostatic portion of binding as sums of matrix-vector products:

$$\Delta G_{\text{elec}} = q_l' L q_l + q_r' R q_r + q_l' C q_r$$  \hspace{1cm} (7)

The first term, $q_l' L q_l$, represents the ligand desolvation penalty. The second term $q_r' R q_r$ is the analogous receptor desolvation penalty. The third term $q_l' C q_r$ accounts for the interaction formed upon binding. $q_l$ and $q_r$ are vectors of the partial atomic charges on ligand and receptor respectively. $L$ and $R$ are the matrices of ligand and receptor desolvation unit potentials, respectively, derived from solving the Poisson Equation (once in the bound and unbound states for each charge on either partner). $L$ and $R$ describe the loss of solvent interaction with the ligand or receptor. They also account for the change in screened interaction between the charges within the ligand or receptor upon complex formation. The matrices are in units of energy per charge. Thus, to obtain desolvation energies, we multiple the matrix by the charge vector and its transpose. Finally, $C$ is the matrix of unit interaction potentials between the ligand and receptor. $C$ accounts for the screened Coulombic interactions between the ligand and receptor upon binding.

The desolvation matrices, $L$ and $R$, are positive semidefinite, which means that they have nonnegative eigenvalues. They are symmetric with dimensions equal to the number of partial atomic charges on each molecule. Each $ij$th element of the matrix captures the $\frac{1}{2}$ potential difference per unit charge between the bound and unbound states in the interaction of the $i$th and $j$th charge on the molecule. When calculating the potential of a charge interacting with its own
reaction field using linear response theory; the entropic penalty incurred is half of the unit potential. Thus, the diagonal elements of these desolvation matrices measure half of the change in potential per charge at each point $i$ in the molecule between the bound and unbound states. The factor of $\frac{1}{2}$ accounts for the entropic penalty incurred. Additionally, each off-diagonal element is divided by 2 to avoid double-counting.

$C$ is rectangular, with dimensions of $m$, the number of charges on the receptor, by $n$, the number of charges on the ligand. The $ij$th element in the complex interaction matrix measures the potential difference between the bound and unbound states in the interaction of the $i$th charge on the receptor and the $j$th charge on the ligand. $L$, $R$, and $C$ are independent of the charge distribution of the ligand and receptor and depend only on the shapes of the molecules and locations of the partial atomic charges.

Recall that we aim to find a hypothetical optimal charge distribution for our drug or ligand: $q_i$ is treated as a variable. We can treat the above $\Delta G_{\text{elec}}$ equation as a simple quadratic optimization problem:

$$\Delta G_{\text{elec}} = q_i^T L q_i + q_r^T R q_r + q_i^T C q_r$$

As with finding any maxima or minima, we can take the derivative of $\Delta G_{\text{elec}}$ with respect to $q_i$ and set it to 0. Because $L$ is positive semidefinite, a plot of $\Delta G_{\text{elec}}$ vs $q_i$ will yield an upward facing multi-dimensional paraboloid.

$$\frac{\partial \Delta G_{\text{elec}}}{\partial q_i} = 0 = 2q_{i,\text{opt}}^T L + C q_r \quad (8)$$

Thus, we can find the minimum of $\Delta G_{\text{elec}}$ by solving for $q_{i,\text{opt}}$.

$$q_{i,\text{opt}} = -0.5 L^{-1} C q_r \quad (9)$$
Finally, we can substitute $q_{l,\text{opt}}$ — the optimal charge distribution that yields a minimal $\Delta G_{\text{elec}}$ — into our Gibbs free energy equation to obtain an optimal binding free energy:

$$\Delta G_{\text{elec, opt}} = q_{l,\text{opt}}' L q_{l,\text{opt}} + q_r' R q_r + q_{l,\text{opt}}' C q_r$$  \hspace{1cm} (10)

However, the completely unconstrained optimization derived here may lead to physically unrealistic results. If the binding free energy is not very sensitive to certain charges on the drug, those charges may optimally have high charge magnitudes to achieve only minor gains in affinity. To address this, constraints are often used to yield more realistic results. In this study, we constrain the charge magnitude of each atom to be less than 1.0e. Where appropriate, we also apply various constraints to the overall monopole of the drug.

We note that, to apply charge optimization, there is no need to explicitly calculate $R$ or $C$. Instead, we calculate a $Cqr$ vector and a receptor desolvation $q_r'R q_r$ term.

**Component Analysis**

We also use component analysis to study binding in this system. Component analysis allows us to evaluate the contributions of individual chemical moieties or residues to the free energy of binding.\(^{46}\) By applying component analysis on chemical moieties or functional groups on the drug, we can potentially identify portions of the drug that could be modified to improve the electrostatics of binding. Similarly, we can apply component analysis on residues on the Abl kinase to quantify their role in binding. By setting the charges of a chemical moiety on the drug or amino acid on the receptor to zero, we can reevaluate Equation (7) to obtain a new free energy of binding with those charges set to zero; we define a $\Delta\Delta G_{\text{moiety}}$ to assess the effect of that moiety to binding.

$$\Delta\Delta G_{\text{moiety}} = \Delta G_{\text{moiety, zeroed}} - \Delta G_{\text{orig}}$$  \hspace{1cm} (11)
A positive $\Delta \Delta G_{\text{moiety}}$ indicates that the moiety contributes favorably to binding. A negative $\Delta \Delta G_{\text{moiety}}$ indicates that the moiety makes an electrostatically unfavorable contribution to binding.

**METHODS**

**Structure Preparation**

Seven crystalline structures of CML drugs bound to the wild-type (WT) or mutant structure of Abl kinase were used: Imatinib in complex with WT Abl kinase (PDB ID 2HYY\textsuperscript{2}), Nilotinib in complex with WT Abl kinase (3CS9\textsuperscript{15}), Dasatinib in complex with WT Abl kinase (2GQG\textsuperscript{12}), Ponatinib in complex with WT Abl kinase (3OXZ\textsuperscript{17}), Ponatinib in complex with T315I mutant Abl kinase (3IK3\textsuperscript{17}), PPY-A in complex with WT Abl kinase (2QOH\textsuperscript{1}), and PPY-A in complex with T315I mutant Abl kinase (2Z60\textsuperscript{1}). Table 1 summarizes the structures used, along with resolutions and number of water molecules retained. All seven structures contained multiple complexes in the unit cell. The complex with the least missing density close to the binding site were chosen. Residues adjacent to missing density regions were patched with methylamide groups for N-terminals and acetamide groups for C-terminals.

<table>
<thead>
<tr>
<th>PDB ID</th>
<th>resolution (Å)</th>
<th>water molecules retained</th>
</tr>
</thead>
<tbody>
<tr>
<td>2HYY\textsuperscript{2}</td>
<td>2.4</td>
<td>2</td>
</tr>
<tr>
<td>3CS9\textsuperscript{15}</td>
<td>2.21</td>
<td>13</td>
</tr>
<tr>
<td>2GQG\textsuperscript{12}</td>
<td>2.4</td>
<td>20</td>
</tr>
<tr>
<td>3OXZ\textsuperscript{17}</td>
<td>2.2</td>
<td>15</td>
</tr>
<tr>
<td>3IK3\textsuperscript{17}</td>
<td>1.9</td>
<td>24</td>
</tr>
<tr>
<td>2QOH\textsuperscript{1}</td>
<td>1.95</td>
<td>20</td>
</tr>
<tr>
<td>2Z60\textsuperscript{1}</td>
<td>1.95</td>
<td>19</td>
</tr>
</tbody>
</table>

*Table 1. Structures used from Protein Data Bank.*
To prepare the structure for use in our study, we eliminated all water molecules that had less than three potential hydrogen bonding interactions or were more than 3.3 Å away from either binding partner. Remaining explicit waters were assigned to the receptor during calculations. Explicit ions were also eliminated. After careful examination, the amide group on the side chain of asparagine and glutamine were flipped as necessary based on any potential hydrogen bonding interactions in the proximity. As well, the imidazole group of each histidine was flipped as necessary; each histidine was also assigned as the epsilon or delta tautomer based on potential hydrogen bonds with surrounding residues or drug. In complex 2GQG, a phosphotyrosine was converted to a regular tyrosine residue.

To obtain partial atomic charges for the drug molecules, the geometry of each drug molecule was first optimized with Gaussian 2003\textsuperscript{47} using Hartree-Fock theory and 6-31G basis set. The electronic wave functions and electrostatic potentials at the Merz-Kollman surface were obtained using Hartree-Fock level theory and 6-31G* basis set.\textsuperscript{48} The resulting electrostatic potentials were then used in a two-stage fitting RESP method\textsuperscript{40} to obtain charge distribution for drug structures.

Hydrogen atoms were incorporated into each structure using CHARMM’s\textsuperscript{49} hydrogen-building (HBUILD) facility.\textsuperscript{50} This study used the CHARMM22 force field and parameter set.\textsuperscript{41} CHARMM22 atom types were assigned to each atom in the drugs. Because several atoms in 2QOH, 2GQG, and 3OXZ lacked parameters, alternative but similar atom types were assigned. Because their positions were fixed during structure preparation and calculations, assigning an alternative atom type is not expected to have significant effects in the structure preparation process.
**Charge Optimization**

To obtain all matrix and vector components necessary for charge optimization (described in the theory section), a finite difference solver\(^{51}\) was used to solve the Poisson Equation. The \(L\) matrix, \(Cqr\) vector, and receptor desolvation penalty \((q_r’ R q_r)\) were obtained for each complex. To solve the Poisson Equation, a cubic grid ranging from \((193 \times 193 \times 193)\) to \((201 \times 201 \times 201)\) grid points was used to maintain a consistent grid resolution of 6 grids/Å for each complex. Three stages of focusing were used to solve the Poisson Equation: the structure first occupied 23% of the grid, followed by 92% of the grid, and finally 184% of the grid. The structure was centered at the binding interface. PARSE radii were used for all atoms; Parm99 AMBER van der Waals radii were used for F atoms.\(^{52}\) The dielectric boundary of the protein and drug was determined using a 1.4 Å probe sphere.\(^{53}\) A dielectric constant of 4 was used for proteins and small molecules.

The General Algebraic Modeling System\(^{54}\) was used to conduct all charge optimization calculations subject to constraints. In all instances, charge magnitudes were constrained to be 1.00e. Because all drugs studied naturally have an overall charge of 0, the drug monopole was constrained to be 0 to enable a direct comparison. Follow-up optimizations with additional non-zero total monopole constraints were performed as well.

**Component Analysis**

To quantify the contributions of certain moieties of the drug to electrostatic portion of binding, we can set the atoms of that moiety to zero and calculate the free energy of binding. In equation (7), \(qL\) elements corresponding to a moiety's atoms were set to zero and the free energy of binding was reevaluated. Matlab (The Mathworks, Inc., Natick, MA) was used to execute drug component analysis calculations.
In applying component analysis to the receptor's residues, because we did not calculate explicit $R$ matrix elements, we set the charges on the residue to zero; the binding free energy was reevaluated.

**Sensitivity Analysis**

Diagonal $L$ matrix elements approximately represent the sensitivity of the free energy of binding to changes in charge on that atom.\textsuperscript{55} These diagonal elements were extracted; radii sizes of atoms were scaled accordingly.

**Figure Generation**

All figures were generated using VMD\textsuperscript{56} and compiled using Adobe Illustrator CS5.1.

**RESULTS**

In this study, charge optimization, component analysis, and sensitivity analysis were used to probe the electrostatic determinants of binding in CML drug complexes. This section presents the data obtained from these computational techniques. For each drug studied, three types of figures are shown. Figures 7a, 9a, 12a, 15a, and 19a show the structure of the drug colored by atom type and labeled to facilitate discussion of the data. Figures 7b, 9b, 12b, 15b, and 19b depict results of component analysis, with the chemical moieties separated by pink bonds. For ponatinib and PPY-A, part c of the figure shows component analysis results of the mutant Abl kinase complex. Chemical moieties are labeled in parentheses to facilitate discussion of moieties. Key residues interacting via hydrogen bonds are shown in yellow.

Figures 2, 10, 13, 16, and 20 also depict the binding interface and key interactions with Abl kinase for each drug. Drugs are highlighted in green and colored by atom type; potential hydrogen bonds are displayed in red dotted lines. Interacting residues are also labeled.
Charge optimization and sensitivity analysis results are shown in figures 8, 11, 14, 17, 18, 21, and 22. Part a of each panel shows the drug’s optimal charge distribution. Part b of each panel shows the difference in optimal and natural charges on the drug.

**First and Second Generation Drugs**

*Imatinib*

![Diagram of Imatinib](image)

**Figure 7. Structure of imatinib and component analysis show favorable and unfavorable moieties of the drug.** A) Structure of imatinib colored by atom type. Atoms are labeled. B) Component analysis was conducted on each chemical moiety (separated by pink bonds and labeled with numbers). Favorability of moiety is indicated with color and intensity. White represents neutral moieties, red represents unfavorable moieties, and blue represents favorable moieties. $\Delta\Delta G_{moiety}$ are labeled.

Component analysis (figure 7b) shows that much of imatinib contributes favorably to binding (~1 kcal/mol to 2 kcal/mol). Moiety 1, or the methylpiperazine group, contributes most
unfavorably to binding, with a $\Delta \Delta G_{\text{moiety},1} = -1.968$ kcal/mol. Groups 3, 5, and 7 all form hydrogen bonds with residues on the receptor Abl kinase (refer to Figure 2). However, only groups 3 and 7 contribute favorably to binding ($\Delta \Delta G_{\text{moiety},3} = 1.793$ kcal/mol, $\Delta \Delta G_{\text{moiety},7} = 1.968$ kcal/mol). While the amine (5) does form a hydrogen bond with the side chain of Thr315, it is slightly unfavorable to binding ($\Delta \Delta G_{\text{moiety},5} = -0.6815$ kcal/mol).

**Figure 8.** Charge optimization results and sensitivity analysis of imatinib in complex with WT Abl. A) Optimal charge distribution of imatinib when bound to WT Abl. Charges range from -1.0e (red) to 1.0e (blue). Hydrophobic atoms are colored in white. Total charge on drug is constrained to be 0. B) Charge differences of imatinib bound to WT between optimal and natural charges. Red indicates atoms that are too positive and would be more optimal if they were more negative. Blue indicates atoms that are too negative and would be more optimal if they were more positive. Radii in figures A and B correspond to sensitivity of free energy to changes in charge on that atom, with a larger radius indicating greater sensitivity.
Figure 8a shows the optimal charge distribution of imatinib in complex with WT Abl kinase. Upon charge optimization, the free energy of binding improved by 10.39 kcal/mol. Refer to Table 2 for all binding calculations and optimization results. The drug is largely hydrophobic at its optimal charge distribution (figure 8a). However, a few atoms on imatinib are highly charged when optimal. Figure 8b shows the difference between optimal and natural charge distributions. Interestingly, both the methylpiperazine group and the aromatic group (7) have high magnitude charges when optimal. The methylpiperazine’s natural charge distribution is far from optimal while group 7 is closer to optimal. In particular, one of the hydrogens (H62) on the methyl group of moiety 1 would be charged when optimal. Furthermore, N29 in that same moiety is positive in its optimal charge distribution, rather than negative; thus, the charge on N29 is far from optimal.

In group 7, N5 has a highly negative partial atomic charge and is already optimal. C11, which is sp² hybridized and bound to three nitrogens, is originally positive. However, in its optimal distribution, those atoms (C11, N10, N12, N13) are hydrophobic. Finally, the amide group (moiety 3, atoms N20, C21, O30, H41) that interacts with the side chain carbonyl of Glu286 and backbone NH of Asp381 via hydrogen bonds is highly optimal (figure 8b). The free energy of binding show high sensitivity to changes in the partial charges of the atoms directly interacting in the hydrogen bond (H41 and O30) (figure 8).

While these figures show charge optimization of imatinib subject to the constraint that the drug monopole be 0.0e, an unconstrained optimization found that the optimal monopole for the drug to be close to 1.67e. The free energy of binding improved by 12.726 kcal/mol ($\Delta G = 1.092$ kcal/mol), which yielded an additional gain of 2.336 kcal/mol from the constrained optimization. The total charge of the optimal drug is quite high and warrants further
investigation. Table 2 at the end of the results section lists all optimizations and their free energy calculation results.

*Nilotinib*

We applied a similar computational analysis to nilotinib, another drug developed by Novartis.

![Figure 9. Structure of nilotinib and component analysis show favorable and unfavorable moieties of the drug.](image)

A) Structure of nilotinib colored by atom type. Atoms are labeled. B) Component analysis was conducted on each chemical moiety (separated by pink bonds). Favorability of moiety indicated with color and intensity. White represents neutral moieties, red represents unfavorable moieties, and blue represents favorable moieties. ΔΔG moiety are labeled.

Nilotinib makes hydrogen bonds with the same key residues that imatinib does (figure 10 below shows the binding site interactions with Abl kinase). Component analysis results (figure 9b) show that no moiety of nilotinib is particularly unfavorable. All unfavorable moieties have...
\( \Delta G_{\text{moiety}} \) magnitudes of less than 0.5 kcal/mol. Moreover, the amide group (moiety 5) contributes very favorably to binding (\( \Delta G_{\text{moiety,5}} = 2.334 \) kcal/mol). No other chemical moiety contributes as strongly to binding. This amide group, like the amide in imatinib, interacts with both Glu286 and Asp381. The pyridine N (moiety 1) interacts via a hydrogen bond with the backbone of Met318 (figure 10).

Figure 10: Binding of nilotinib (green) to Abl kinase (magenta). Potential hydrogen bonds are displayed in red dotted lines. Interacting residues are labeled and highlighted.
Figure 11. Charge optimization results and sensitivity analysis of nilotinib in complex with WT Abl. A) Optimal charge distribution of nilotinib when bound to WT Abl. Charges range from -1.0e (red) to 1.0e (blue). Hydrophobic atoms are colored in white. Total charge on drug is constrained to be 0. B) Charge differences of nilotinib bound to WT between optimal and natural charges. Red indicates atoms that are too positive and would be more optimal if they were more negative. Blue indicates atoms that are too negative and would be more optimal if they were more positive.

Nilotinib’s charge distribution was optimized to bind to WT Abl kinase. Figure 11a shows nilotinib’s optimal charge distribution, which yielded an improvement in binding of 4.547 kcal/mol. Figure 11a shows that most of the drug is hydrophobic when optimized. While C37, N38, and C36 of moiety 1 are polar when optimal, they are already quite close to their optimal partial atomic charges (figure 11b).

Likewise, the amide (moiety 5) also contains high magnitude partial atomic charges (atoms H60, N3, C2, O1). These atoms interact via hydrogen bonds with the side chain carbonyl
of Glu286 and backbone NH of Asp381. Figure 11b shows that the amide’s natural charges are optimal. Our sensitivity analysis shows that the free energy of binding is highly sensitive to change in charge on the carbonyl of the amide.

Interestingly, moieties 2 (pyrimidine) and 3 (amine) are the furthest from optimal. While the hydrogen of amine, H61, that interacts with the side chain OH of Thr315 is close to optimal, N27 and C28 is entirely hydrophobic in the optimal charge distribution. N27 and C28 are almost 1.0e far from optimal.

An unconstrained optimization found the optimal monopole for the drug to be 1.60e. This yielded an improvement in the free energy of binding of 6.971 kcal/mol, which is a 2.424 kcal/mol gain from the constrained optimization.

**Dasatinib**

![Figure 12. Structure of dasatinib and component analysis show favorable and unfavorable moieties of the drug. A) Structure of dasatinib colored by atom type. Atoms are labeled. B) Component analysis was conducted on each chemical moiety (separated by pink bonds). Favorability of moiety indicated with color and intensity. White represents neutral moieties, red represents unfavorable moieties, and blue represents favorable moieties. ΔΔG_moiety are labeled.](image)
Dasatinib interacts with both Thr315 and Met318 via hydrogen bonding interactions. Figure 13 (below) shows key hydrogen bonding interactions; Met318 makes two hydrogen bonds with dasatinib, suggesting that Met318 is a major contributor to binding. Component analysis (figure 12b) shows that moiety 4, which interacts with Met318, contributes very favorably to binding ($\Delta\Delta G_{\text{moiety,4}} = 1.428$ kcal/mol). We analyze Met318 in more detail in the discussion section. However, like with imatinib and nilotinib, the amide group that makes a hydrogen bond with Thr315 contributes unfavorably to binding ($\Delta\Delta G_{\text{moiety,5}} = -2.103$ kcal/mol).

Figure 13. Binding of dasatinib (green) to Abl kinase (magenta). Potential hydrogen bonds are displayed in red dotted lines. Interacting residues are labeled and highlighted.
Figure 14. Charge optimization results and sensitivity analysis of dasatinib in complex with WT Abl. A) Optimal charge distribution of dasatinib when bound to WT Abl. Charges range from -1.0e (red) to 1.0e (blue). Hydrophobic atoms are colored in white. Total charge on drug is constrained to be 0. B) Charge differences of dasatinib bound to WT between optimal and natural charges. Red indicates atoms that are too positive and would be more optimal if they were more negative. Blue indicates atoms that are too negative and would be more optimal if they were more positive.

After charge optimization (figure 14), the free energy of binding showed an improvement of 5.57 kcal/mol. Just as with the two previous early generation drugs, when the monopole of the drug is constrained to 0.0e, the drug’s optimal charge distribution is largely hydrophobic. Only the aminothiazole group (moiety 4), parts of the piperazine group (moiety 2), and the hydroxyl group of moiety 1 are polar. Figure 14b shows that the aminothiazole, which interacts with Met318 via two hydrogen bonds, is already very close to its optimal charge distribution. As well, the piperazine group is close to optimal. Moreover, our sensitivity analysis suggests that the free energy of binding is highly sensitive to changes in charge on these atoms.
Component analysis results (figure 12) indicate that the amide group (moiety 5, O27, C3, N22, H37) contributes very unfavorably to binding. This seems to contradict the idea that this is a key interaction in the binding of dasatinib to Abl kinase. However, the partial atomic charges on the amide are fairly optimal (figure 14b), suggesting that nearly any other charge distribution on the amide would be even more unfavorable. This is discussed in more detail later.

Additionally, the free energy of binding is slightly sensitive to changes in charge on atoms H37 and O27 of the amide. However, the hydroxyl group in moiety 1 is furthest from its optimal charge distribution. In fact, atoms O33 and H56 are oppositely charged from its natural charges when optimized: when optimal, O33 is slightly positive, while H56 is negatively charged.

An unconstrained optimization found the optimal monopole for the drug to be 6.6e. The free energy of binding improved by 16.608 kcal/mol (table 2). This represents an additional gain of 11.018 kcal/mol from the constrained optimization. The rather high magnitude of the optimal monopole is surprising and seems physically unrealistic, even given the constraints applied. This will be further addressed in the discussion section.

**Third Generation Drugs**

In this section, we present data on two third generation drugs that bind to both the WT and T315I mutant.

*Ponatinib*

Ponatinib, a novel therapeutic drug developed by Ariad Pharmaceuticals, shows promise in battling drug resistance conferred by T315I mutant.
Figure 15. Structure of ponatinib and component analysis show favorable and unfavorable moieties of the drug to WT and mutant Abl kinase. A) Structure of ponatinib colored by atom type. Atoms are labeled. B) Component analysis was conducted on each chemical moiety (separated by pink bonds) of ponatinib while bound to WT. C) Component analysis was conducted on each chemical moiety of ponatinib while bound to T315I mutant. Favorability of moiety indicated with color and intensity. White represents neutral moieties, red represents unfavorable moieties, and blue represents favorable moieties. ΔΔG<sub>moiety</sub> are labeled.

Like earlier generation drugs, ponatinib interacts with a few key residues on the Abl binding interface (see Figure 16 below). However, ponatinib bypasses any interaction with Thr315 or Ile315. Moiety 6 (acetylene) is closest to residue 315 in both WT and mutant complexes. Component analysis results (figures 15b, 15c) on moiety 6 show that it does not contribute significantly favorably or unfavorably to binding (|ΔΔG<sub>moiety,6</sub>| < 1 kcal/mol).

Component analysis shows us that moiety 1 (methylpiperazine) is highly unfavorable to binding in both WT and mutant complexes: ΔΔG<sub>moiety,1,WT</sub> = -1.749 kcal/mol; ΔΔG<sub>moiety,1,T315I</sub> = -1.880 kcal/mol. In both the WT and mutant complex, two groups contribute significantly favorably to binding: moiety 4 and 7. Like imatinib and nilotinib, the amide group, or moiety 4, interacts via two hydrogen bonds with Glu286 and Asp381 (figure 16). Component analysis reveals that this amide group contributes favorably (over 2 kcal/mol) in both mutant and WT complex. Similarly,
moiety 7, which interacts with the backbone of Met318, contributes favorably (over 1 kcal/mol) in both mutant and WT complex.

Figure 16. Binding of ponatinib (green) to A) WT Abl kinase (magenta) and B) T315I mutant Abl kinase. Potential hydrogen bonds are displayed in red dotted lines. Interacting residues are labeled and highlighted. Mutated residue is shown in black.
Figure 17. Charge optimization results and sensitivity analysis of ponatinib in complex with WT Abl. A) Optimal charge distribution of ponatinib when bound to WT Abl. Charges range from -1.0e (red) to 1.0e (blue). Hydrophobic atoms are colored in white. Total charge on drug is constrained to be 0. B) Charge differences of ponatinib bound to WT between optimal and natural charges. Red indicates atoms that are too positive and would be more optimal if they were more negative. Blue indicates atoms that are too negative and would be more optimal if they were more positive.

Ponatinib was optimized to bind maximally to the WT Abl kinase. Charge optimization yielded an improvement of 10.139 kcal/mol. With the monopole of the drug constrained to 0.0e, several parts of ponatinib are optimal when hydrophobic. In particular, moieties 5 and 6 are both hydrophobic and already close to optimal in its natural charge distribution. The methylpiperazine group (moiety 1) has high magnitude partial atomic charges in its optimal charge distribution (indicated by the white and blue colored atoms in Figure 17a). However, in addition to contributing unfavorably to binding, moiety 1 is also far from its optimal charge distribution. This suggests that moiety 1 can be improved to contribute less unfavorably to binding.
Furthermore, the free energy of binding is highly sensitive to changes in charge on the atoms of the methyl group in moiety 1. Moiety 3 (atoms O28, C27, N29, and H56) interacts with both Asp381 and Glu286, just as earlier generation drugs do. The amide group is highly polar and already close to optimal. Similarly, moiety 7’s nitrogen (N1) is highly charged when optimized. Figure 17b shows that N1 is already optimal; N1 interacts with the backbone NH of Met318.

An unconstrained optimization yielded an optimal monopole of the drug to be 1.6e kcal/mol when bound to the WT, improving the free energy of binding by 12.039 kcal/mol (table 2). This yields an additional gain in favorability of 1.9 kcal/mol.

Figure 18. Charge optimization results and sensitivity analysis of ponatinib in complex with T315I mutant Abl. A) Optimal charge distribution of ponatinib when bound to T315I mutant Abl. Charges range from -1.0e (red) and 1.0e (blue). Hydrophobic atoms are colored in white. Total charge on drug is constrained to be 0. B) Charge differences of ponatinib bound to mutant Abl between optimal and natural charges. Red indicates atoms that are too positive and would be more optimal if they were more negative. Blue indicates atoms that are too negative and would be more optimal if they were more positive.
Ponatinib was also optimized to bind with maximal tightness to T315I mutant Abl kinase. Charge optimization yielded an improvement of 8.83 kcal/mol (figure 18). Like ponatinib bound to the WT Abl kinase, much of the drug is hydrophobic when optimized. All except moieties 1 and 4 show an optimal distribution of mostly hydrophobic atoms (figure 18a). The methyl-piperazine group (moiety 1) contributes unfavorably to binding and shows an optimal charge distribution with charged atoms, especially atoms N39 and C15, both of which indicate optimal charges of close to positive 1.0e. These two atoms are also very far from their optimal charge distribution (figure 18b).

Another optimally charged group is moiety 4, the amide (atoms O28, C27, N29, and H56). This group interacts with the side chains of Asp381 and Glu286 and shows an optimally polar charge distribution. Sensitivity analysis shows that the free energy of binding is highly sensitive to changes in O28’s charge. Figure 18b shows that moiety 4 is already quite optimal.

Moiety 7’s N1 atom, which interacts with the backbone of Met318, is very polar and fairly close to being optimally charged. Finally, atom C33 in moiety 3 is quite far from optimal (indicated by the intense red color in Figure 18b). In its optimal charge distribution, moiety 3 is overall hydrophobic.

An unconstrained optimization was also applied to this system, yielding an optimal monopole of the drug to be 2.02e. The free energy of binding was improved by 12.781 kcal/mol (table 2), yielding an additional gain of 3.951 kcal/mol.

**PPY-A**

PPY-A is a pyrypyridine inhibitor, binding to both the T315I mutant and WT Abl kinase (Figure 20 shows the binding site and key interacting residues for both). It is also worth noting that PPY-A binds to the active form of the Abl kinase, like dasatinib.
Like ponatinib, PPY-A bypasses interacting with residue 315 in both the WT and mutant complexes. Interestingly, almost all moieties contribute favorably to binding in PPY-A to both mutant and WT Abl kinase (figures 19b, 19c). The favorable moieties all form hydrogen bonds with at least one residue on Abl kinase (figure 20). Moiety 5, or the ether group, contributes slightly unfavorably to binding, but with $|\Delta \Delta G_{\text{moiety},6}| < 1$ kcal/mol, we cannot say that it is significantly unfavorable. Between the WT and mutant complex, all moieties except moiety 4 contribute more significantly in the mutant complex than the WT complex. For example, $\Delta \Delta G_{\text{moiety,1},\text{WT}} = 0.528$ kcal/mol, while $\Delta \Delta G_{\text{moiety,1},\text{mutant}} = 1.089$ kcal/mol; $\Delta \Delta G_{\text{moiety,2},\text{WT}} = 0.199$ kcal/mol, while $\Delta \Delta G_{\text{moiety,2},\text{mutant}} = 0.919$ kcal/mol. PPY-A’s methyl-ether group, or moiety 5, contributes more unfavorably in the WT than the mutant ($\Delta \Delta G_{\text{moiety,5},\text{WT}} = -0.229$ kcal/mol; $\Delta \Delta G_{\text{moiety,5},\text{mutant}} = -0.0775$ kcal/mol).
Figure 20. Binding of PPY-A (green) to A) WT Abl kinase (magenta) and B) T315I mutant Abl kinase (blue). Potential hydrogen bonds are displayed in red dotted lines. Interacting residues are labeled and highlighted.
Figure 21. Charge optimization results and sensitivity analysis of PPY-A in complex with WT Abl. A) Optimal charge distribution of PPY-A when bound to WT Abl. Charges range from -1.0e (red) to 1.0e (blue). Hydrophobic atoms are colored in white. Total charge on drug is constrained to be 0. B) Charge differences of PPY-A bound to WT between optimal and natural charges. Red indicates atoms that are too positive and would be more optimal if they were more negative. Blue indicates atoms that are too negative and would be more optimal if they were more positive.

PPY-A was optimized to bind to WT Abl kinase. Charge optimization yielded an improvement of 2.85 kcal/mol (figure 21). While much of the drug is optimal when hydrophobic, the moiety 1’s hydrogens (H46, H47, H48) are very polar in its optimal charge distribution. They are also far from optimal. This suggests that this portion of the drug can potentially be improved. Another notable polar section of the drug is atoms N1 and C6 in moiety 3. As part of the double-ring aromatic group, these atoms are naturally close to their optimal charge distribution. N1 and H37 interact via hydrogen bonds with residues Met318 and Glu316, respectively.
An unconstrained optimization was also applied to this system, yielding an optimal monopole of the drug to be $3.61e$. This optimization improved the free energy of binding by $8.769 \text{ kcal/mol}$ (table 2), which is an additional gain in favorability of $5.919 \text{ kcal/mol}$.

Figure 22. Charge optimization results and sensitivity analysis of PPY-A in complex with T315I mutant Abl. A) Optimal charge distribution of PPY-A when bound to mutant Abl. Charges range from $-1.0e$ (red) to $1.0e$ (blue). Hydrophobic atoms are colored in white. Total charge on drug is constrained to be 0. B) Charge differences of PPY-A bound to mutant Abl between optimal and natural charges. Red indicates atoms that are too positive and would be more optimal if they were more negative. Blue indicates atoms that are too negative and would be more optimal if they were more positive.

Similarly, PPY-A was optimized to bind to T315I mutant Abl kinase. This yielded an improvement in the free energy of binding to be $2.99 \text{ kcal/mol}$. When compared to the WT complex, more of PPY-A is polar when optimized to bind to the mutant Abl kinase. Two moieties show a difference between the optimal and natural charge distribution: moieties 1 and 2. The methyl group of moiety 1, again (H46, H47, H48, C27), proved to be polar when optimized.
The carbonyl group of moiety 1 (C23, O24) is also extremely polar. The carbonyl interacts with the backbone of Gly249 in a hydrogen bond and is moderately far from optimal. Moiety 2, the pyrimidine, also contains highly charged atoms in its optimal charge distribution. Atoms C20 and C21 are more positive and negative, respectively, in their optimal charge distribution and are far from their optimal charges.

An unconstrained optimization was also applied to this system, yielding an optimal monopole of the drug to be 2.62e. The free energy of binding improved by 7.797 kcal/mol (table 2). This is an additional gain of 4.807 kcal/mol from the constrained optimization.

Both the WT and mutant complexes interact with the same residues via hydrogen bonds and have no apparent interactions with residue 315. The portions of the drug near residue 315 (moieties 4 and 5) are already highly optimal.

<table>
<thead>
<tr>
<th>PDB ID</th>
<th>ΔG original</th>
<th>ΔG optimal (total monopole = 0)</th>
<th>ΔG optimal (total monopole = 1)</th>
<th>ΔG optimal (total monopole = 2)</th>
<th>ΔG optimal (total monopole = 3)</th>
<th>ΔG optimal (total monopole = 6)</th>
<th>ΔG optimal (total monopole = 7)</th>
<th>ΔG optimal (no total constraint)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2HYY</td>
<td>13.818</td>
<td>3.375</td>
<td>1.457</td>
<td>1.144</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>1.092</td>
</tr>
<tr>
<td>3CS9</td>
<td>9.064</td>
<td>4.517</td>
<td>2.469</td>
<td>2.267</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>2.093</td>
</tr>
<tr>
<td>2GQG</td>
<td>8.16</td>
<td>2.570</td>
<td>-0.111</td>
<td>-2.531</td>
<td>-4.629</td>
<td>-8.338</td>
<td>-8.399</td>
<td>-8.448</td>
</tr>
<tr>
<td>3OXZ</td>
<td>12.461</td>
<td>2.322</td>
<td>0.667</td>
<td>0.494</td>
<td>1.305</td>
<td>n/a</td>
<td>n/a</td>
<td>0.422</td>
</tr>
<tr>
<td>3IK3</td>
<td>11.69</td>
<td>2.842</td>
<td>-0.059</td>
<td>-1.091</td>
<td>-0.163</td>
<td>n/a</td>
<td>n/a</td>
<td>-1.091</td>
</tr>
<tr>
<td>2QOH</td>
<td>6.769</td>
<td>3.919</td>
<td>1.436</td>
<td>-0.452</td>
<td>-1.724</td>
<td>n/a</td>
<td>n/a</td>
<td>-2.000</td>
</tr>
<tr>
<td>2Z60</td>
<td>6.497</td>
<td>3.531</td>
<td>1.041</td>
<td>-0.842</td>
<td>-1.129</td>
<td>n/a</td>
<td>n/a</td>
<td>-1.300</td>
</tr>
</tbody>
</table>

Table 2. Free energy binding calculations of optimized structures subject to various total monopole constraints (all values in kcal/mol). All structures were optimized subject to total monopole constraints of 0, 1e, and 2e. Drugs with optimal overall monopoles of greater than 2e were optimized with greater monopole constraints as well.
DISCUSSION

In this study, we systemically applied the computational techniques of charge optimization and component analysis to probe the electrostatic interaction between chronic myeloid leukemia drugs and their biological target, the Abl kinase. We comparatively analyzed three DFG-out inhibitors, imatinib, nilotinib, and ponatinib, which bind to the inactive conformation of the Abl kinase. We also analyzed two DFG-in inhibitors, dasatinib and PPY-A, which bind to the active conformation of the Abl kinase. Of these drugs, only ponatinib and PPY-A remain potent against the T315I mutant Abl kinase. While this study focused on the electrostatic interactions, it is worth noting that each drug forms key van der Waals and aromatic interactions with the Abl kinase in the hydrophobic binding pocket as well.

DFG-out inhibitors imatinib, nilotinib, and ponatinib form hydrogen bonds with residues Glu286, Met318, and Asp381 on the Abl kinase (figures 2, 10, 16). The amide moieties on each drug that interact with Asp381 and Glu286 are already quite optimal to binding in their natural charge distributions (figures 8, 11, 17, 18), suggesting that these are key moieties to conserve when generating novel drugs for CML. This amide moiety contributes favorably to binding in all three drugs.

To further investigate the role of these key residues in the interaction with CML drugs, we also performed component analysis on the Abl residues Glu286, and Asp381 for each CML drug-Abl kinase complex. While component analysis on Glu286 did not reveal any significantly favorable or unfavorable contribution to binding, Asp381 yielded an unfavorable $|\Delta G_{381}|$ of greater than 1 kcal/mol for all DFG-out drugs. Asp381 forms a hydrogen bond with the amide carbonyl with all three drugs, intuitively suggesting that Asp381 might contribute favorably to binding. Upon closer examination, we find that Asp381 pays a heavy desolvation penalty in
DFG-out inhibitors (2.33 – 3.64 kcal/mol). Thus, when rendered into its hydrophobic isostere, while the interaction with each drug’s amide carbonyl is weakened, the dampened desolvation penalty of a hydrophobic residue 381 improves the overall free energy of binding. However, it is worth noting that while Asp381 contributes unfavorably to binding due to its desolvation penalty, the amide moiety it interacts with contributes favorably to the binding reaction. Thus, though the receptor’s contribution might be electrostatically unfavorable, the optimal nature of the drug suggests that any other drug moiety or charge distribution could be even more unfavorable. Additionally, this interaction could possibly contribute to the specificity of this drug.57

DFG-in inhibitors dasatinib and PPY-A bind in the more flexible active conformation of the Abl kinase. Interestingly, dasatinib and PPY-A interact with different residues on the Abl kinase (figures 13 and 20). Dasatinib interacts heavily with the backbone of Met318 via two hydrogen bonds through the N of the thiazole group and NH of the amino group. Similarly, PPY-A forms key hydrogen bonds with Met318 and other residues on the Abl kinase. Specifically, the two nitrogen atoms in the pyrrolopyridine core of PPY-A (moiety 3 in figure 19) interact with the backbone NH of Met318 and carbonyl O of Glu316. Furthermore, the N of the pyridine group (moiety 2) forms a hydrogen bond with the side chain of Asn322; finally, the carbonyl O of dimethylaminocarbonyl group (moiety 1) interacts with the backbone N of Gly249 in the glycine-rich loop of the Abl kinase.

Qualitatively, the differential binding interactions between PPY-A and the other drugs included in this study indicate potential future direction for future improvement in designing drugs to target not just the hydrophobic binding pocket of the Abl kinase used by previous drugs, but also the other areas of the ATP-binding pocket, including the Gly-rich loop, to gain
selectivity and tight binding. To study the electrostatic interaction between the Gly-rich loop residue Gly316 and PPY-A, we set the charges of that residue to zero and recalculated the free energy of binding. We found that this interaction with a residue in the Gly-rich loop is very favorable to binding ($\Delta \Delta G_{316} > 2$ kcal/mol). The interacting partner on PPY-A (the pyyrolopyridine core, moiety 3) similarly contributes favorably to binding in both the WT and mutant complex ($\Delta \Delta G_{\text{moiety3}} > 1$ kcal/mol). These mutually favorable interactions can be capitalized on in future generation drugs. Indeed, Zhou et al.’s structural analysis of PPY-A showed potent activity and tight binding to the T315I Abl mutant kinase.\textsuperscript{1}

Interestingly, whether the drug is a DFG-out or DFG-in inhibitor, each drug in this comparative study forms key hydrogen bonds with Met318. Furthermore, the moieties of each drug that interact with the backbone of Met318 are already optimal to yield the most favorable electrostatic free energy of binding. To confirm our charge optimization findings, we also conducted component analysis on Met318, finding that in all drug complexes except nilotinib, Met318 contributes very favorably to binding ($\Delta \Delta G_{318}$ ranges from 1 - 2 kcal/mol; for nilotinib $\Delta \Delta G_{318} = -0.1225$ kcal/mol). Dasatinib in particular forms two hydrogen bonds with Met318; component analysis yielded $\Delta \Delta G_{316} > 2$ kcal/mol. Because of the nature of backbone interaction with Met318, it is unclear whether or not mutations at this position will confer any resistance. No current point mutations at Met318 have been documented. This may be because in the normal function of the Bcr-Abl kinase, the backbone of Met318 interacts with N1 of ATP.\textsuperscript{8} Thus, point mutations on residue 318 may completely eliminate the activity of the Abl kinase. However, this residue warrants further investigation.

In conducting this comparative study of five CML drugs, we also found some conserved moieties in the drugs, including a methylpiperazine group in ponatinib and imatinib. A moiety
often used for solubilization, this methylpiperazine group in ponatinib and imatinib contributes highly unfavorably to binding (figures 7 and 15). Furthermore, charge optimization shows that these groups are also highly suboptimal for binding (figures 8, 17, and 18). While this group helps make the drug soluble and thus feasible as a CML therapy, it is worth noting that nilotinib, which was rationally improved from computational and structural studies of imatinib15, no longer has this group. Instead, the methylpiperazine was replaced with a methyl-imidazole group. Thus, the methylpiperazine in imatinib and ponatinib can potentially be altered to significantly improve binding and retain solubility properties of the drug.

Interestingly, dasatinib also possesses a piperazine moiety in its structure. However, component analysis shows that piperazine does not contribute significantly favorably or unfavorably to the electrostatic binding interaction. The free energy of binding, however, is very sensitive to changes in charge on the piperazine’s atoms. Because of the high degree of sensitivity of the binding free energy, we can speculate that this moiety could potentially contribute favorably to binding if altered.

In addition, trifluoromethane is common in both ponatinib and nilotinib. Fluorinated compounds are often used in drug design, as it confers greater membrane permeability. They are also harder to oxidize, thus increasing the half life of the drug after consumption.58 To quantify the effects of the trifluoromethane on binding, we investigated the moiety with component analysis, revealing that in both ponatinib (WT and mutant) complexes and nilotinib complex, the contribution of trifluoromethane to the electrostatic portion of binding is negligible.

Finally, we would like to highlight the T315I mutation and the important role of electrostatics in the relatively hydrophobic binding pocket of Abl kinase. First and second generation drugs imatinib and nilotinib form key hydrogen bonds with the gatekeeper residue
Thr315. In addition to the steric clash introduced by the bulkier side chain of Ile, the mutation from Thr315 to Ile315 also eliminates this key hydrogen bond with the side chain OH of Thr. Previous studies and phase 2 trials have confirmed imatinib’s and nilotinib’s lack of activity against T315I Abl kinase. Like imatinib and nilotinib, dasatinib also forms a hydrogen bond with gatekeeper residue Thr315’s OH via the NH of the drug’s amide. However, upon mutation to Ile315, dasatinib loses efficacy against battling CML in patients. Charge optimization (figure 14b) shows that the amide is already fairly close to optimal to yield the most favorable electrostatic binding free energy, suggesting that any other charge distribution would be worse for binding. However, this amide is not optimal in either imatinib or nilotinib. Future work should focus on a more detailed analysis of this particular interaction in various drugs.

Both PPY-A and ponatinib show potent activity against mutant T315I Abl kinase. These third generation drugs bypass any interaction with the residue at position 315. Ponatinib uses its ethynyl group (moiety 6) to avoid contact with the bulky Ile in T315I mutant. Component analysis on this group reveals that it does not contribute significantly favorably or unfavorably to the electrostatic portion of binding. Similarly, component analyses on nearby residue T315 and I315 on the WT and mutant complex respectively show that neither residue contributes significantly to binding either. Taken together, these results suggest a lack of interaction between ponatinib and residue 315. Ponatinib may capitalize on interactions with other residues to inhibit the Abl kinase, rendering it effective against both the WT and T315I mutant. Analysis of PPY-A reveals a similar theme in binding to Abl kinase while bypassing interaction with residue 315.

Despite numerous other hydrogen bonding and hydrophobic interactions with the Abl kinase binding pocket, the point mutation T315I is enough to render the early generation drugs almost useless against CML. For example, though dasatinib forms two important hydrogen
bonds with Met318 and one hydrogen bond with Thr315, the mutation to Ile315 is enough to confer resistance of CML to dasatinib. Similarly, other early generation drugs imatinib and nilotinib also lose efficacy to the T315I mutant Abl kinase. While studies\textsuperscript{14,63} have proposed the loss of hydrogen bond in the T315I mutant as a cause of loss in treatment sensitivity to imatinib, no study has systematically analyzed residue 315 across multiple drugs. In contrast to this hypothesis, an MD simulation of T315I Abl kinase with imatinib indicates the major cause of resistance to be conformational changes resulting from the mutation rather than the loss of hydrogen bond to Thr 315.\textsuperscript{64} However, currently no crystallographic data on early generation drugs and T315I Abl kinase exists on the Protein Databank. Future steps of this project include conducting a docking study, in which we can build a T315I mutant model and dock imatinib, nilotinib, and dasatinib into the binding site to investigate the effects of the T315I mutation in the electrostatic portion of binding to these inhibitors.

Currently, the most common mutations found in clinical treatment of CML are at seven residue positions (250, 254, 255, 315, 351, 359, and 396), with T315I being the most resistant mutation.\textsuperscript{65} Given the differential binding interactions of different existing drugs, CML therapy could potentially include specific drug cocktails with previous generation drugs and future generation drugs to encompass as much interaction as possible with the Abl kinase. Already, second generation drugs such as dasatinib and nilotinib are used in conjunction with frontline therapy imatinib to battle resistance.\textsuperscript{14,66}

Upon unconstrained charge optimization, we found each drug’s optimal overall monopole to be quite high, ranging from 1.60e to as large as 6.60e. DFG-in inhibitors imatinib and nilotinib had an optimal overall monopole of 1.67e and 1.60e respectively when bound to the WT Abl kinase. Ponatinib had an optimal overall monopole of 1.60e to the WT and 2.02e to the
mutant Abl kinase. The original Abl kinase has an overall monopole of -7 to -10, depending on the missing residues in the crystal structure (described in methods section). The binding interface of Abl includes charged residues Glu286, Lys271, Asp381, and His361. One speculation may be that the long-range electrostatic effects of the overall negative Abl may affect the overall optimal monopole of the drug.

DFG-out inhibitors dasatinib and PPY-A are more difficult to rationalize. Dasatinib’s overall optimal monopole is extremely positive (6.6e), while PPY-A’s is 3.61e when optimized to the WT and 2.62e when optimized to the mutant Abl kinase. These values are substantially more positive than the optimal overall monopoles of the DFG-in inhibitors. This may be a result of the distinction between inactive and active conformations of the Abl kinase. In dasatinib and PPY-A’s case, only polar (non-charged) residues at the binding site interact with these drugs. In contrast, imatinib, nilotinib, and ponatinib interact with charged residues Glu286 and Asp381. Because dasatinib and PPY-A are less buried in the binding interface, these DFG-in inhibitors may be more influenced by the long-range electrostatic effect of the overall negative Abl kinase, making its optimal monopole more positive. It is also worth noting that the overall charge on the Abl kinase structures bound to dasatinib and PPY-A are more negative due to missing density of the crystal structure. Thus, the high magnitude of the optimal overall monopole of dasatinib and PPY-A may be an artifact of crystallographic missing density.

Interestingly, dasatinib’s overall optimal monopole is significantly higher at 6.6e. Upon closer examination, a few atoms dominate by adopting the highest possible charge given the constraints: 1.0e. These atoms are: C17, C31, H49, H52, H54, and H56. In figures 12 and 14, we see that the free energy of binding shows very low sensitivity to changes in these atoms. The carbons belong to the piperazine group, and the hydrogens belong to both the piperazine and
methanol group. We find that these atoms do not experience much desolvation upon binding; in both the bound and unbound states, they are exposed to solvent. Thus, an optimization may overcompensate with a high interaction term for the lack of desolvation. This could generate physically unrealistic high partial atomic charge magnitudes on those atoms. More investigation and follow-up studies are required to fully understand the implications of dasatinib’s high proclivity to a positive monopole.

Despite the unconstrained optimization results that indicate the optimality of positively charged drugs, small molecule inhibitors with an overall monopole is often physiologically undesirable. While this ultimately depends on the nature of the treatment, charged drugs have difficulty permeating the bilipid cell membrane. Furthermore, the overall charge on a drug affects solubility and kinetics as well.\textsuperscript{33}

We must note that this study only investigates the electrostatic portion of binding. As mentioned above, other components of binding also affect the binding of CML drugs to the Abl kinase, including aromatic and hydrophobic interactions. Experimental studies have shown a decrease in binding affinity between imatinib and T315I mutant relative to the WT Abl kinase.\textsuperscript{67} Furthermore, experimental data confirms our free energy of binding calculations that nilotinib binds more tightly to the WT Abl than imatinib.\textsuperscript{67} However, binding assays show a significant difference in binding of ponatinib to the WT and mutant Abl kinase.\textsuperscript{17} While our electrostatic-only calculation revealed a slight improvement in binding to the mutant Abl, ponatinib shows a five to seven fold reduction in inhibitory efficacy to the mutant relative to the WT. Thus, while electrostatics is often used to better understand the binding of drugs to their biological targets computationally, future studies that encompass other components of binding will be needed to quantitatively link computational analyses to experimental data.
Though it is beyond the scope of this work, our results could be extended with a molecular dynamic simulation. Our study assumes rigid binding. However, as mentioned previously, the Abl kinase can undergo conformational change upon binding to either DFG-in or DFG-out inhibitors. Future studies should apply charge optimization to molecular dynamic simulation structures, accounting for the conformational flexibility of the binding site when studying the electrostatics of binding within this system.

Additionally, as with previous studies that used charge optimization theory, it would be advantageous to calculated root mean square deviations between the actual and optimal charge distributions of each drug to provide a quantitative comparison to complement the visual inspections of the figures. Future work should quantify how optimal a charge is.

Finally, it is worth noting that we conducted our study in the absence of salt in the solvent. To address this potential discrepancy, we conducted a free energy binding calculation of imatinib with WT Abl in salt, yielding a difference between salt and non-salt calculations to be 0.06 kcal/mol. Follow up experiments should investigate the effect of salt on charge optimization and component analysis. Future studies should take into account salt in the solvent, especially because salt screens electrostatic interactions and therefore could potentially yield artificially high optimal monopoles when considering long-range interaction target monopoles.

In this study, we identified highly conserved moieties of drugs that contribute favorably to binding. In particular, while previous drugs failed to capitalize on the Gly-rich loop of the active conformation, this is a future direction that could potentially guide better development of novel CML therapeutics. Even after identifying moieties of a drug that could potentially be altered to improve binding, there are inherent challenges in proposing a modification that maintains the integrity of the drug’s structure while altering the partial atomic charges to reflect
the optimal distribution. Furthermore, changing a functional group may reduce solubility or introduce other pharmacologically unfavorable properties to the drug. Nevertheless, these potential alterations may provide a future direction for ration drug design.

Taken together, this study analyzes the differential binding interactions of five CML drugs. To the best of our knowledge, this is the first study that systematically investigates the interactions of DFG-out and DFG-in inhibitors with both WT and mutant Abl kinase through charge optimization, component analysis, and sensitivity analysis. We hope that this study will provide some insight that may help facilitate understanding of the principles of electrostatic binding and recognition in biological systems.
REFERENCES


5. Weisberg, E.; Griffin, J. D., Mechanism of resistance to the ABL tyrosine kinase inhibitor STI571 in BCR/ABL-transformed hematopoietic cell lines. Blood 2000, 95 (11), 3498-3505.


31. Minkara, M. S.; Davis, P. H.; Radhakrishnan, M. L., Multiple drugs and multiple targets: An analysis of the electrostatic determinants of binding between non-nucleoside HIV-1 reverse transcriptase inhibitors and variants of HIV-1 RT. *Proteins-Structure Function and Bioinformatics* 2012, 80 (2), 573-590.
52. Wang, J. M.; Morin, P.; Wang, W.; Kollman, P. A., Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. *Journal of the American Chemical Society* 2001, 123 (22), 5221-5230.
Blood Analysis of the structural basis of specificity of inhibition of the ABL kinase by STI571.


