Understanding Src Kinase Regulatory Mechanisms and Drug Resistance Using Deep Mutational Scanning and Chemical Biology

Sujata Chakraborty

A dissertation completed for the partial fulfilment

of the requirements of the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Dustin J. Maly, Chair

Jesse Zalatan

Champak Chatterjee

Program authorized to offer degree:

Chemistry

University of Washington

ABSTRACT

Understanding Src Kinase Regulatory Mechanisms and Drug Resistance Using Deep Mutational Scanning and Chemical Biology

Sujata Chakraborty

Chair of Supervisory Committee:

Professor Dustin J. Maly

Department of Chemistry

Protein phosphorylation controls a wide variety of cellular processes such as growth, differentiation, proliferation, and apoptosis in eukaryotes. Kinases are signaling enzymes that dictate cellular phosphorylation state by phosphorylating specific protein substrates. Over 530 protein kinases are encoded by the human genome and roughly half of these enzymes have at least one accessory domain in addition to the catalytic domain. The Src Family Kinases (SFKs) are a well-studied family of multi-domain, non-receptor tyrosine kinases. SFKs participate in numerous signal transduction pathways, and their misregulation is implicated in a variety of diseases, including cancer. Therefore, SFKs are of general interest as a model for understanding multi-domain kinase regulation and as potential drug targets.

In this thesis, I describe my efforts to use the SFKs as models for understanding inhibitor selectivity and drug resistance, which are major challenges in the field. I show that it is possible to develop ATP-competitive inhibitors that are highly selective for the SFK Lyn over other members of the SFKs. Obtaining such selectivity is significant because the ATP-binding sites of SFKs are almost identical and obtaining selectivity amongst such closely related kinases has proven to be particularly challenging. I show that it is possible to achieve selectivity for Lyn by developing inhibitors that target a region

called the helix αC. With a series of sequence swap experiments, I demonstrate that sensitivity to these Lyn-selective inhibitors is due to the identity of the linker residues that control the conformational flexibility of helix αC rather than any specific ATP-binding site interactions. Our strategy may hold promise for selectivity targeting other protein kinases. I also describe efforts to better understand how kinases develop resistance to ATP-competitive inhibitors. To do this, we used saturation mutagenesis and a yeast screening assay to comprehensively identify sites of ATP-competitive inhibitor resistance in Src. Using this methodology, resistance to the drug dasatinib and to a panel of conformation-selective, ATP-competitive inhibitors was profiled. Our efforts led to identification of mutations that provide general resistance to ATP-competitive inhibitors and mutations that uniquely affect specific modes of ATP-competitive inhibition. Interestingly, some of the strongest resistance mutations identified do not directly affect inhibitor binding but, instead, appear to confer resistance through other mechanisms. I describe comprehensive biochemical and biophysical analyses to better understand how a region on the N-terminal lobe of Src's catalytic domain, which was previously identified as participating in the regulation of Src through an unknown mechanism, confers broad resistance. As a major challenge in the field of kinase drug discovery is the emergence of drug resistance, our efforts may help better inform the development of chemotherapeutic treatments that are less prone to resistance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES

ACKNOWLEDGEMENTS

This dissertation would not have come to fruition without the incredible support, guidance, and encouragement from my mentors, family, friends, and colleagues. First, I would like to thank my Ph.D. advisor, Dustin Maly, for being an outstanding mentor. From Dusty, I learned how to be a thoughtful, rigorous, and productive scientist. I will forever be grateful towards him for the faith he showed in me, even when my own faltered. He ensured that my capabilities as a scientist went beyond hands-on experimental expertise. He persistently encouraged me to not only develop a comprehensive understanding of my projects but also be proficient in presenting my work to my peers. Next, I would like to express my sincerest gratitude towards my doctoral committee members, Dr. Jesse Zalatan, Dr. Champak Chatterjee, Dr. Forrest Michael, and my GSR Dr. Dana Miller for their constructive feedback, word of wisdom and helpful guidance.

I have had the opportunity to collaborate with various members of the UW scientific community, all of whom have had a massive impact on this work. First, I am incredibly grateful to Dr. Douglas Fowler and his extraordinarily talented student Dr. Ethan Ahler. Doug is always so enthusiastic. His depth of knowledge in every subject awed me every time we met to discuss the progress of our project. He is genuinely inspirational, who also taught me how to be a better presenter. I thank Ethan who single-handedly performed all the DMS work that preceded the work discussed in Chapter 2 of this dissertation. I thank Dr. Miklos Guttman, who taught me everything about Hydrogen-Deuterium exchange, which resulted in some crucial findings.

I have been extremely fortunate to have worked with some of the most talented and intelligent co-workers. I want to thank Dr. Takayuki Inukai, an extraordinarily gifted and hard-working visiting scientist from Japan. He and I collaborated on a project that resulted in a fruitful peer-reviewed article discussed in Chapter 1 of this dissertation. I have collaborated with Linglan, Emily, and Zack, who I

would like to dearly thank for continually motivating me towards excelling by providing helpful, constructive feedback. I would like to thank my seniors, especially Dr. Dan Cunningham-Bryant and Dr. Jack Rose, for teaching me techniques and better scientific practices. I thank Dr. Ames Register for setting the platform on which my projects stand. I want to thank Dr. Hannah Feldman, from whom I have learned how to be a better presenter and inspire me to be a high performing scientist. I would like to thank Chloe for her support and being my first friend in grad school. Grad school is a challenging endeavor. I want to thank all past and present Malylab members for making the lab a fun, positive, and friendly environment to work in.

I want to thank all my grad school friends. Julia, Andrea, and Wei Pin, you have made grad school enjoyable. I appreciate you guys dearly and want to thank you for supporting and encouraging me through tough times. Jefferey, thank you for taking care of Millie and for being such a kind-hearted person. I thank Suchetana for being such a dear and caring friend. Despite being miles apart, you could always sense my anxiety. Thank you for making me smile and forget my stress.

Last but not least, I would like to thank my family. I honestly will not have come this far if it was not for Ma's endless encouragement. She is my forever cheerleader to whom I owe all my present and future successes. Thank you, Ma, for making small and big sacrifices silently so I can chase my dreams, I know how hard it must have been for you, even though I don't always express it. I thank Baba, who is never tired of telling me how proud he is of me. He has always encouraged me to be passionate about my dreams and never shown doubt in me. I want to thank my brother and sister-in-law for always cheering for me, and for all the advice. I especially want to thank them for bringing my nephew into my life. Watching him grow up always fills me up with so much joy and positivity.

DEDICATION

To my loving parents and anyone who reads this

INTRODUCTION

I.1 Opening comments

In 1909 Peyton Rous, a pathologist at the Rockefeller Institute in New York City with a keen interest in cancer was handed a Plymouth Rock hen to investigate a large tumor developing in its breast. He determined that the tumor was a sarcoma.[1, 2] Later, in 1911, Rous did a straightforward experiment that changed the future of how viral infections and cancer is perceived. He crushed the tumor and passed in through a fine-mesh filter through which even bacteria cannot pass.[3] To his surprise, this 'cell-free' filtrate resulted in sarcoma when injected into other healthy chickens of the same species and consequently in birds of other species, upon later passages.[4] This discovery was so ahead of its time that his contemporaries did not pay much attention to it. Years later, upon the subsequent discovery of viral genetic material and the process of reverse transcription in retroviruses, the Nobel Prize committee recognized the significance of Peyton Rous' groundbreaking discovery. He was finally honored in the year 1966. The virus came to be known as the Rous Sarcoma Virus (RSV) which paved the way for endless breakthroughs in cancer research.[2, 5]

RSV, like other simple retroviruses, contains four genes, but only one gene known as v-Src (short for sarcoma) is oncogenic, or in other words, the v-Src gene results in cancer. Decades later, Harold Varmus and Michael Bishop, scientists at the University of California, San Francisco, made another stunning discovery.[6-8] They found that the v-Src gene is not a foreign entity belonging to the virus. In fact, the v-Src gene is almost identical to a gene found in birds and virtually in all animals, including human beings. They called this gene c-Src (cellular). c-Src became the first proto-oncogene ever discovered. In healthy cells, the c-Src gene is vital for proper cellular functions. It is posited that RSV was a 'genetic hitchhiker' that picked up the normal gene, during an invasion and attached it to its own genetic material, subsequently also deleting a critical regulatory component that converted c-Src into

an oncogene. This paradigm altering discovery won Bishop and Varmus, the Nobel Prize in 1989 in the field of Physiology or Medicine, a second in the field of Src biology.

Src gene encodes a protein that belongs to a class of signaling enzyme known as kinases.[6, 9] This dissertation documents my efforts in advancing our understanding of Src kinase and members of the Src Family of Kinases. A century and over 30,000 publications later, Src continues to be a gift that keeps on giving. My work aims at understanding the fine-scale regulation of Src kinase for better development of kinase inhibitors.

I.2 Src Family Kinases

A cell encounters a variety of extracellular stimuli that prompts it to respond in a specific manner. In order to accurately respond to the continually changing landscape of signals, cells enlist a diverse collection of signal transduction proteins. Protein kinases are one such class of signaling proteins. There are as many as 538 different protein kinases encoded by the human genome. Protein kinases all contain a catalytic domain that folds into a bilobal structure that interacts with the cofactor ATP. Kinases can be envisioned as highly dynamic molecular switches which, upon receiving a signal, are turned 'on' to transfer the γ-phosphate from ATP to specific protein substrates. These phosphorylation events mark protein substrates as part of a cascade of cellular responses. In the absence of a cellular signal, most kinases must remain in an 'off' state with low activity. In order to only function when specific signals are received, kinases are tightly regulated in cells by a variety of mechanisms. About 50% of kinases in humans are multi-domain, implying that in addition to the phosphotransferase catalytic domain, they have accessory domains that serve as regulatory apparatuses that control kinase 'on'/ 'off' states. Mutations or deletions of one or more accessory domain have been demonstrated to result in unregulated kinase activity; in many cases leading to aberrant cellular behavior.

A well-studied family of multi-domain protein kinases is the Src Family of Kinases (SFKs), which are a non-receptor tyrosine kinase family. There are 8 SFKs categorized into two subfamilies (**Figure I-1**). The Src-A family consists of Src, Yes, Fyn and Fgr, and the Src-B family consists of Lck, Blk, Hck and Lyn. SFKs are pleiotropic, participating in a plethora of cellular responses that include cell differentiation, adhesion, migration, invasion, proliferation, morphology, and cell survival.[10, 11] SFKs are indispensable to proper cell functioning, and their misregulation has been implicated in various diseases, including cancer and diabetes.

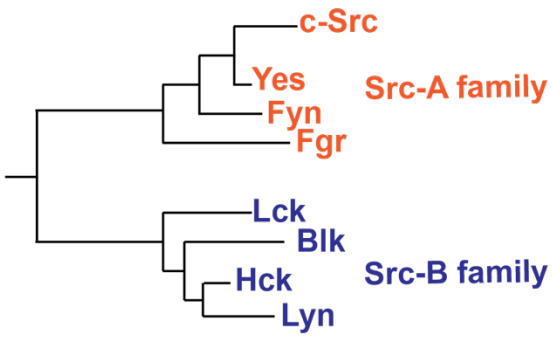


Figure I-1: Phylogenetic tree of Src Family Kinases (SFKs). Src-A family SFKs are shown in orange. Src-B family SFKs are shown in purple. [12]

I.3 Regulation of Src kinase

SFKs are ~50-60 kDa proteins that share high sequence conservation, domain architecture, and have a similar mechanism of action. Here, we will more closely focus on Src to understand these structural features and their functions (**Figure I-2 A**). At the N-terminus, there is a disordered 20 amino acid long membrane-associating sequence, often referred to as the Src Homology 4 (SH4) domain. This domain is modified by fatty acids co-translationally. In Src, the SH4 domain is co-translationally myristoylated (14-carbon long fatty acid chain). For all other members, one or two additional palmitoyl groups (16-carbon long fatty acid chain) get added to the SH4 domain. The fatty acid and SH4 domain form a bipartite membrane-associating module that localizes the kinase to various compartments in the cell.

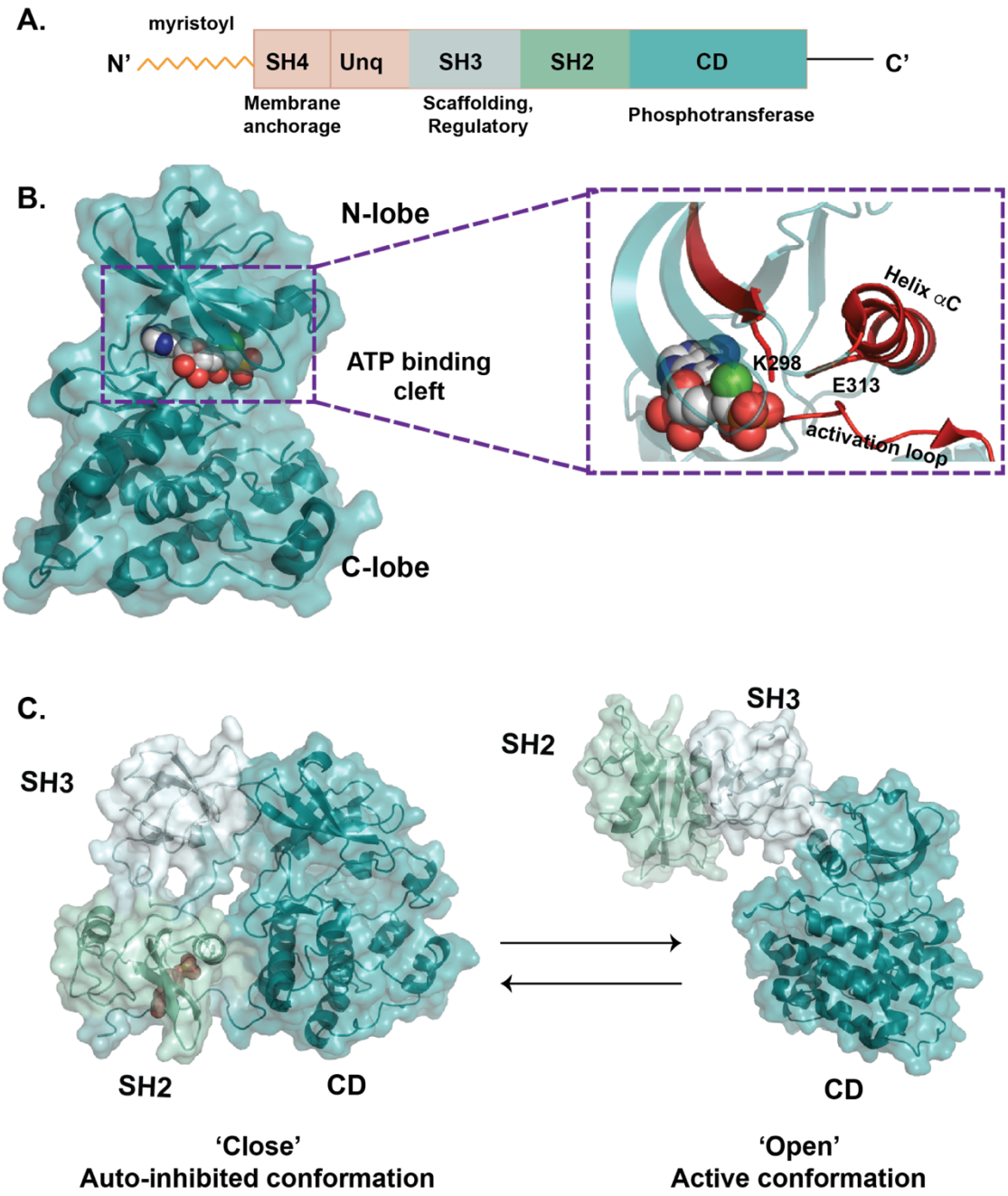


Figure I-2: Domain architecture and regulation of Src kinase.

A. Line structure of Src kinase, starting from the N-terminal myristoylated SH4 domain, Unique (Unq) domain, SH3, SH2 and C-terminal catalytic domain (CD). **B.** (Left) Structure of the Src catalytic domain (PDB: 3DQW), showing the N-terminal and C-terminal lobes of the kinase along with the ATP-binding cleft. (Right) Zoom-in view of the Src active site showing catalytically important residues K298 on the β3 strand and E313 on the helix αC. The activation loop is shown in red. **C.** The structure of Src kinase in the 'close' auto-inhibited (left, PDB id: 2SRC) and the 'open' active (right, PDB id: 1Y57) global conformations.

Recently, we have uncovered a regulatory feature of the SH4 domain.[13] Next, lies the Unique (Unq) domain, where the kinase sequence varies greatly between family members, and this domain has member-specific roles. Followed by this extended disordered region, there is the 60 amino acid long folded SH3 domain and 100 amino acid long folded SH2 domains. The functions of these domains are well documented in decades of SFK structural and biochemical studies. These domains are non-catalytic and act as scaffolds for various SFK protein partners, mediating significant protein-protein interactions with other effector proteins. The SH3 domains of SFKs interact with a consensus P-(XX)-P polyproline motif present in various effector proteins, such as the HIV-1 Nef protein.[14]

The SH2 domain interacts with phosphorylated tyrosine side chains present on a protein partner, such as the focal adhesion proteins.[15] C-terminal to the non-catalytic SH3/SH2 domains is the catalytic domain (CD), which is bilobal (**Figure I-2 B**) in structure. Between the N- and C-lobes lies the catalytic cleft where ATP binds. Upon activation, the kinase active site undergoes a dynamic conformational change to allow substrate binding and phosphorylation. In this active state, two key residues, a conserved lysine (K298) on the β3 strand and a conserved glutamic acid (E313) on helix αC form a salt bridge, necessary for substrate phosphorylation. Additionally, the activation loop is oriented such that substrate binding can be facilitated. Finally, C-terminal to the CD is a short-disordered tail. The tyrosine kinases Csk (C-terminal Src Kinase) and CHK (Csk Homologous Kinase) are known to phosphorylate a tyrosine (Y530) on this tail. This phosphorylated tyrosine (pY530) can then dock onto Src's SH2 domain to form an autoinhibitory intramolecular interaction.[16, 17]

The SH3 domain can also intramolecularly interact with a proline-containing linker present between the SH2 and the CD. Together, these two crucial interactions lock the kinase in an auto-inhibited 'closed' or 'off' state (**Figure I-2 C**). In this 'closed' state, the kinase active site is allosterically shifted to a conformation where the critical salt bridge between K298 and E313 is disrupted, and the kinase is no longer capable of phosphorylating substrate. Therefore, the SH3/SH2 regulatory apparatus is

responsible for Src's canonical mechanism of regulation. As shown in **Figure I-2** C, the SH3 and SH2 domains maintain Src's equilibrium between the 'closed' and 'open' states. In the presence of higher affinity ligands or activating mutations these auto-inhibitory interactions are broken, and Src is activated ('open'). This allosteric mechanism of Src regulation is well characterized, thus, making this kinase a prototype for understanding the regulation of other multi-domain protein kinases.

I.4 SFK in diseases and inhibitor development

SFKs interact with and often phosphorylate a diverse repertoire of cellular proteins that include other kinases, scaffolding-proteins, integrins, and actin. SFKs are, therefore, involved in a multitude of signal transduction pathways that control cell adhesion, growth, cell migration, and proliferation. Fibroblast cells lacking Src show slowed motility and Src/Yes/Fyn knock-out mouse fibroblast cells show reduced adhesion on fibronectin as well as diminished tyrosine phosphorylation. A common SFK effector protein is the FAK (Focal Adhesion Kinase), which upon integrin stimulation recruits Src by interacting with its SH2 domain. This interaction promotes cell migration through a variety of signaling proteins. In addition, SFKs sit downstream of many Receptor Tyrosine Kinases (RTKs) and T-cell and B-cell receptors. Therefore, SFKs are significant players in several canonical signaling pathways. It is not surprising that the misregulation of SFKs leads to aberrant cell signaling, contributing to diseases including hematopoietic disorders, tumor progression, and several forms of cancers (colon, breast, pancreas, liver, gastric, ovarian, stomach, epithelial, and lymphoid). [18]

For this reason, significant attention is devoted to the development of SFK inhibitors, and several kinase inhibitors are already proving to be highly efficient towards cancer therapeutics.[19, 20] However, a challenge in the field is the development of member-specific inhibitors. Despite structural homology and high sequence identity, each SFK member has non-redundant roles in a cell. For example, in MES (Murine Embryonic Stem) cells, it has been demonstrated that while Src promotes cell differentiation, Yes promotes cell renewal and diminishes differentiation.[21-23]

It is of utmost importance to design highly selective inhibitors that target a single SFK member to avoid the undesired off-target effects of inhibiting other members, given the non-redundant roles of SFKs. In **Chapter 1** of this dissertation, I describe my efforts in designing member-specific inhibitors of SFKs. Using conformation-selective inhibitors, we were able to obtain Lyn-selective inhibitors that showed over 200-fold selectivity for Lyn over Src. A second challenge in the field of kinase drug discovery is the emergence of inhibitor resistance mutations. In **Chapter 2**, I discuss how we used Deep Mutational Scanning (DMS) on Src to understand the basis of drug resistance and to guide more rational kinase drug discovery.

I.5 References

[1] (2005) Peyton Rous: father of the tumor virus, *J Exp Med 201*, 320.

[2] Weiss, R. A., and Vogt, P. K. (2011) 100 years of Rous sarcoma virus, *J Exp Med 208*, 2351-2355.

[3] Rous, P. (1911) A Sarcoma of the Fowl Transmissible by an Agent Separable from the Tumor Cells, *J Exp Med 13*, 397-411.

[4] Brugge, J. S., Jarosik, G., Andersen, J., Queral-Lustig, A., Fedor-Chaiken, M., and Broach, J. R. (1987) Expression of Rous sarcoma virus transforming protein pp60v-src in Saccharomyces cerevisiae cells, *Mol Cell Biol 7*, 2180-2187.

[5] Martin, G. S. (2004) The road to Src, *Oncogene 23*, 7910-7917.

[6] Hunter, T. (2015) Discovering the first tyrosine kinase, *Proc Natl Acad Sci U S A 112*, 7877-7882.

[7] Stehelin, D., Varmus, H. E., Bishop, J. M., and Vogt, P. K. (1976) DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA, *Nature 260*, 170-173.

[8] Varmus, H. (2019) Of oncogenes and open science: an interview with Harold Varmus, *Dis Model Mech 12*.

[9] Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002) The protein kinase complement of the human genome, *Science 298*, 1912-1934.

[10] Thomas, S. M., and Brugge, J. S. (1997) Cellular functions regulated by Src family kinases, *Annu Rev Cell Dev Biol 13*, 513-609.

[11] Aleshin, A., and Finn, R. S. (2010) SRC: a century of science brought to the clinic, *Neoplasia 12*, 599-607.

[12] Shah, N. H., Lobel, M., Weiss, A., and Kuriyan, J. (2018) Fine-tuning of substrate preferences of the Src-family kinase Lck revealed through a high-throughput specificity screen, *Elife 7*.

[13] Ahler, E., Register, A. C., Chakraborty, S., Fang, L., Dieter, E. M., Sitko, K. A., Vidadala, R. S. R., Trevillian, B. M., Golkowski, M., Gelman, H., Stephany, J. J., Rubin, A. F., Merritt, E. A., Fowler, D. M., and Maly, D. J. (2019) A Combined Approach Reveals a Regulatory Mechanism Coupling Src's Kinase Activity, Localization, and Phosphotransferase-Independent Functions, *Mol Cell*.

[14] Pene-Dumitrescu, T., Shu, S. T., Wales, T. E., Alvarado, J. J., Shi, H., Narute, P., Moroco, J. A., Yeh, J. I., Engen, J. R., and Smithgall, T. E. (2012) HIV-1 Nef interaction influences the ATP-binding site of the Src-family kinase, Hck, *BMC Chem Biol 12*, 1.

[15] Mitra, S. K., and Schlaepfer, D. D. (2006) Integrin-regulated FAK-Src signaling in normal and cancer cells, *Curr Opin Cell Biol 18*, 516-523.

[16] Okada, M. (2012) Regulation of the SRC family kinases by Csk, *Int J Biol Sci 8*, 1385-1397.

[17] Okada, M., Nada, S., Yamanashi, Y., Yamamoto, T., and Nakagawa, H. (1991) CSK: a protein-tyrosine kinase involved in regulation of src family kinases, *J Biol Chem 266*, 24249-24252.

[18] Kumar, A., Jaggi, A. S., and Singh, N. (2015) Pharmacology of Src family kinases and therapeutic implications of their modulators, *Fundam Clin Pharmacol 29*, 115-130.

[19] Kim, L. C., Song, L., and Haura, E. B. (2009) Src kinases as therapeutic targets for cancer, *Nat Rev Clin Oncol 6*, 587-595.

[20] Roskoski, R., Jr. (2015) Src protein-tyrosine kinase structure, mechanism, and small molecule inhibitors, *Pharmacol Res 94*, 9-25.

[21] Zhang, X., Meyn, M. A., 3rd, and Smithgall, T. E. (2014) c-Yes tyrosine kinase is a potent suppressor of ES cell differentiation and antagonizes the actions of its closest phylogenetic relative, c-Src, *ACS Chem Biol 9*, 139-146.

[22] Meyn, M. A., 3rd, and Smithgall, T. E. (2009) Chemical genetics identifies c-Src as an activator of primitive ectoderm formation in murine embryonic stem cells, *Sci Signal 2*, ra64.

[23] Anneren, C., Cowan, C. A., and Melton, D. A. (2004) The Src family of tyrosine kinases is important for embryonic stem cell self-renewal, *J Biol Chem 279*, 31590-31598.

CHAPTER 1[1]: Targeting dynamic ATP-binding site features allows discrimination between highly homologous protein kinases

1.1 Abstract

ATP-competitive inhibitors that demonstrate exquisite selectivity for specific members of the human kinome have been developed. Despite this success, the identification of highly selective inhibitors is still very challenging, and it is often not possible to rationally engineer selectivity between the ATP-binding sites of kinases, especially amongst closely related family members. Src-family kinases (SFKs) are a highly homologous family of eight multi-domain, non-receptor tyrosine kinases that play general and specialized roles in numerous cellular processes. The high sequence and functional similarities between SFK members make it hard to rationalize how selectivity can be gained with inhibitors that target the ATP-binding site. Here, we describe the development of a series of inhibitors that are highly selective for the ATP-binding sites of the SFKs Lyn and Hck over other SFKs. By biochemically characterizing how these selective ATP-competitive inhibitors allosterically influence the global conformation of SFKs, we demonstrate that they most likely interact with a binding pocket created by the movement of the conformationally-flexible helix αC in the ATP-binding site. With a series of sequence swap experiments, we show that sensitivity to this class of selective inhibitors is due to the identity of residues that control the conformational flexibility of helix αC rather than any specific ATP-binding site interactions. Thus, the ATP-binding sites of highly homologous kinases can be discriminated by targeting heterogeneity within conformationally flexible regions.

[1] This chapter has been adapted with only minor revisions from: **Sujata Chakraborty***, Takayuki Inukai et.al. "Targeting dynamic ATP-binding site features allows discrimination between highly homologous protein kinases." Copyright © ACS chemical biology, 2019.

1.2 Introduction

Protein kinases are key regulators of signal transduction pathways and represent a significant portion of the human proteome.[1] All kinases possess a catalytic domain (CD) composed of an N-lobe and a C-lobe with a catalytic cleft that interacts with ATP located between them. Dysregulated kinase activity is associated with a number of human diseases and, for this reason, protein kinase inhibition is an active area of drug discovery.[2, 3] There have been significant advancements in our ability to target specific kinases with small molecule inhibitors over the last decade. Despite the fact that a majority of potent and selective small molecule inhibitors target the highly conserved ATP-binding site, it is possible to identify inhibitors that demonstrate exquisite selectivity for a specific kinase target.[4-6] However, the development of highly selective kinase inhibitors is still very challenging and it is often not possible to rationally engineer discrimination between similar targets because the molecular determinants of specificity are often not well understood. This selectivity problem is particularly acute for closely related subgroups of kinases that possess ATP-binding sites with very few sequence differences.

Most ATP-competitive kinase inhibitors can be classified into three broad categories based on the active site conformation that they stabilize (**Figure 1-1**).[7] Many inhibitors–often called *type 1*–stabilize or are compatible with an "active" ATP-binding site conformation, where all catalytic residues are aligned in an orientation compatible with phosphate transfer. ATP-competitive inhibitors can also stabilize two distinct "inactive" kinase conformations, in which catalytically important residues are displaced from a phosphotransferase-competent orientation. The most common inactive conformation stabilized by ATP-competitive inhibitors is the DFG-out conformation, which is defined by displacement of the Asp-Phe-Gly (DFG)-motif at the base of the activation loop from an active conformation.[8, 9] Inhibitors that stabilize the DFG-out conformation–often called *type 2*–usually possess moieties that make hydrogen bonds with a conserved Glu residue in helix αC and the

backbone of the DFG-motif and a large hydrophobic substituent that occupies a pocket created by movement of the Phe residue in the DFG-motif. Potent inhibitors that stabilize an inactive conformation characterized by displacement of helix αC from an active conformation have also been developed.[10-12] Inhibitors that stabilize this inactive form, which is often referred to as the helix αC -out conformation, usually possess substituents that occupy a pocket created by movement of helix αC. It has been widely speculated that inhibitors that stabilize the DFG-out inactive conformation are inherently more selective than type 1 inhibitors but recent studies call into question the magnitude of this effect.[8, 11] Fewer inhibitors that stabilize the helix αC -out conformation have been characterized and the inherent selectivity of this ligand class is less understood.

Human Src-family kinases (SFKs) are a family of eight multidomain, nonreceptor tyrosine kinases that play general and specialized roles in numerous cellular processes.[7, 13] SFKs all share regulatory SH2 and SH3 domains that are N-terminal to a highly conserved kinase CD. SFKs are closely related in sequence–64-90% CD sequence identity–and only significantly diverge at their membrane-interacting N-termini and the linker (SH2-kinase domain linker) that connects the SH2 domain to the kinase domain. The SFKs are grouped into two subfamilies–Src-A and Src-B subfamilies (**Figure 1-2**)–based on their SH2-kinase domain linker motifs.[14] Despite the importance of SFKs in numerous fundamental processes, there has been limited success in identifying cell-permeable inhibitors that discriminate between family members.[4, 15-17] Here, we describe a series of ATP-competitive inhibitors that are highly selective for the Src-B subfamily over the Src-A subfamily SFKs. Using a panel of biochemical assays, we show that these inhibitors promote the autoinhibited form of SFKs, most likely by stabilizing the helix αC -out conformation of the ATP-binding site. Finally, by performing a series of sequence swap experiments, we demonstrate that flexible linker residues that connect helix αC to the N-terminal lobe of the ATP-binding site are responsible for inhibitor sensitivity. Together,

these experiments show that it is possible to discriminate between highly similar ATP-binding sites by targeting differences in the conformational flexibility of helix αC.

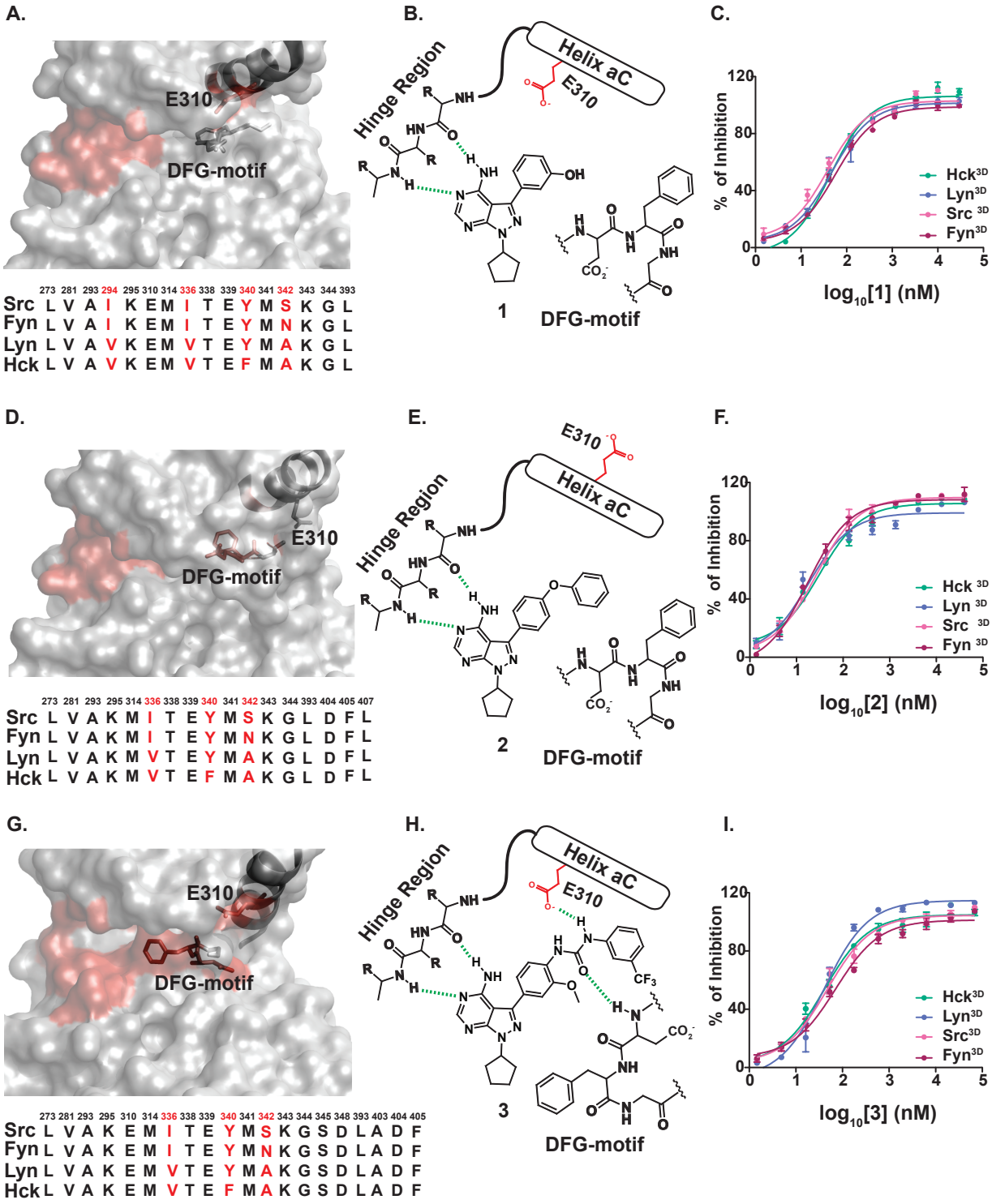


Figure 1-1: The ATP-binding sites of SFKs are structurally and functionally similar.

A. Co-crystal structure of Src bound to dasatinib (PDB ID, 3G5D; top). Residues that form hydrophobic or polar contacts–within 4 Å–with dasatinib (identified using Ligplot+)[18] are shown in dark red. Sequence alignment of the 16 hydrophobic or polar dasatinib contact residues for Src, Fyn, Hck, and Lyn (*bottom*). Residues shown in black are conserved, while residues shown in red vary between these four SFKs. **B**. Binding pose of inhibitor **1**, which stabilizes an active ATP-binding site conformation. **C**. Inhibitory curves for inhibitor **1** against Src^{3D}, Fyn^{3D}, Hck^{3D}, and Lyn^{3D} activity in the presence of 1 mM ATP (n=3). **D**. Crystal structure of Src bound to a helix αC -out-stabilizing analog of dasatinib (PDB ID, 4YBK; *top*). Residues that form hydrophobic or polar contacts–within 4 Å–with helix αC -out-stabilizing dasatinib are shown in dark red. Sequence alignment of the 17 hydrophobic or polar residues that form contacts with helix αC -out-stabilizing dasatinib for Src, Fyn, Hck, and Lyn (*bottom*). **E**. Binding pose of inhibitor **2**, which stabilizes the helix αC -out conformation of the ATP-binding site. **F**. Inhibitory curves for inhibitor **2** against Src^{3D}, Fyn^{3D}, Hck^{3D}, and Lyn^{3D} activity in the presence of 1 mM ATP (n=3). **G**. Crystal structure of Src bound to a DFG-out-stabilizing analog of dasatinib (PDB ID, 4YBJ; *top*). Residues that form hydrophobic or polar contacts–within 4 Å–with DFG-out-stabilizing dasatinib are shown in dark red. Sequence alignment of the 20 hydrophobic or polar DFG-out-stabilizing dasatinib contact residues for Src, Fyn, Hck, and Lyn (*bottom*). **H**. Binding pose of inhibitor **3**, which stabilizes the DFG-out conformation of the ATP-binding site. **I**. Inhibitory curves for inhibitor **3** against Src^{3D}, Fyn^{3D}, Hck^{3D}, and Lyn^{3D} activity in the presence of 1 mM ATP (n=3).

1.3 Results and Discussion

The kinase domains of SFKs are highly conserved, with very few sequence differences in the residues comprising their ATP-binding sites. To structurally define the similarity of SFK ATP-binding sites, we analyzed co-crystal structures of the CD of Src bound to the drug dasatinib and two conformation-selective analogs of dasatinib (**Figure 1-1 A-C**).[11] Using these three co-crystal structures, we identified residues that make hydrophobic or polar contacts with each inhibitor (within 4 Å) and performed sequence alignments of inhibitor contact residues for the Group A SFKs Src and Fyn and the Group B SFKs Hck and Lyn (**Figure 1-2 B-D**). As expected, there are very few sequence differences between Src, Fyn, Hck, and Lyn in the regions that make contact with all three inhibitors (**Figure 1-2 B-D**). There are sixteen residues in Src's ATP-binding site, which is in the active conformation, that make contact with dasatinib. For Src, Fyn, Hck, and Lyn, twelve of these residues are identical, three are similar, and only one–residue 342 (chicken Src numbering)–varies substantially (**Figures 1-1 A, 1-2 B**). Seventeen ATP-binding site residues in Src interact with the helix αC-out version of dasatinib, with three residues making unique interactions relative to dasatinib (**Figures 1-1 D, 1-2 C**). For Src, Fyn, Hck, and Lyn, fourteen of these residues are identical, two are similar, and, again, only residue 342 varies substantially. Src forms the largest interaction surface with the DFG-out version of

dasatinib, with twenty ATP-binding site residues making inhibitor contacts (**Figure 1-1 G, 1-2 D**). However, seventeen of these residues are identical and two are similar in Src, Fyn, Hck, and Lyn. Like the other two inhibitors, only contact residue 342 greatly varies between these four kinases. (**Figure 1-1 D**). Residue 342 is located on the hinge region and its side chain is perpendicular to the ATP-binding site, making it an unlikely source of inhibitor selectivity.

To functionally define the similarities of SFK ATP-binding sites, we tested the sensitivities of Src, Fyn, Hck, and Lyn constructs that contain the SH3, SH2, and CDs (SFK^{3D}) to pyrazolopyrimidine-based inhibitors that stabilize either the active (**1**), helix αC -out (**2**), or DFG-out (**3**) conformation of the ATP-binding site (**Figures 1-1 B**, **E**, **and H**). Consistent with the high sequence similarity of their ATP-binding sites, we found that Src^{3D}, Fyn^{3D}, Hck^{3D}, and Lyn^{3D} are equally sensitive to all three inhibitors (**Figures 1-1 C**, **F**, **I, 1.3)**. Notably, conformation-selective inhibitors **2** and **3** were as non-selective amongst the SFKs as type 1 inhibitor **1**.

Lyn/Hck-selective inhibitors. The sequence and functional similarities between Src, Fyn, Hck, and Lyn makes it hard to rationalize how selectivity can be gained by targeting specific interactions within their ATP-binding sites. However, we have observed that differences in the linker connecting the SH2 domain to the CD (SH2-CD linker) of SFKs lead to heterogeneous allosteric coupling between their regulatory and CDs.[19, 20] When SFKs are in a closed autoinhibited state, their SH3 domains and SH2-CD linkers form extensive contacts with helix αC in the CD and these regions are highly coupled through allostery. Based on these observations, we predicted that it may be possible to distinguish between the ATP-binding sites of SFKs with inhibitors that make extensive contacts with the region adjacent to helix αC. Therefore, we generated a panel of inhibitors based on a pyrrolo[2,3-d]pyrimidine scaffold that should project substituents from its *C*-5 position

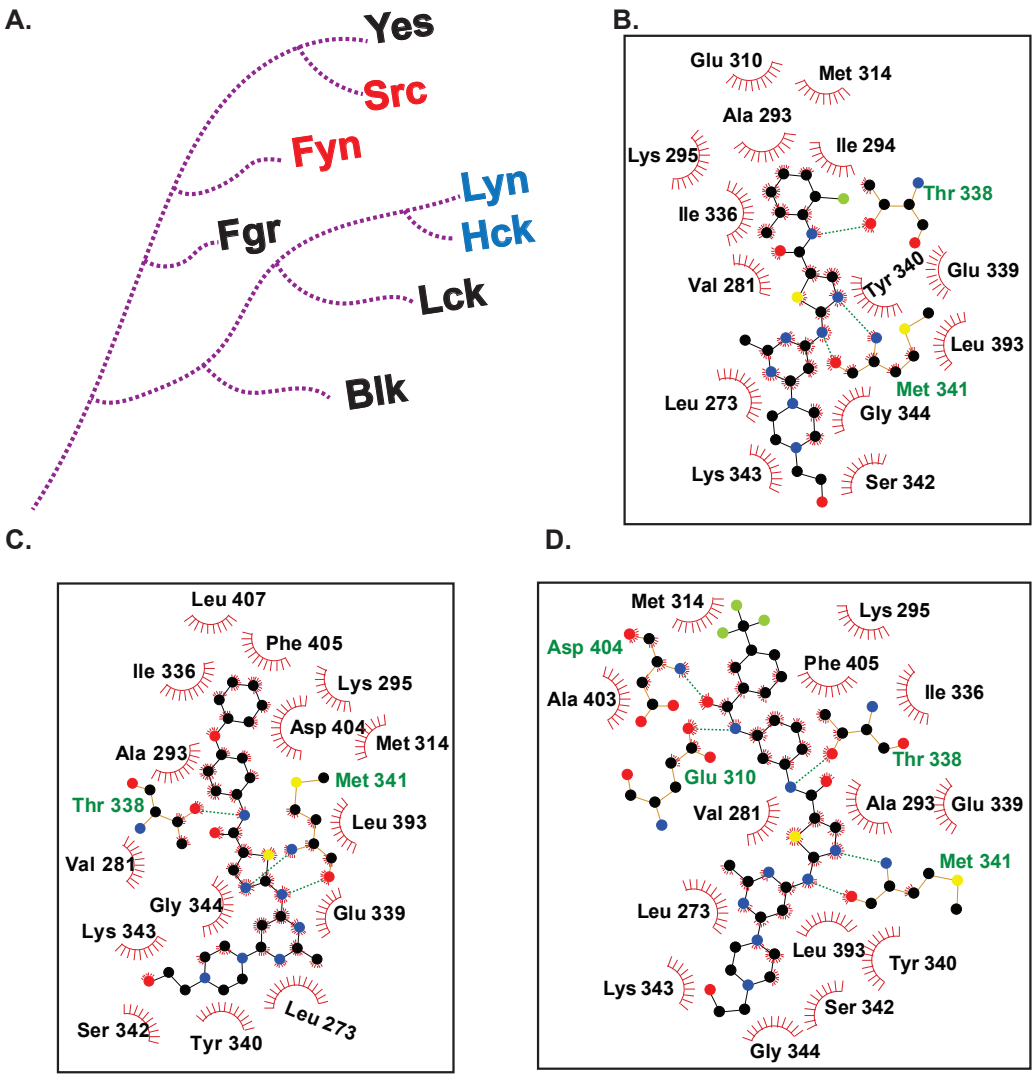


Figure 1-2: Ligplot+ analysis of Src bound to dasatinib or dasatinib analogs.

A. Dendogram of SFK members. Colored family members (red-Group A and blue-Group B) were characterized in this study. **B**-**D**. Schematic diagram of the binding interactions of dasatinib (**B**), a helix αC-out-stabilizing analog of dasatinib (**C**), and a DFG-out-stabilizing analog of dasatinib (**D**). Red semicircles represent a hydrophobic interaction. Green dotted lines indicate a hydrogen-bond between the inhibitor and Src. Schematics were generated using Ligplot+.[18]

towards helix αC when forming hydrogen-bonding interactions with the hinge region of SFKs.[21] To provide selectivity for kinases that contain threonine gatekeeper residues, like the SFKs, a 2-methoxyaniline was introduced at the *C*-5 position of the pyrrolo[2,3]pyrimidine scaffold. Acylation

of the 2-methoxyaniline allowed us to introduce diverse substituents that we predicted would project towards the helix αCs of SFKs (**Figure 1-4 A**). Various benzofurans, benzothiophenes, indoles, pyrroles, and pyrazoles were introduced at this position.

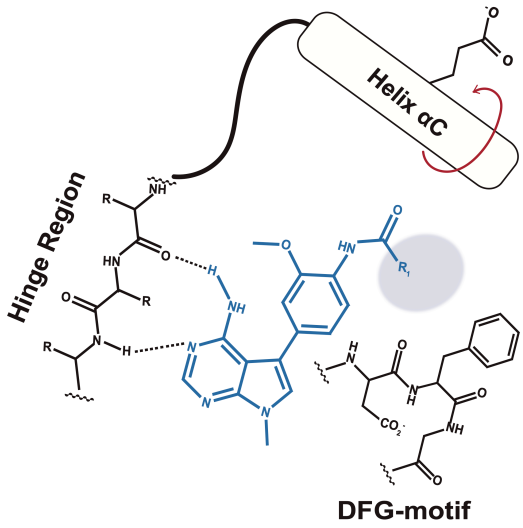
	K_I (nM)		
SFK	**1**	**2**	**3**
Hck ^{3D}	**3.2 ± 0.4**	**2.2 ± 0.2**	**2.3 ± 0.1**
Lyn ^{3D}	**2.9 ± 0.1**	**1.9 ± 0.1**	**2.1 ± 0.3**
Src ^{3D}	**2.8 ± 0.2**	**1.1 ± 0.1**	**2.8 ± 0.4**
Fyn ^{3D}	**3.5 ± 0.4**	**1.2 ± 0.3**	**3.8 ± 0.7**

Figure 1-3: K_I values of inhibitors 1-3, calculated using the Cheng-Prusoff Equation against Hck^{3D}, Lyn^{3D}, Src^{3D}, and Fyn^{3D} (mean ± sem, n=3).

We first obtained IC_{50}s and K_I values for each inhibitor in our panel against Src^{3D}, Fyn^{3D}, Hck^{3D} and Lyn^{3D} constructs using a fluorescent reporter activity assay (**Figure 1-4 B, 1.5**).[22, 23] We observed that while indole **4** potently inhibited all four SFKs, all other derivatives in this series demonstrated high selectivity within the SFKs. Strikingly, all inhibitors are more potent against group B SFKs Hck and Lyn over the group A members Src and Fyn. Notably, **7**, **15**, **17**, **19**, **20**, **22**, and **23** potently inhibited both Hck^{3D} and Lyn^{3D} but had minimal activity against Src^{3D} and Fyn^{3D} at the highest concentration tested (30 μM). Within this group of Hck/Lyn-selective inhibitors, a structural motif consisting of a 5-membered heterocycle with a 1, 3-relationship between the aryl group and the amide linkage to the 2-methoxyaniline scaffold is common. Some of the inhibitors in this series were also able to discriminate between Lyn^{3D} and Hck^{3D}, with **10**, **11**, and **12** demonstrating at least 40-fold selectivity for Lyn^{3D} over Hck^{3D}–with no inhibition of Src^{3D} or Fyn^{3D}. These Lyn-selective inhibitors all contain a 2-carboxyindole with a substituent at the 4-position. Thus, by introducing structural diversity off of

the 2-methoxyaniline at the *C*-5 position it is possible to discriminate between the highly conserved ATP-binding sites of the SFKs.

A.



B. K_I (nM) Table

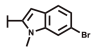
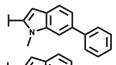
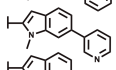
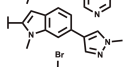
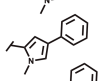
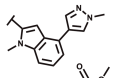
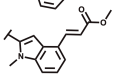
Compound	R_1	Hck^{3D}	Lyn^{3D}	Src^{3D}	Fyn^{3D}	Compound	R_1	Hck^{3D}	Lyn^{3D}	Src^{3D}	Fyn^{3D}
4		< 3.0	< 3.0	6.5 ± 1.5	16 ± 1	13		330 ± 20	40 ± 10	> 2000	> 2000
5		4.7 ± 0.5	< 3.0	180 ± 20	> 2000	14		> 2000	29 ± 8	> 2000	> 2000
6		3.9 ± 1.5	< 3.0	130 ± 20	> 2000	15		26 ± 3	3.3 ± 0.4	> 2000	> 2000
7		14 ± 2	6.5 ± 1.5	> 2000	> 2000	16		> 2000	61 ± 1	> 2000	> 2000
8		5.4 ± 0.8	< 3.0	110 ± 10	> 2000	17		< 3.0	< 3.0	> 2000	> 2000
9		210 ± 10	58 ± 1	> 2000	> 2000	18		310 ± 20	250 ± 20	> 2000	> 2000
10		> 2000	45 ± 12	> 2000	> 2000	19		3.5 ± 0.6	5.2 ± 0.4	> 2000	> 2000
11		> 2000	8.7 ± 1.1	> 2000	> 2000	20		7.2 ± 0.1	4.9 ± 0.7	> 2000	> 2000
12		280 ± 10	5.6 ± 1.5	> 2000	> 2000	21		12 ± 1	80 ± 10	50 ± 10	260 ± 20
						22		13 ± 2	25 ± 5	> 2000	> 2000

***Figure 1-4*: Pyrrolo[2,3-d]pyrimidine-based SFK inhibitors.**

A. Structure and proposed ATP-binding mode of the pyrrolopyrimidine-based inhibitors generated and tested. **B**. K_I values of inhibitors **4**-**22** against Src^{3D}, Fyn^{3D}, Hck^{3D}, and Lyn^{3D} (mean ± sem, n=3).

Kinase	K_m [ATP] (μM)
Hck ^{3D}	88 ± 12
Hck ^{CD}	57 ± 7
Lyn ^{3D}	85 ± 15
Lyn ^{CD}	55 ± 5
Src ^{3D}	67 ± 11
Src ^{CD}	37 ± 14
Fyn ^{3D}	77 ± 12
Abl ^{3D}	68 ± 8

Figure 1-5: Michaelis-Menten constant calculation. (K_m [ATP]) of SFK^{3D}, SFK^{CD} and Abl^{3D} (mean ± sem, n=3).

Global profiling of inhibitor selectivity. Next, we investigated how the selectivity we observed within the SFKs was reflected throughout the broader kinome. To determine this, we tested four representative inhibitors–**7**, **9**, **12**, and **19**–in a mass spectrometry-based lysate profiling assay (**Figure 1-6 A)**.[24-27] This kinase profiling assay involves determination of the level of competition an inhibitor provides for enriching endogenous kinases in an HCT116/HEK293 cell lysate mixture by an immobilized, non-selective, ATP-competitive inhibitor matrix. We tested each inhibitor at a single high concentration–30 μM–to identify strong and weak kinase interactors. In total, ~160 kinases were quantitatively profiled in these experiments. We found that **7**, **9**, **12**, and **19** possess fairly narrow specificity profiles and that Lyn was their most competed target (**Figure 1-6 B-C, 1-7**). Amongst the SFKs, Yes and Hck showed the highest level of competition after Lyn. Similar to our *in vitro* inhibition assays with purified constructs, Src and Fyn, as well as closely related FRK, showed significantly lower

levels of competition. Furthermore, we observed that our pyrrolopyrimidine-based inhibitors were generally selective for the SFKs, with only a handful of non-SFK targets that appeared to be of low

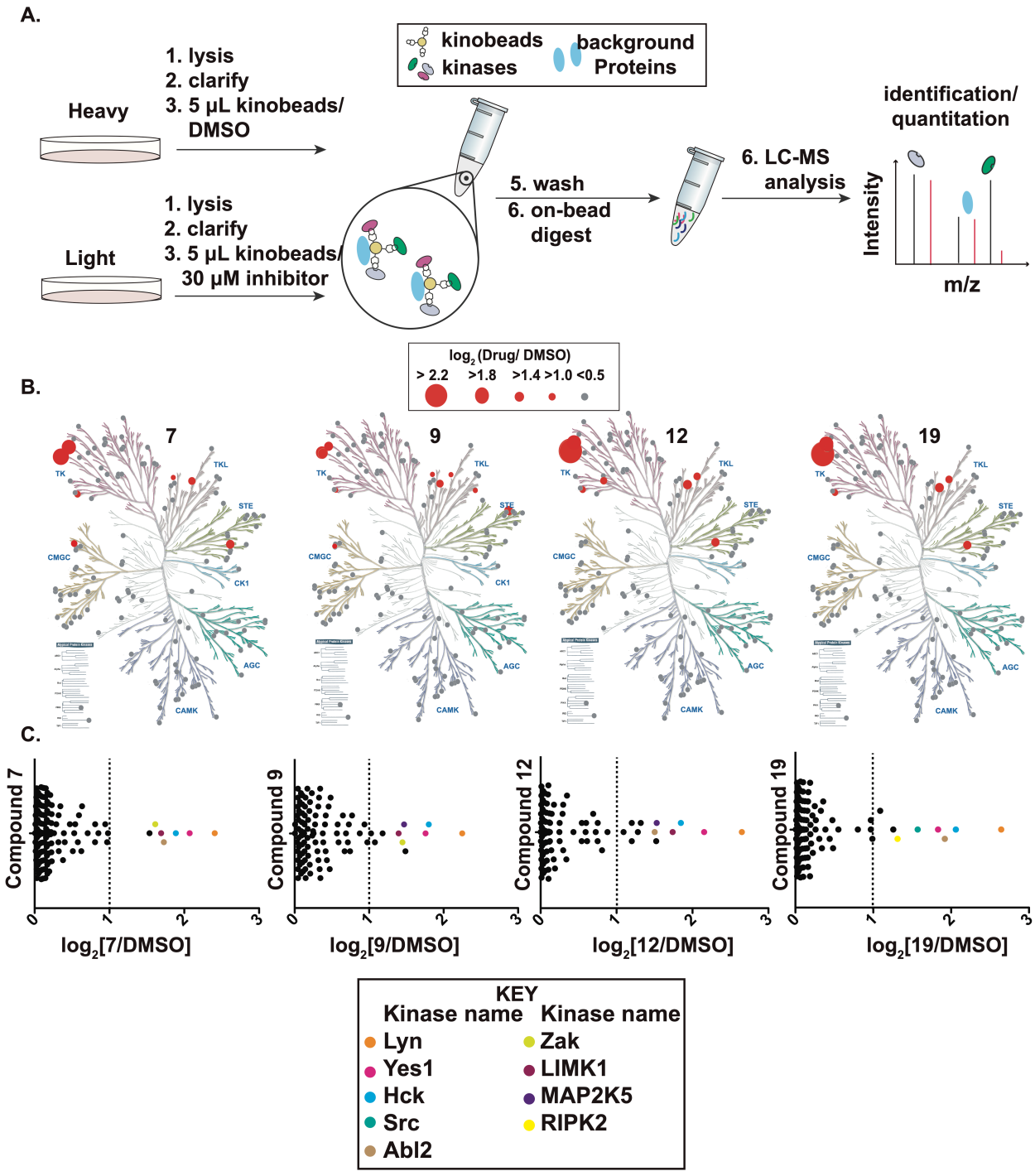


Figure 1-6: Chemoproteomic profiling of representative inhibitors.

A. Schematic of the kinobead-based profiling method used for profiling inhibitors **7**, **9**, **12** and **19**.[27] SILAC-labeled HEK293/HCT116 lysates were incubated with DMSO or 30 μM of the inhibitor being profiled and an immobilized matrix of non-selective kinase inhibitors (kinobeads). Captured proteins were subjected to on-bead digestion, identified and quantified by LC/MS analysis. **B.** Phylogenetic trees showing the kinases that were profiled in the chemoproteomic profiling experiments.[28] All profiled kinases are represented by grey circles and labeled by name. The size of the circle corresponds to relative percent competition (larger circle, more competed). Each profiling experiment was performed in duplicate. **C**. Transverse water-fall plots of the profiling data shown in **B.**

to moderate affinity. To obtain an indication of what type of inhibitory potency this low to moderate competition translates to, we performed activity assays with a recombinant construct of Abl1, which was competed at comparable, albeit slightly lower, levels as Abl2/Arg by **9**, **12**, and **19** but did not match either our selected cutoff for quantification of significance (compounds **12** and **19**) or was not quantified (compound **7**) (**Figure 1-8**). We observed that **7**, **9, 12**, and **19** were at least 60-fold less potent in inhibiting the activity of Abl1 (**Figure 1-9**). Thus, our pyrrolopyrimidine-based inhibitors are, in general, selective for Lyn within the SFKs and across the wider kinome.

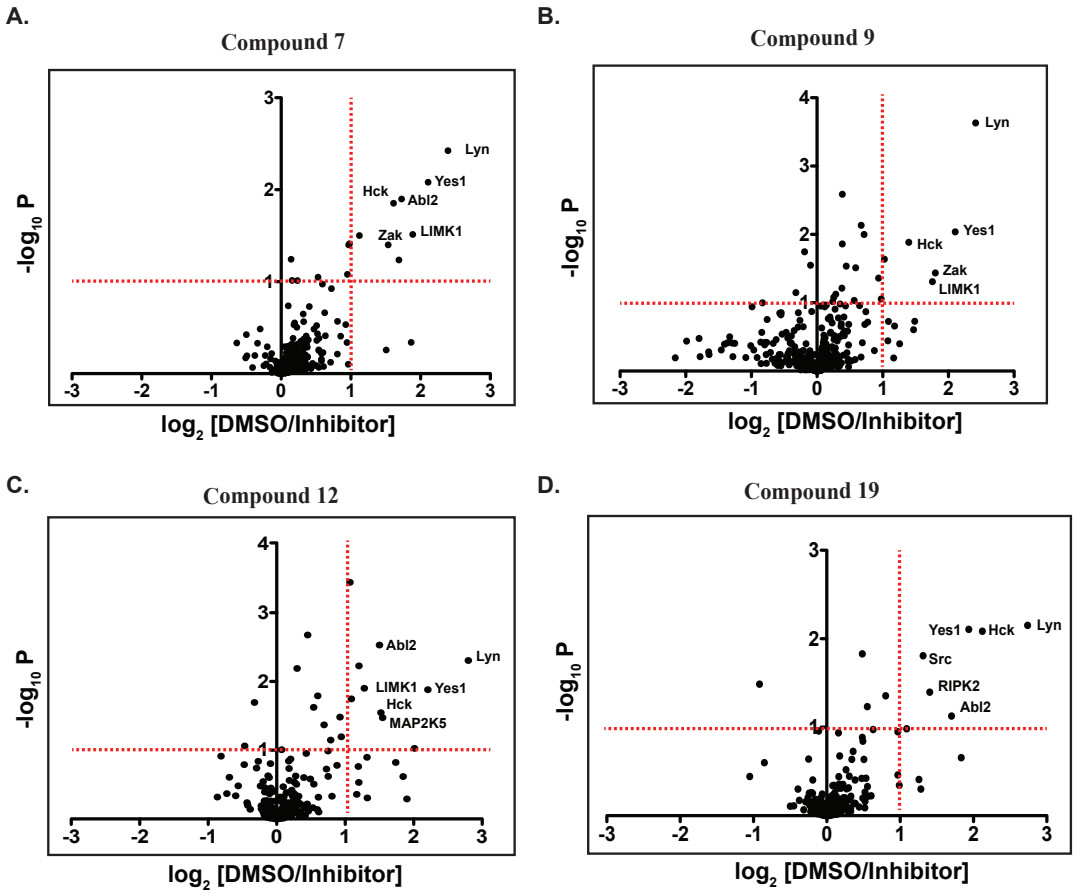


Figure 1-7: Volcano plots of chemoproteomic profiling of representative inhibitors.

A. plot showing profiling data for compound **7**. **B**. Volcano plot showing profiling data for compound **9**. **C**. Volcano plot showing profiling data for compound **12**. **D**. Volcano plot showing profiling data for compound **19**. The red horizontal line denotes our statistical cutoff ($-\log_{10}P < 1$) for quantification of hits.

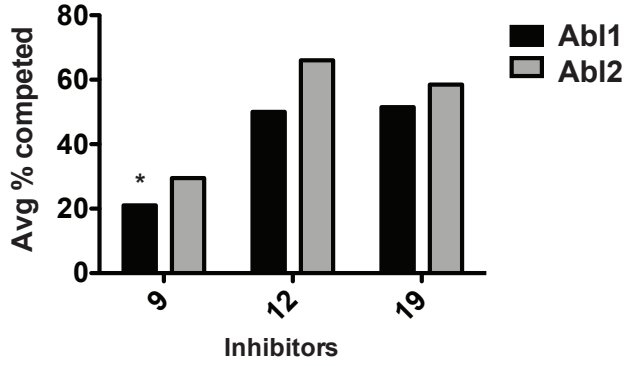
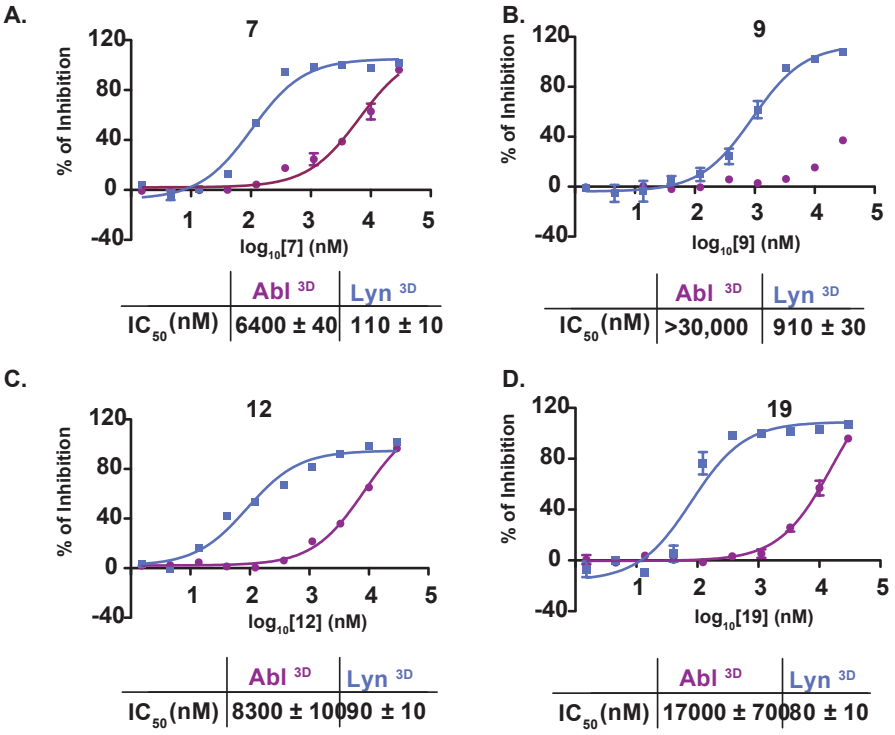


Figure 1-8: Comparison of average percent competition for Abl1 and Abl2 by compounds 9, 12 and 19 in the kinobead chemoproteomic profiling experiments (mean, n=2).

The values shown for Abl1 are not represented in **Figure 1.6** because quantified values were below our selected cutoff ($-\log_{10}P < 1$) for statistical significance. For compound **9**, the value shown is from a single replicate (denoted by *) because Abl1 was only quantified in one replicate. Abl1 was not quantified ($-\log_{10}P < 1$) in either replicate of competition experiments with compound **7**.

Influence of activation loop phosphorylation. Next, we determined how activation loop phosphorylation affects the potency and selectivity of our Lyn/Hck-selective inhibitors. We observed that the activation loops of Lyn, Hck, and Src undergo near quantitative phosphorylation on Tyr 416 during the pre-incubation step in our activity assays (**Figure 1-10**). Because activation loop phosphorylation can affect the potencies of some inhibitors for non-receptor tyrosine kinases,[29, 30] we investigated if our inhibitors are sensitive to this modification. To do this, we developed an assay for ATP-binding site occupancy that does not require measuring kinase activity and can be used with quantitatively dephosphorylated SFKs. Specifically, we used a time-resolved fluorescence resonance energy transfer (TR-FRET)-based competitive binding assay between a Europium-labeled, anti-His_6 antibody–bound to the N-termini of SFKs–as a FRET donor and a Cy5-labeled, ATP-competitive probe (SCP-2) as a FRET acceptor (**Figures 1-11 A**,**B, 1-12**).[31]



	Abl 3D	Lyn 3D
IC_{50} (nM)	6400 ± 40	110 ± 10

	Abl 3D	Lyn 3D
IC_{50} (nM)	>30,000	910 ± 30

	Abl 3D	Lyn 3D
IC_{50} (nM)	8300 ± 100	90 ± 10

	Abl 3D	Lyn 3D
IC_{50} (nM)	17000 ± 700	80 ± 10

E.

Abl 3D	7	12	19
K_I (nM)	410 ± 10	530 ± 50	>1000

Figure 1-9:Inhibitory curves of representative inhibitors against purified Abl and Lyn.

A-D. (top) Inhibition curves of **7**, **9**, **12**, and **19** against Abl^{3D} and Lyn^{3D} using a fluorescent reporter activity assay and 1 mM ATP (mean ± sem, n=3). (bottom) IC_{50} values for Abl^{3D} and Lyn^{3D}. **E**. K_I (nM) values of inhibitor **7**, **12**, and **19** for Abl^{3D} using the Cheng-Prusoff equation.

With our TR-FRET assay, we first determined the *K*$_I$ values of eight–**5**, **6**, **7**, **9**, **10**, **12**, **19** and **22**–representative inhibitors for dephosphorylated Lyn. Unlike DFG-out-stabilizing inhibitor **3**, all eight inhibitors were nearly equipotent for phosphorylated and dephosphorylated Lyn (**Figure 1-11 C, 1-13**). Thus, despite targeting a conformationally flexible region of the SFKs, our Lyn/Hck-selective inhibitors are not sensitive to kinase phosphorylation state. We also tested **5**, **6**, **7**, **9**, **10**, **12**, **19** and **22** in the TR-FRET assay with Hck and Src. Like Lyn, Hck's and Src's sensitivity to this panel of inhibitors was minimally influenced by their phosphorylation states (**Figure 1-11 D, 1-13**). In sum, the phosphorylation state of the SFKs does not appear to have a major effect on their affinities for our

panel of Lyn/Hck-selective inhibitors. Furthermore, differences in SFK activation loop phosphorylation in our activity assays cannot account for the dramatic selectivity we observe.

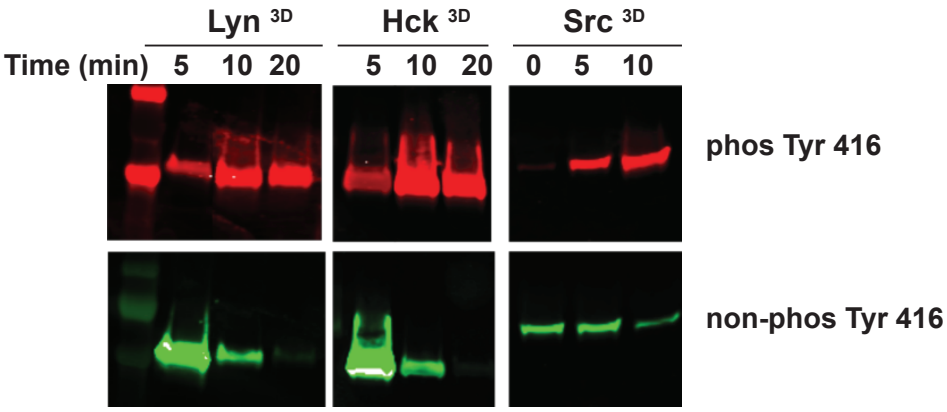


Figure 1-10: Activation loop phosphorylation of SFKs.

Western blots showing phosphorylated Tyr 416 (phos Tyr 416) and non-phosphosphorylated Tyr 416 (non-phos Tyr 416) at the time points indicated for Lyn^{3D}, Hck^{3D}, and Src^{3D}.

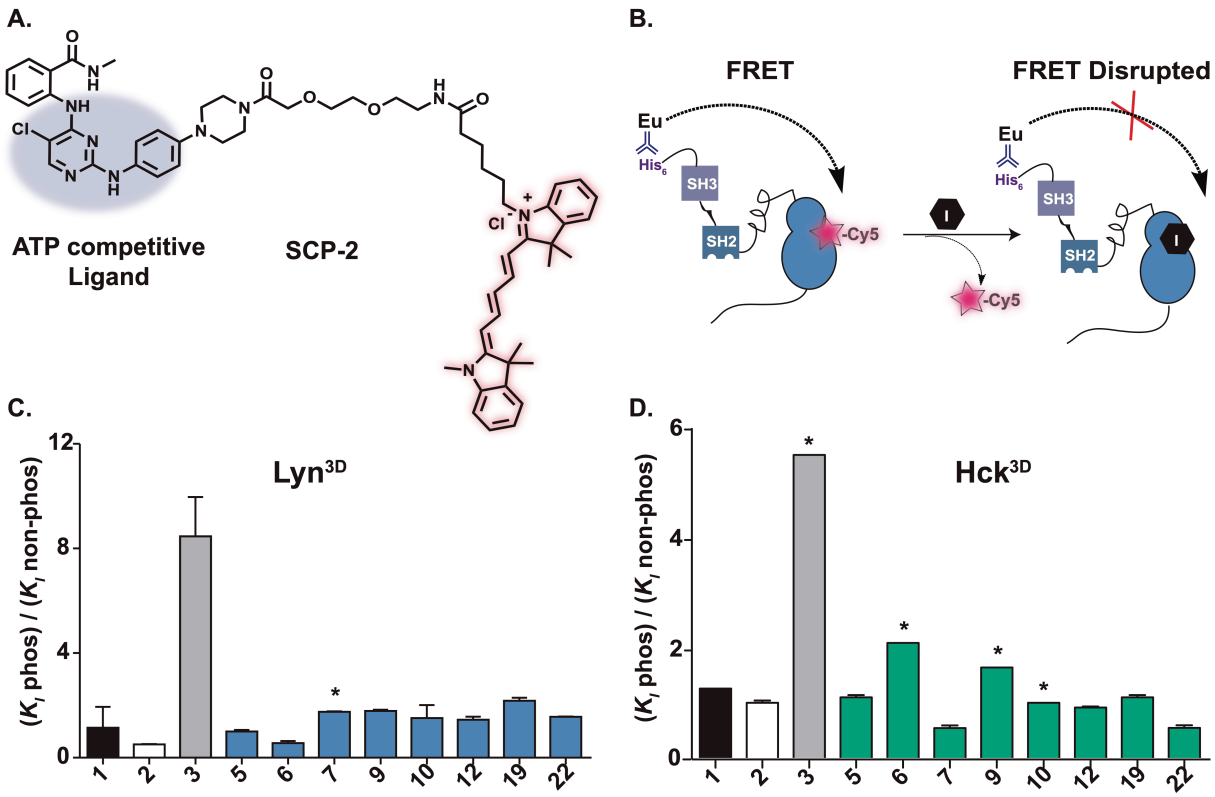


Figure 1-11:SFK activation loop phosphorylation minimally influences inhibitor potency.

A. Chemical structure of Cy5-labeled probe SCP-2. **B**. Schematic of the TR-FRET-based competitive binding assay. A Europium-labeled, anti-His_6 antibody acts as a FRET donor for the Cy5-labeled, ATP-competitive ligand SCP-2. Displacement of SCP-2 with an inhibitor leads to a loss in TR-FRET. **C**. Ratio of K_I values determined for phosphorylated and non-phosphorylated Lyn^{3D} (n=3). **D**. Ratio of K_I values determined for phosphorylated and non-phosphorylated Hck^{3D} (n=3). K_I values for phosphorylated Lyn^{3D} and Hck^{3D} were determined with an activity assay. K_I values for non-phosphorylated Lyn^{3D} and Hck^{3D} were determined with the competitive binding assay shown in **B**. If one or more of the K_I values measured approached the enzyme concentration used in an assay, ratios are denoted with a *. K_I values from the TR-FRET-based competitive binding assay are shown in **Figure 1.12B**.

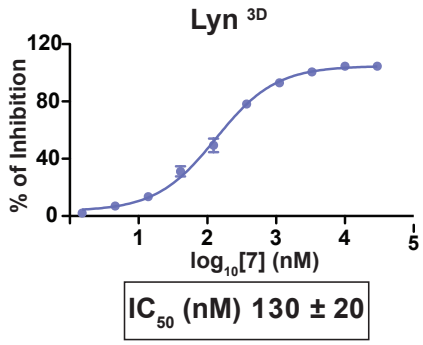
Kinase	K_D [SCP-2](nM)
Hck^{3D}	130 ± 10
Lyn^{3D}	49 ± 5
Src^{3D}	43 ± 1

Figure 1-12: Dissociation constant of SCP-2. K_D values of SCP-2 against Lyn^{3D}, Hck^{3D}, and Src^{3D} (mean ± sem, n=3).

Stabilization of the helix αC-out conformation. We next performed mechanistic studies to further define how Lyn/Hck-selective inhibitors interact with the ATP-binding sites of SFKs. Our pyrrolopyrimidine-based inhibitors were designed to project their *C*-5 substituents into a pocket that is available if helix αC in the N-terminal lobe of the kinase domain moves into an inactive conformation–the helix αC-out inactive conformation. We have previously demonstrated that inhibitors that stabilize the helix αC-out inactive conformation of SFKs allosterically promote a more closed global conformation with enhanced intramolecular SH2 and SH3 domain engagement.[19, 20, 32, 33] In contrast, we have observed that inhibitors that stabilize the DFG-out inactive conformation of the ATP-binding site promote a more open global SFK conformation with SH2 and SH3 domains that are freely accessible to intermolecular binding partners. To test whether our Lyn/Hck-selective

inhibitors promote the helix αC-out inactive conformations of Lyn and Hck as designed, a representative panel of SFK-inhibitor complexes were tested in two biochemical assays of SH2 and SH3 domain accessibility.[19, 20, 23, 32, 33]

A.



B.

Compound	phos Hck 3D	non-phos Hck 3D	phos Lyn 3D	non-phos Lyn 3D	phos Src 3D	non-phos Src 3D
	K_I (nM)					
5	4.7 ± 0.5	4.3 ± 0.5	< 3.0	< 3.0	180 ± 20	160 ± 10
6	3.9 ± 0.5	<3.0	< 3.0	< 3.0	130 ± 10	210 ± 40
7	14 ± 2	26 ± 5	6.5 ± 1.5	3.7 ± 0.9	> 2000	>2000
9	210 ± 10	130 ± 10	58 ± 4	33 ± 12	> 2000	> 2000
10	>2000	>2000	44 ± 8	29 ± 5	> 2000	> 2000
12	250 ± 20	280 ± 20	5.6 ± 1.4	3.7 ± 0.2	> 2000	> 2000
19	3.5 ± 0.6	3.2 ± 1.1	5.2 ± 0.4	2.4 ± 0.6	> 2000	> 2000
22	13 ± 2	24 ± 3	25 ± 5	16 ± 2	> 2000	> 2000

C.

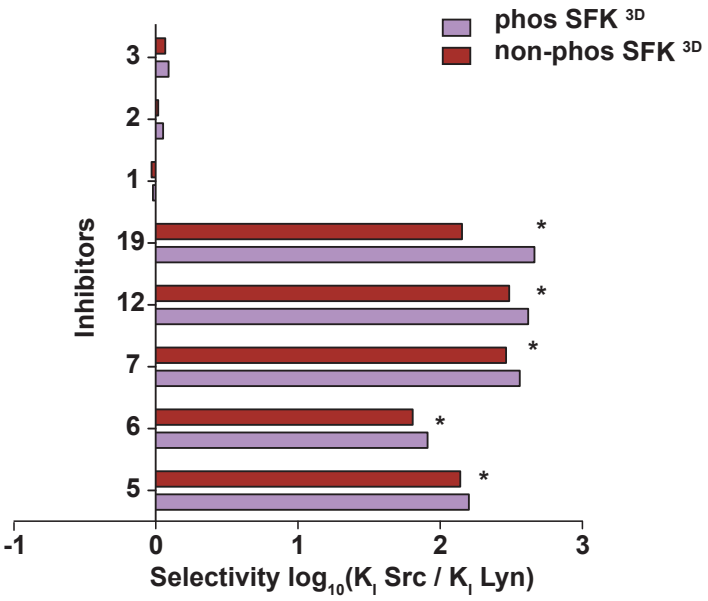


Figure 1-13: K_I comparisons of phosphorylated and non-phosphorylated SFKs.

A. Representative IC_{50} curve for compound **7** with Lyn^{3D} (mean ± sem, n=3) obtained from the TR-FRET based competitive binding assay. **B**. K_I values (nM) of inhibitors for phosphorylated and non-phosphorylated HCK^{3D}, Lyn^{3D}, and Src^{3D} (mean ± sem, n=3). **C**. Ratio of K_I values quantitative comparison of $\log_{10}$ of Src to K_I values of Lyn show selectivity of compounds remain unaltered between phosphorylated and non-phosphorylated SFKs. K_I > 2000 nM for Src against inhibitors **5**, **6**, **7**, **12** and **19**, is therefore a lower estimate, but a value of 2000 was used to measure the ratio and ratios are hence indicated by a *.

First, we tested the intermolecular accessibility of representative SFK^{3D}-inhibitor complexes in a pulldown assay with an immobilized SH3 domain ligand. To allow quantitative comparisons of the amount of Lyn pulled down in this assay, we generated a Lyn^{3D} construct with a tetramethylrhodamine (TMR) fluorophore at its *N*-terminus using the transpeptidase Sortase A (**Figures 1-14, 1-15A**).[34] Using the TMR-labeled Lyn construct, we confirmed that inhibitors **1**-**3** lead to expected levels of intermolecular SH3 domain accessibility (**Figure 1-14B, 1-15B**). TMR-Lyn^{3D} bound to helix αC-out-stabilizing inhibitor **2**, which promotes a closed global SFK conformation, demonstrated minimal pulldown in this assay, while TMR-Lyn^{3D} bound to DFG-out-stabilizing inhibitor **3**, which promotes an open global SFK conformation, was efficiently retained by the immobilized SH3 domain ligand. As expected, complexation to inhibitor **1**, which has a minimal influence on the global conformation of SFKs, led to an intermediate level of retained TMR-Lyn^{3D}. With these benchmarks in hand, we next tested the intermolecular accessibility of TMR-Lyn^{3D}'s SH3 domain when complexed to a representative set of pyrrolopyrimidine-based inhibitors. We found that all TMR-Lyn^{3D}-inhibitor complexes exhibited limited pulldown with the immobilized SH3 domain ligand, with retained levels comparable to the TMR-Lyn-**2** complex (**Figure 1-14 B, 1-15 B**). These results strongly suggest that the pyrrolopyrimidine-based inhibitors in our panel promote displacement of Lyn's helix αC from an active conformation, resulting in a closed global conformation with strengthened intramolecular regulatory interactions. We also observed that TMR-Hck^{3D} exhibited limited pulldown (**Figure 1-14 B, 1-15 C)** when bound to a representative set of pyrrolopyrimidine-based inhibitors, demonstrating that the binding mode of our Lyn/Hck-selective inhibitors is likely conserved amongst the SFKs.

To probe how inhibitor binding affects intramolecular SFK SH2 domains engagement, we monitored C-terminal tail phosphorylation of SFK^{3D}-inhibitor complexes by C-terminal Src kinase (Csk).[35-38] We have previously demonstrated that the amount of C-terminal tail phosphorylation by Csk is highly sensitive to the global conformation of SFK^{3D} constructs. Open, regulatory domain-disengaged SFKs

are efficient Csk substrates, while closed, regulatory domain-engaged SFKs exhibit limited phosphorylation by Csk (**Figure 1-14 C, 1-14 D**).[19, 20] Prior to testing inhibitors in the C-terminal tail phosphorylation assay, we first determined whether the required concentrations (3-5 μM) used to form near quantitative SFK-inhibitor complexes directly inhibit the activity of the Csk construct used in this assay. To do this, we monitored Csk's ability to phosphorylate a Lyn construct–Lyn^{CD}–that consists of only the catalytic domain (CD) and C-terminal tail. Because Lyn^{CD} does not contain an SH2 domain capable of shielding the C-terminal tail from phosphorylation, any diminution observed must be due to direct inhibition of Csk. Because our prototypical helix αC-out-stabilizing inhibitor **2** reduced the ability of Csk to phosphorylate Lyn^{CD}, it could not be used as a benchmark for C-terminal tail phosphorylation levels when Lyn is in the closed global conformation. However, inhibitors **1** and **3** displayed no such liability. Consistent with their high kinome selectivity, all representative pyrrolopyrimidine-based inhibitors tested minimally affected Csk's ability to phosphorylate Lyn^{CD} (**Figure 1-16 A**). Similar to our observations with other SFKs, Lyn^{3D} was a more efficient substrate for Csk when complexed to DFG-out-stabilizing inhibitor **3**, which promotes an open global conformation, than when bound to **1**.[19] All representative pyrrolopyrimidine-based inhibitor-Lyn^{3D} complexes tested displayed lower levels of C-terminal tail phosphorylation than the **1**-Lyn^{3D} complex, which further suggests that they stabilize the helix αC-out, ATP-binding site conformation and promote a closed, autoinhibited global conformation. We also observed that representative pyrrolopyrimidine-based inhibitors similarly influenced Csk's ability to phosphorylate the C-terminal tail of Hck^{3D} (**Figure 1-14 D, 1-16 B**). Together, these biochemical studies suggest that our pyrrolopyrimidine-based inhibitors stabilize the helix αC-out conformation of SFKs and that contacts within this region most likely are responsible for inhibitor potency and selectivity

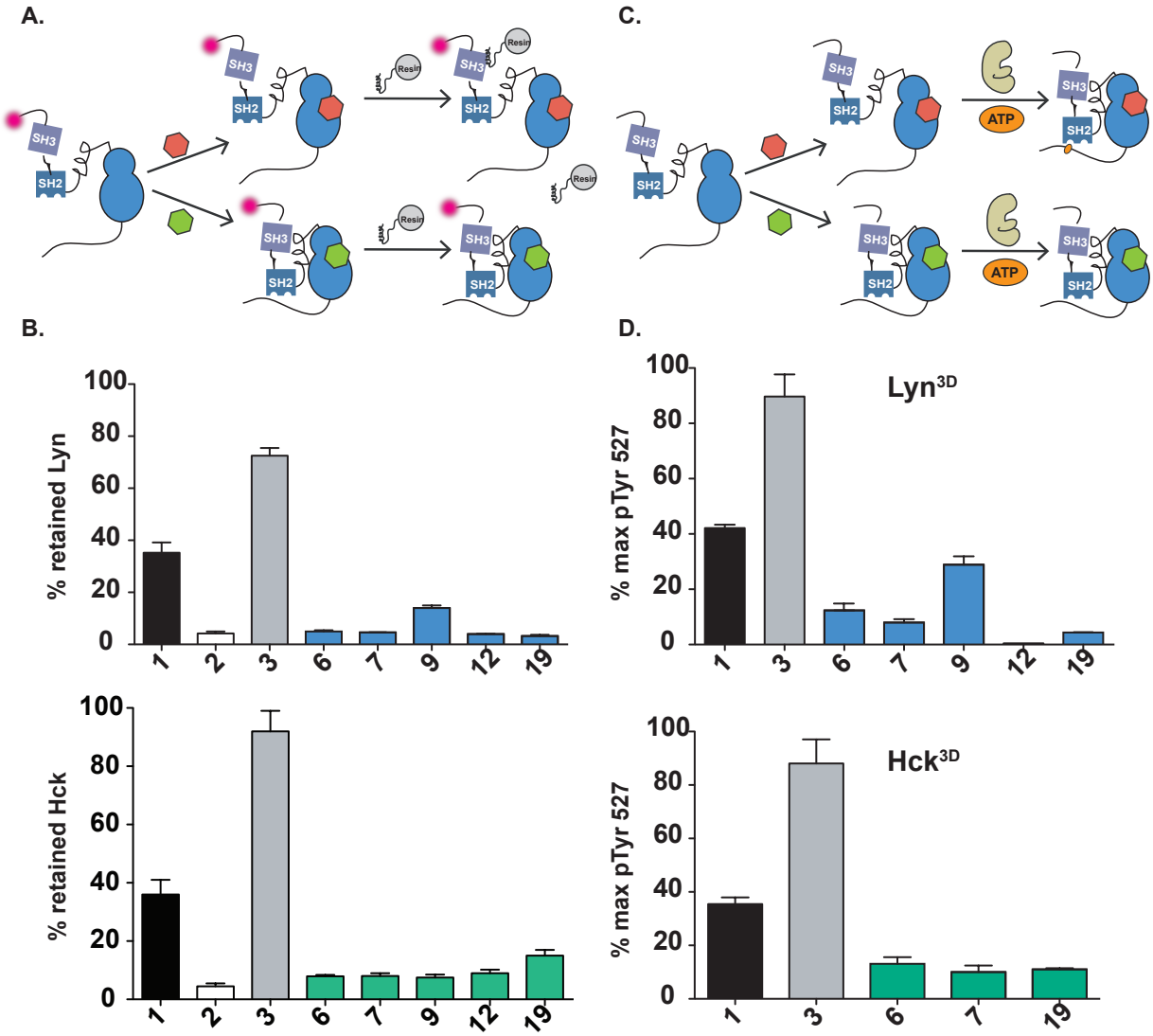


Figure 1-14: Lyn/Hck-selective inhibitors behave like helix αC-out-stabilizing ligands.

A. Schematic of the SH3 domain ligand pulldown assay using TMR-Lyn^{3D} and TMR-Hck^{3D}. SFK^{3D}-inhibitor complexes are incubated with an immobilized SH3 domain ligand and the amount of retained kinase is determined after incubation and resin washing. Retained SFKs are quantified after elution using fluorescence imaging. **B**. Percent retained Lyn^{3D} (*top*) and Hck^{3D} (*bottom*) inhibitor complexes in the SH3 pulldown assay. Values shown are mean ± sem (n=3). **C.** Schematic of the SH2 domain accessibility assay using Csk. SFK^{3D}-inhibitor complexes are incubated with a drug-resistant mutant of Csk (T338I) and γ^{32}P-ATP. Radioactive phosphate transfer to the C-terminal tail of SFKs is quantified. Closed, SH2-engaged Src cannot be phosphorylated by Csk whereas open, SH2-disengaged Src is efficiently phosphorylated by Csk. **D.** Quantification of Lyn's (*top*) and Hck's (*bottom*) C-terminal phosphorylation by Csk T338I. Values shown are mean ± sem (n=3).

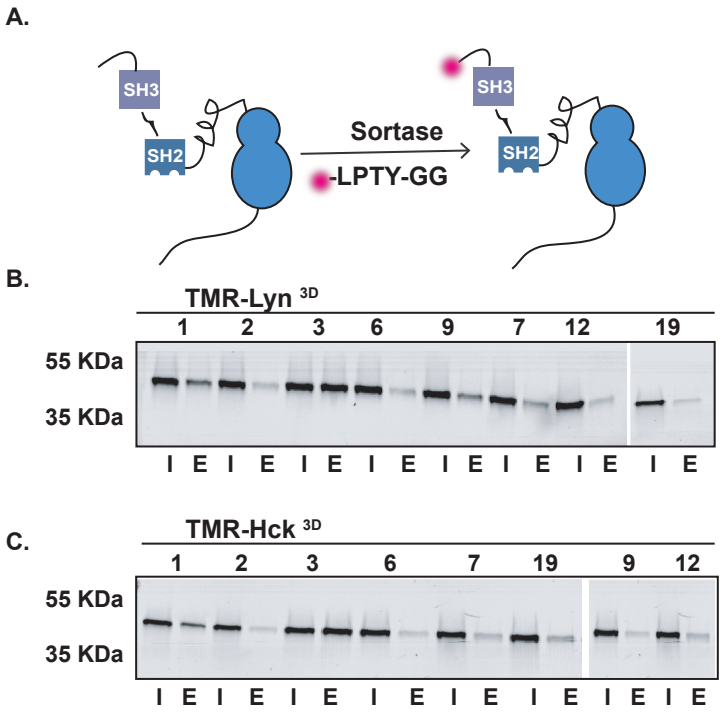


Figure 1-15: SH3 pulldown assay.

A. Schematic for generating TMR-Lyn^{3D} and TMR-Hck^{3D} using trans-petidase enzyme Sortase. **B.** Representative rhodamine gels showing input (I) and eluted (E) amounts of TMR-Lyn^{3D}. **C**. Representative rhodamine gels showing input (I) and eluted (E) amounts of TMR-Hck^{3D}.

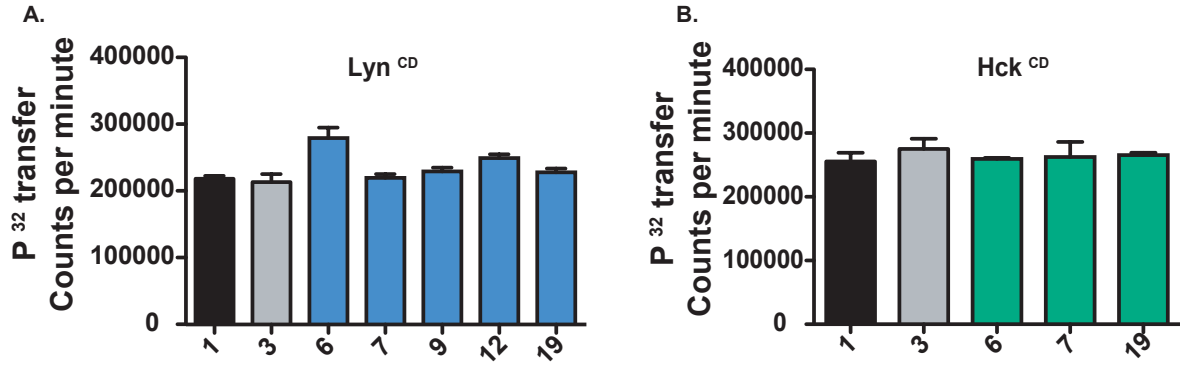


Figure 1-16: Csk assay.

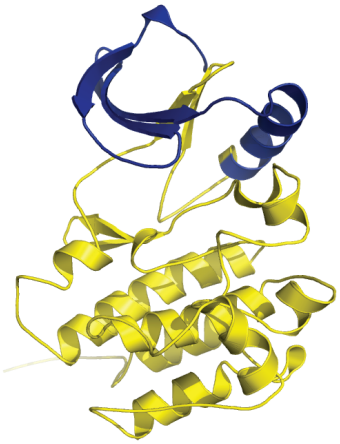
A. Raw phospho-counts of Lyn^{CD} C-terminal phosphorylation by Csk T338I in presence of different inhibitors. **B**. Raw phospho-counts of Hck^{CD} C-terminal phosphorylation by Csk T338I in presence of different inhibitors (mean ± sem, n=3).

Sequence determinants of inhibitor sensitivity. The mechanistic studies above suggest that our pyrrolopyrimidine-based inhibitors stabilize the helix αC-out conformation and likely require large movements within this region of the ATP-binding site. This suggests that their selectivity within the SFKs is most likely due to sequence differences within helix αC or in residues that influence the conformation of this region of the ATP-binding site. To determine if we could identify these sequence differences, we generated a series of chimeric constructs consisting of an inhibitor-sensitive SFK–Lyn–and inhibitor-resistant SFK–Src (**Figure 1-17 A**). Until this point, all inhibition and biochemical assays, except **Figure 1-16**, were performed with SFK constructs that possess an SH3, SH2 and CD (SFK^{3D} constructs). To limit sample search space and allow more facile protein expression and purification, we first determined whether the regulatory domains of Src or Lyn influence their sensitivities to pyrrolopyrimidine-based inhibitors. We found that Src and Lyn constructs containing only the catalytic domain (Src^{CD} and Lyn^{CD}) were similarly inhibited by a representative set of inhibitors as SFK^{3D} constructs, although **5** and **6** were less potent against Src^{CD} than Src^{3D} (**Figures 1-4 B**, **1-17 B**). Therefore, we generated and tested chimeras that only contain CDs and lack SH2 and SH3 domains. First, we generated a Lyn chimera ($LynS^{CD1}$) that consists of a majority of Src's N-terminal lobe by swapping residues 258-315 between the two SFKs (**Figure 1-17 A, 1-18**). Consistent with the N-terminal lobe serving as a key region of pyrrolopyrimidine-based inhibitor sensitivity, we found that $LynS^{CD1}$ was almost completely insensitive to a representative set of inhibitors (**Figure 1-17 B**).

Having established that Src's N-terminal lobe is sufficient to provide resistance to our pyrrolopyrimidine-based inhibitors, we next determined whether we could sensitize Src through the introduction of residues from Lyn's N-terminal lobe. To do this, we generated two Src chimeras–$SrcL^{CD1}$ and $SrcL^{CD2}$–that possess different regions of Lyn's N-terminal lobe (**Figure 1-17 C**). The Src chimera $SrcL^{CD1}$, which contains residues 258-314 from Lyn, displayed increased sensitivity to several inhibitors 5, **6**, **7**, **12**, and **19** (**Figure 1-17 D**). Furthermore, a Src chimera– $SrcL^{CD2}$–containing a

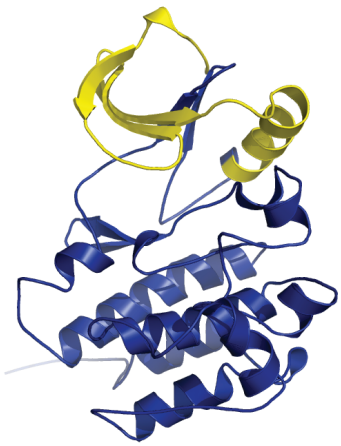
smaller region of Lyn's N-terminal lobe (residues 294-314) displayed a similar level of inhibitor sensitivity as SrcL^{CD1} (**Figure 1-17 E**). As both SrcL^{CD1} and SrcL^{CD2} contain most of helix αC from Lyn, we narrowed our focus on this region of the ATP-binding site.

A.

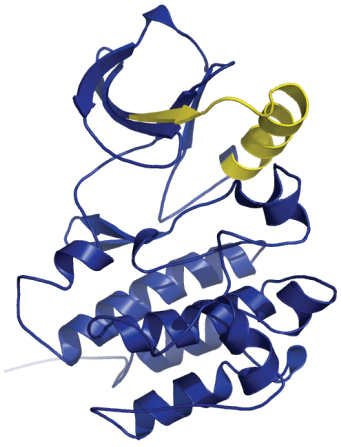


LynS^{CD1}
Residue 244 - 314: Src
Residue 314 - 533: Lyn

C.



SrcL^{CD1}
Residue 315 - 533: Src
Residue 244 - 314: Lyn



SrcL^{CD2}
Residue 244 - 293; 315 - 533: Src
Residue 294 - 314: Lyn

B.

Compound	K_I (nM) Lyn^{CD}	Src^{CD}	LynS^{CD1}
5	< 3.0	> 2000	49 ± 9
6	2.3 ± 0.1	> 2000	> 2000
7	4.6 ± 0.8	> 2000	> 2000
9	41 ± 9	> 2000	> 2000
10	33 ± 5	> 2000	> 2000
12	4.3 ± 0.5	> 2000	> 2000
19	2.5 ± 0.1	> 2000	> 2000
22	51 ± 7	> 2000	> 2000

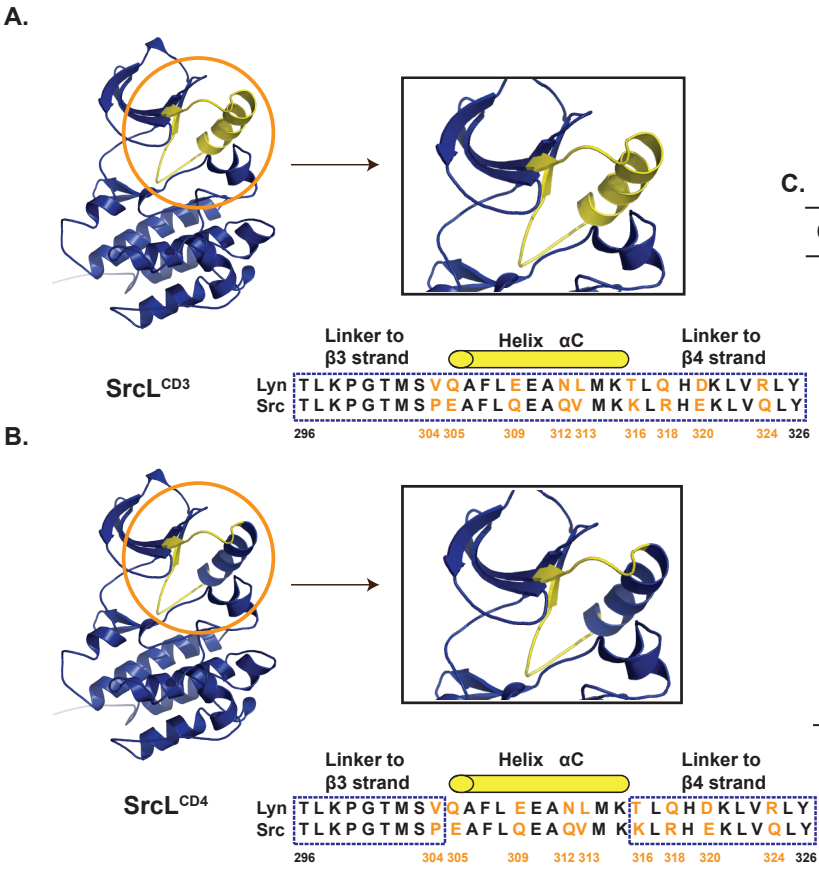
D.

Compound	K_I (nM) SrcL^{CD1}	SrcL^{CD2}
5	85 ± 5	78 ± 11
6	89 ± 11	66 ± 6
7	140 ± 30	152 ± 45
9	> 2000	> 2000
10	> 2000	> 2000
12	310 ± 80	> 2000
19	98 ± 8	150 ± 20
22	> 2000	> 2000

Figure 1-17: Inhibitor sensitivity of Lyn-Src chimeras.

A. Crystal structure of the SH1 domain (Src^{CD}) from PDB ID, 2SRC (Note: only the SH1 domain (CD) from 2SRC, which contains the SH3, SH2, and SH1 domains, is shown for all crystal structures in **Figures 6** and **7**

because all K_I values shown were determined against constructs that lack SH2 and SH3 domains). The regions of Lyn (*yellow*) that were swapped with regions from Src (*blue*) in chimera LynS^{CD1} are shown. **B.** K_I values of a representative panel of Lyn/Hck-selective inhibitors for Lyn^{CD}, Src^{CD}, and LynS^{CD1} (values shown are mean ± sem (n=3)). **C**. Crystal structures of Src^{CD} showing the regions of Src (blue) that were swapped with regions from Lyn (yellow) in chimeras SrcL^{CD1} (*left*) and SrcL^{CD2} (*right*). **D**. K_I values of a representative panel of Lyn/Hck-selective inhibitors for SrcL^{CD1} and SrcL^{CD2} (values shown are mean ± sem (n=3)).



C.

Compound	K_I (nM) SrcL^{CD-3}	SrcL^{CD-4}
5	11 ± 2	16 ± 4.6
6	13 ± 3	15 ± 2
7	21 ± 1	44 ± 8
9	98 ± 14	310 ± 10
10	85 ± 12	450 ± 20
12	18 ± 3	22 ± 3
19	29 ± 7	45 ± 7
22	57 ± 9	37 ± 9

Figure 1-18: Inhibitor-sensitive Src chimeras containing Lyn's helix αC.

A, **B.** Crystal structures of the SH1 domain of Src (Src^{CD}) (PDB ID, 2SRC) showing the regions of Src (*blue*) that were swapped with regions from Lyn (*yellow*) in chimeras SrcL^{CD3} and SrcL^{CD4}. Insets show an enlarged version of the area circled in the left panel. A sequence alignment of Lyn and Src showing the residues that have been swapped (*boxed*) in Src is shown below each inset. Residues that differ between Lyn and Src are colored orange. **C**. K_I values of a representative panel of Lyn/Hck-selective inhibitors for LynS^{CD3} and LynS^{CD4} (values shown are mean ± sem (n=3)).

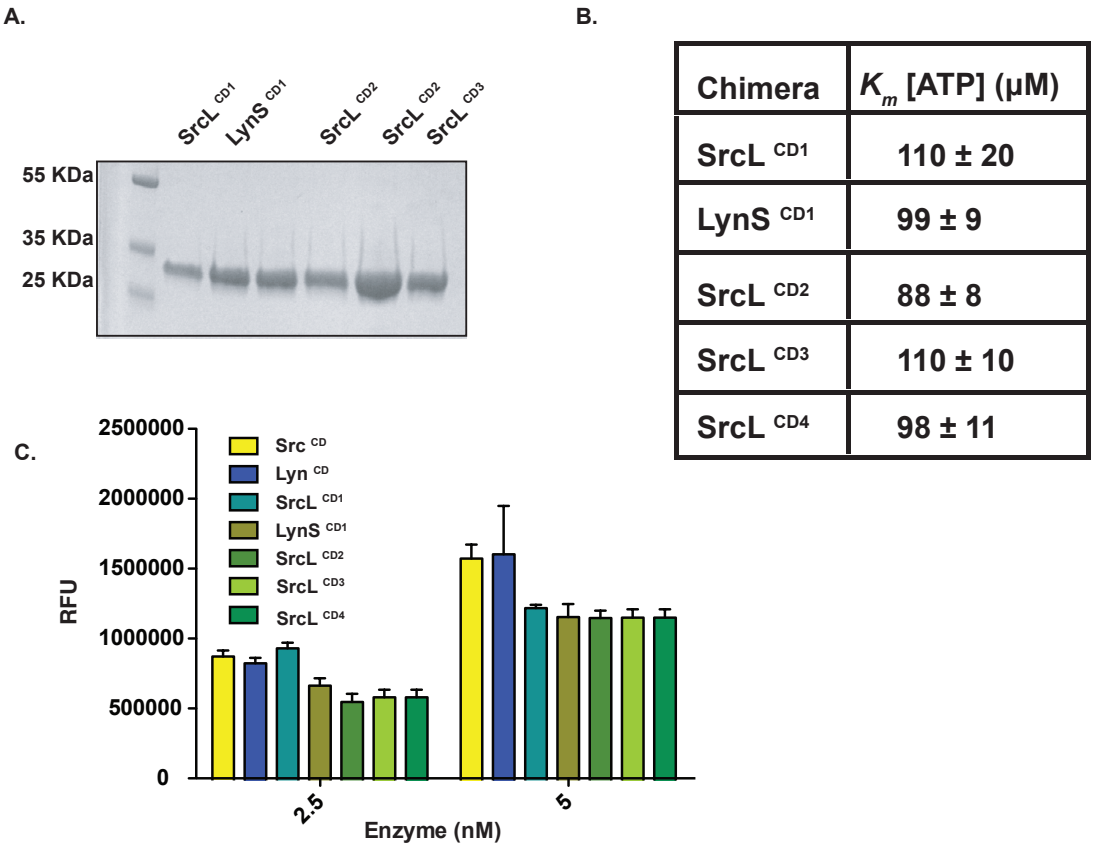
Next, we generated a Src chimera, SrcL^{CD3}, that contains the entire helix αC of Lyn and the linker residues that connect it to the β3 and β4 strands of the N-terminal lobe by swapping residues 304-324 between these two SFKs (**Figure 1-18 A, 1-19**). Remarkably, introducing this ~20 residue region

of Lyn into Src provided high sensitivity to all of the inhibitors in our representative panel (**Figure 1-18 B**). There are nine residues that differ between Src and Lyn in the region that spans residues 304-324. Four of these differing residues–305, 309, 312, and 313–are in the structured region of helix αC and five are in the residues that connect helix αC to the N-terminal lobe. To probe whether residues within helix αC or the linkers that connect this structured region to the N-terminal lobe are responsible for inhibitor sensitivity, we generated the Src chimera $SrcL^{CD4}$ (**Figure 1-18 C**). $SrcL^{CD4}$ contains the entire C- and N-lobes of Src, including the helical region of helix αC, except for the five sequence differences in the residues that connect helix αC to the N-terminal lobe from Lyn. Consistent with the conformational preference of Lyn's helix αC being responsible for its high sensitivity to Lyn/Hck-selective inhibitors, $SrcL^{CD4}$ is almost equally sensitive to our representative panel of Lyn/Hck-selective inhibitors as $SrcL^{CD3}$ (**Figure 1-18 D**). Thus, the conformational flexibility of Lyn's helix αC, which is controlled by N-lobe connecting residues, appears to be responsible for its sensitivity to Lyn/Hck-selective inhibitors rather than the presence or absence of specific interactions. For inhibitors like **9** and **10**, which are several-fold more potent against $SrcL^{CD3}$ than $SrcL^{CD4}$, Lyn's sensitivity may stem from both the conformational preference of helix αC and specific contacts with helix αC itself. Together, our data show that high selectivity amongst the SFKs can be achieved by targeting the conformational preference of helix αC.

1.4 Conclusions

Diverse strategies are needed for identifying inhibitors that are capable of discriminating between the ATP-binding sites of kinases that possess high sequence homology, like the SFKs. Here, we show that, unsurprisingly, the ATP-binding sites of SFKs are equally susceptible to different modes of ATP-competitive inhibition. However, by modifying a non-selective scaffold with substituents that project towards the conformationally-flexible helix αC, we were able to generate ATP-competitive inhibitors

that are highly selective for group B SFKs Hck and/or Lyn over Group A members Src and Fyn. Biochemical analyses of how these inhibitors allosterically modulate SFK regulatory domain engagement confirm that they displace helix αC from an active conformation. Our Lyn/Hck-selective inhibitors, in general, possess substituents that would be expected to require large movements of helix αC to be accommodated in SFK ATP-binding sites. Whether the helix αC s of Hck and Lyn are capable of larger movements or pay a smaller energetic penalty for adopting an inactive conformation than Src and Fyn remains to be determined. The ability to identify inhibitors that discriminate between



Chimera	K_m [ATP] (μM)
SrcL ^{CD1}	110 ± 20
LynS ^{CD1}	99 ± 9
SrcL ^{CD2}	88 ± 8
SrcL ^{CD3}	110 ± 10
SrcL ^{CD4}	98 ± 11

Figure 1-19: Src/Lyn chimeras.

A. SDS-PAGE analysis of recombinant chimeras SrcL^{CD1}, LynS^{CD1}, SrcL^{CD2}, SrcL^{CD2}, SrcL^{CD3}. **B**. K_m [ATP] of chimeras (mean ± sem, n=3). **C.** Comparison of the catalytic activities of Src^{CD}, Lyn^{CD} and all 5 chimeric constructs at enzyme concentrations 2.5 nM and 5 nM using a fluorescent reporter activity assay, 1 mM ATP and 20 μM of a fluorescence SFK peptide. Constructs were incubated with ATP and SFK peptide for 60 min and increase in (background subtracted) fluorescence is reported as Relative Fluorescence Unit (RFU,mean ± sem, n=3).

that are highly selective for group B SFKs Hck and/or Lyn over Group A members Src and Fyn. Biochemical analyses of how these inhibitors allosterically modulate SFK regulatory domain engagement confirm that they displace helix αC from an active conformation. Our Lyn/Hck-selective inhibitors, in general, possess substituents that would be expected to require large movements of helix αC to be accommodated in SFK ATP-binding sites. Whether the helix αCs of Hck and Lyn are capable of larger movements or pay a smaller energetic penalty for adopting an inactive conformation than Src and Fyn remains to be determined. The ability to identify inhibitors that discriminate between SFKs is somewhat reminiscent of the mechanism proposed for how the ATP-competitive inhibitor FRAX597 is able to selectivity inhibit Group II p21-Activated Kinases (PAKs) over Group I PAKs.[39] While FRAX597 does not displace helix αC from an active position in PAKs, the apparent increased conformational flexibility of a residue on Group II PAKs' helix αC allows more favorable accommodation of FRAX597 despite Group I PAKs containing identical contact residues in this region. Therefore, it is also a possibility that the helix αC of Lyn and Hck are no more mobile than that of Src and Fyn but that residues in this region can adopt unique rotamer configurations when these SFKs are in the helix αC-out conformation. Regardless of the exact mechanism of discrimination, out sequence swap experiments strongly implicate the flexible linker residues connected to helix αC as being responsible for inhibitor sensitivity.

1.5 Methods

Cloning, protein purification and expression. Hck^{3D}, Hck^{CD} constructs and Abl^{3D} were cloned into the bacterial expression vector pET28-a as an N-terminal His_6 construct. pTH1-SrtA (Sortase) was purchased from Addgene (plasmid # 64973). Everything else was cloned into the bacterial expression plasmid pMCSG7 as a N-terminal His_6 construct. All SFK and Abl constructs were co-transformed in BL21-DE3 cells with tyrosine phosphatase YopH (pCDFDuet-1) and chaperonin GroEL

(pACYCDuet-1) and plated on triple selective plates (Ampicilin for pMCSG7 or Kanamycin for pET28-a along with Chloramphenicol and Streptomycin). A single colony was picked and grown in an overnight culture of 15 mL of Terrific broth containing all three antibiotics. 1 L culture was then inoculated with the starter culture, grown to an OD_{600} of 1.1, temperature was then dropped to 18 ºC and protein expression was induced with 0.4 M IPTG and protein was expressed overnight. Ni-NTA was used to purify His_6-SFK and His_6-Abl constructs using purification buffer (50 mM HEPES, pH 8.0, 300 mM NaCl, 1 mM PMSF, 0.1% (v/v) Triton-X, 10% (v/v) glycerol, 0.2% (v/v) BME, 100 mM imidazole). Dialysis was done in dialysis buffer (50 mM HEPES, pH 8.0, 150 mM NaCl, 1mM DTT, 10% (v/v) glycerol). Finally, anion exchange column (Pierce, #90011) was used to remove YopH and GroEL to yield SFK or Abl at >95% purity.

Csk (T338I) was expressed and purified as a C-terminal GST tagged protein according to previously published protocol. [40]

All Lyn-Src Chimeras were cloned as N-terminal His_6 constructs into pMCSG7. They were expressed and purified like other SFKs.

Western blotting. All antibodies were purchased from Cell Signaling Technology (CST); Src non-phosphorylated Tyr 416 (CST #2102) and Src phosphorylated Tyr 416 (CST #2101).

Kinobead assay. The kinobead affinity reagents were immobilized on ECH Sepharose 4B (GE Healthcare Bio-Sciences) according to the previously published protocol.[27, 41] HCT116 and HEK293T cells were from the American Type Culture Collection (ATCC) and cultured in ATCC suggested media supplemented with Fetal Bovine Serum (FBS). Cells were grown in custom -Lys/-Arg DMEM (Caisson Labs) supplemented with 10% FBS (Sigma), 200 $\mu g\ ml^{-1}$ proline and SILAC amino acids (0.2 mM Lys0/Arg0 for light label, 0.2 mM Lys8/Arg10 for heavy label; Cambridge Isotope Labs, and over, MA). Cells were grown for at least 5 cell doublings in SILAC medium and harvested when they

reached 90% confluency. Concentration of inhibitor used was 30 μM and final concentration of lysate mixture was 2 mg mL^{-1}.

LC-MS/MS and data analysis. Analysis was performed as described previously.[27]

Activity assay to determine IC_{50} of inhibitors. Prior to testing the inhibitors, a kinase titration was carried out in order to determine the concentration of kinase to be used in the inhibitor titration assay. Inhibitors (initial concentration = 30 μM, 3-fold serial dilutions, 10 data points in triplicate) were assayed in triplicate against SFKs ([Hck ^{3D} and Hck ^{CD}] = 10 nM ; [Lyn ^{3D} and Lyn ^{CD}] = 7.5 nM; [Src ^{3D} and Src^{CD}] = 10 nM, [Fyn ^{3D}] = 5 nM and [Abl ^{3D}] = 7.5 nM) in assay buffer (75 mM HEPES, pH 8.0, 15 mM $MgCl_2$, 3.8 mM EGTA, 150 mM NaCl, 0.2 mg mL^{-1} BSA and 750 nM Na_3VO_4). Briefly, kinase was incubated with inhibitors and 1 mM ATP for 30 min in a 384-black assay plate (Corning, #3573). 20 μM of a fluorescence SFK peptide (Ac-EEEIYGE(Dap-Pyrene)-EA-NH_2) or 40 μM of Abl peptide (Ac-AEAIYAA(Dap-pyrene)-LA-NH_2) was then added to plate and incubated for 2 h. Raw fluorescence units were measured on Envision (Perkin Elmer) with excitation wavelength of 344 nm and emission wavelength of 405 nm. Data was analyzed using GraphPad Prism software, and IC_{50} values were determined using "One-site fit log IC_{50}". All IC_{50} values were then converted to K_I by implementing the Cheng-Prusoff's equation using K_m [ATP] determined for individual kinase (Figure S3). In order to check phosphorylation of SFK species, 30 μL of kinase in assay buffer described above containing 1 mM ATP was aliquoted out at various time points and added to 3x SDS-loading dye and amount phosphorylated and non-phosphorylated SFK was visualized using Src phosphorylated Tyr 416 (CST #2101; 1:2000 dilution) and Src non-phosphorylated Tyr 416 (CST #2102; 1:2500 dilution) respectively.

TR-FRET binding assay. Inhibitors (initial concentration = 30 μM, 3-fold serial dilutions, 10 data points) were assayed in triplicate against SFKs ([Hck^{3D}] = 10 nM; [Lyn^{3D}] = 7.5 nM; [Src^{3D}] = 10 nM) in assay buffer (50 mM HEPES, pH 8.0, 150 mM NaCl, 0.2 mg mL^{-1} BSA, 2 nM TR-FRET donor

Eu-anti-His_6, #AD0110). Briefly, kinase was incubated with inhibitors and 10 μM TR-FRET donor SCP-2 for 1 h in a 384-white assay plate (Proxiplate-384 Plus, #6008280).

Raw TR-FRET donor (at 615 nm) and acceptor (at 665 nm) units were measured on Envision (Perkin Elmer) with excitation wavelength of 340 nm. Data was analyzed using GraphPad Prism software, and IC_{50} values were determined using nonlinear regression analysis. All IC_{50} values were then converted to K_I by implementing the Cheng-Prusoff Equation using K_D [SCP-2] determined for individual kinase (Figure 1-12).

Sortase labeling of Hck^{3D} and Lyn^{3D}. 5(6)-Carboxytetramethylrhodamine N-succinimydil ester (Sigma Aldrich, CAS #150408-83-6) was conjugated to ε-5-Caproic acid-LPYTGG (purchased from GenScript) to make the TMR-LPYTGG peptide.

Both Hck^{3D} and Lyn^{3D} was expressed as an N-terminal His_6-SUMO construct with two glycine residues *N*-terminus to Thr 84 (Chicken Src numbering). SFKs were expressed as described above, after the first dialysis, His_6-Ulp1 was added to the dialysis cassette to cleave N-terminal His_6-SUMO releasing the N-terminal 'GG' motif. Anion exchange columns were then used to remove YopH and GroEL. Sortase labeling was performed as described previously.[34] Briefly, 10-fold excess of His_6-Sortase and TMR-LPYTGG peptide was added to the purified SFKs and incubated in the dark for 2 h at 4 °C in labeling buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 20 mM imidazole, 10% (v/v) glycerol, and 1 mM BME). His_6-Sortase and His_6-Ulp1 were removed using Ni-NTA and excess TMR-LPYTGG peptide was removed using Zeba desalting columns (ThermoFisher Scientific, #89882).

SH3 pulldown assay. Formation of the Kinase-Inhibitor Complex: TMR-Lyn^{3D} or TMR-Hck^{3D} (25 nM) was diluted in immobilization buffer (50 mM Tris, pH 7.6, 100 mM NaCl, 1 mM DTT and 0.2 mg mL^{-1} BSA). A saturating amount of the inhibitor of interest ([Compounds 1, 2 and 3 = 3 μM against TMR-Lyn^{3D} or TMR-Hck^{3D}]; [Compounds 6, 7, 9, 12 and 19 = 5 μM against TMR-Lyn^{3D}]; [Compounds 6, 7, and 19 = 5 μM against TMR-Hck^{3D}]; [Compounds 9, and 12 = 40 μM against

TMR-Hck^{3D}]) was added to make a kinase-inhibitor complex. The mixture was allowed to incubate for 30 min at room temperature before being loaded on the resin.

20 μL of a 50% slurry of SNAP-Capture pulldown resin (GE Healthcare, #17-0906-01) was placed in a micro-centrifuge tube. The resin was washed (3x, 10 bed volumes) with immobilization buffer. 8 μM of SNAP tag–polyproline peptide fusion (VSLARRPLPPLP) was loaded onto the resin at a final volume of 50 μL in buffer. The resin was rotated at room temperature for 1 h. After polyproline peptide immobilization, the resin was washed (3x, 10 bead volumes) and 100 μL of the kinase–inhibitor complex was loaded. The resin was allowed to shake at room temperature for 1 h. Post incubation, the flow through was collected, and the resin was washed (3x, 10 bead volumes). To elute the retained kinase, 100 μL of 1x SDS loading buffer was added and the beads were boiled at 95 °C for 10 min. All samples were separated via SDS–PAGE and visualized on GE-Typhoon FLA and quantified as in-gel rhodamine fluorescence. Data was analyzed using GraphPad Prism.

C-terminal Src kinase (Csk) accessibility assay. To determine if the concentrations (3-5 μM) of inhibitor required to form >95% SFK-inhibitor complex directly inhibits drug-resistant Csk (T338I), different inhibitors at concentrations 3-5 μM were incubated with Lyn^{CD} (200 nM) or Hck^{CD} (200 nM) in kinase reaction buffer (76 mM HEPES, pH 8.0, 5 mM $MgCl_2$, 150 mM NaCl, 0.2 mg mL^{-1} BSA, 250 μM Na_3VO_4, 1 mM DTT and 3.8 mM EGTA) for 1 h at room temperature. Following incubation 25 nM Csk was added and phosphorylation was initiated by the addition of γ^{32}P-ATP (0.08 μCi/well) at a final volume of 30 μL. The enzymatic reaction was run at room temperature for 1 h and then terminated by spotting 4.6 μL of the reaction mixture onto a nitrocellulose membrane. Membranes were washed with 0.5% phosphoric acid (3x, 10 min each wash) and dried, and the radioactivity was determined by phosphor-imaging with a GE Typhoon FLA 9000 phosphor-scanner. The scanned membranes were quantified using ImageStudio Lite and plotted using GraphPad prism to determine

γ^{32}P transfer onto SFK^{CD} C-terminal tails for each inhibitor. Our prototypical helix αC-out-stabilizing inhibitor 2 showed lower C-tail phosphorylation of SFK^{CD} and was therefore not used in the subsequent assay with SFK^{3D}.

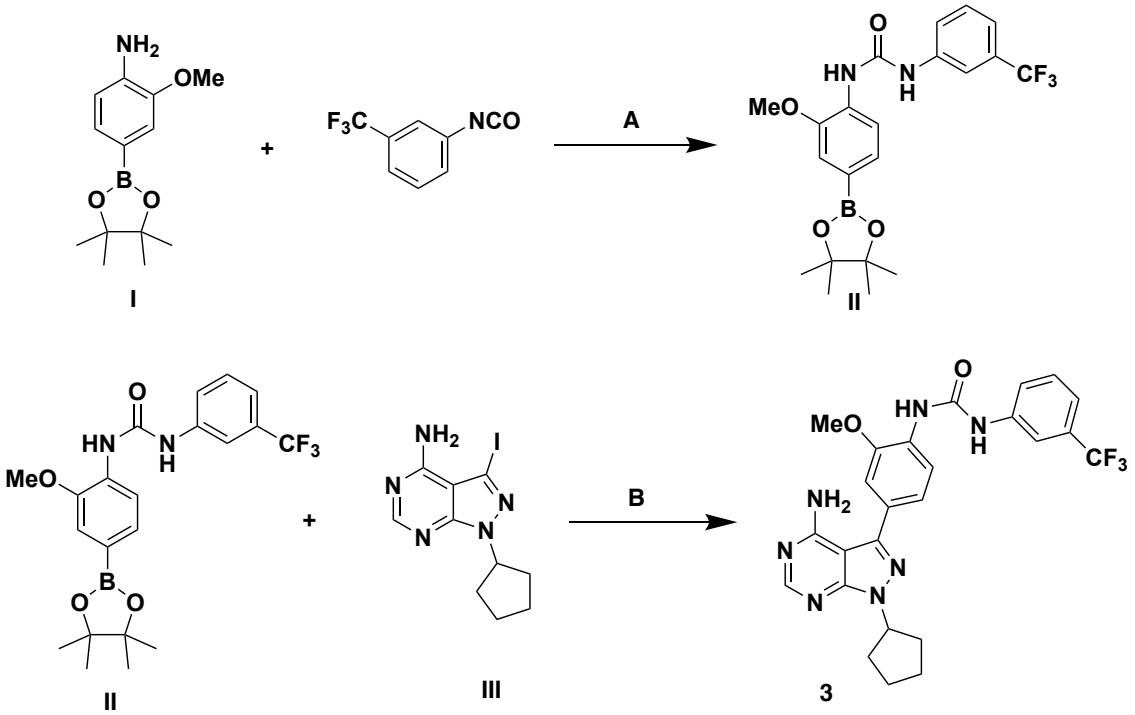
To determine extent of C-tail phosphorylation of SFK^{3D} by Csk (T338I), 400 nM Hck^{3D} or Lyn^{3D} were incubated with 3-5 μM of an inhibitor in kinase reaction buffer. Following an 1 h incubation at room temperature, 40 nM Csk was added and phosphorylation was initiated by the addition of 0.08 μCi/well γ^{32}P-ATP at a final volume of 30 μL. The enzymatic reaction was run at room temperature for 1 h and then terminated by spotting 4.6 μL of the reaction mixture onto a nitrocellulose membrane. Membranes were washed with 0.05% phosphoric acid (3x, 10 min wash) and air-dried, and the radioactivity was determined by phospho-imaging with a GE Typhoon FLA 9000 phosphor scanner. The scanned membranes were quantified with ImageStudioLite, and data were analyzed using GraphPad Prism.

General Synthetic procedures:

All chemicals purchased from commercial suppliers were used without further purification unless otherwise stated. Reactions were monitored with thin-layer chromatography (TLC) using silica gel 60 F254 coated glass plates (EM Sciences). Compound purification was performed with an IntelliFlash 280 automated flash chromatography system using pre-packed Varian SuperFlash silica gel columns (Hexane/EtOAc or CH_2Cl_2/MeOH gradient solvent). A Varian Dynamax Microsorb 100-5 C18 column (250 mm x 21.4 mm), eluting with H_2O/CH_3CN or H_2O/ MeOH gradient solvent (+0.05% TFA), was used for preparatory HPLC purification. The purity of all final compounds was determined by analytical HPLC with an Agilent ZORBAX SB-C18 (2.1 mm x 150 mm) or Varian Microsorb-MV 100-5 C18 column (4.6 mm x 150 mm), eluting with either H_2O/CH_3CN or H_2O/MeOH gradient solvent (+0.05% TFA). Elution was monitored by a UV detector at λ= 220 nm and λ= 254 nm, with

all final compounds displaying > 95% purity. Nuclear magnetic resonance (NMR) spectra were recorded on Bruker 300 or 500 MHz NMR spectrometers at ambient temperature. Chemical shifts were reported in parts per million (ppm) and coupling constants in hertz (Hz). 1H-NMR spectra were referenced to the residual solvent peaks as internal standards (7.26 ppm for $CDCl_3$, 2.50 ppm for d6-DMSO, and 3.34 ppm for CD_3OD). Mass spectra were recorded with a Bruker Esquire Liquid Chromatograph - Ion Trap Mass Spectrometer. Compounds **1** and **2** were synthesized according to previously reported protocol.[42]

Scheme 1:



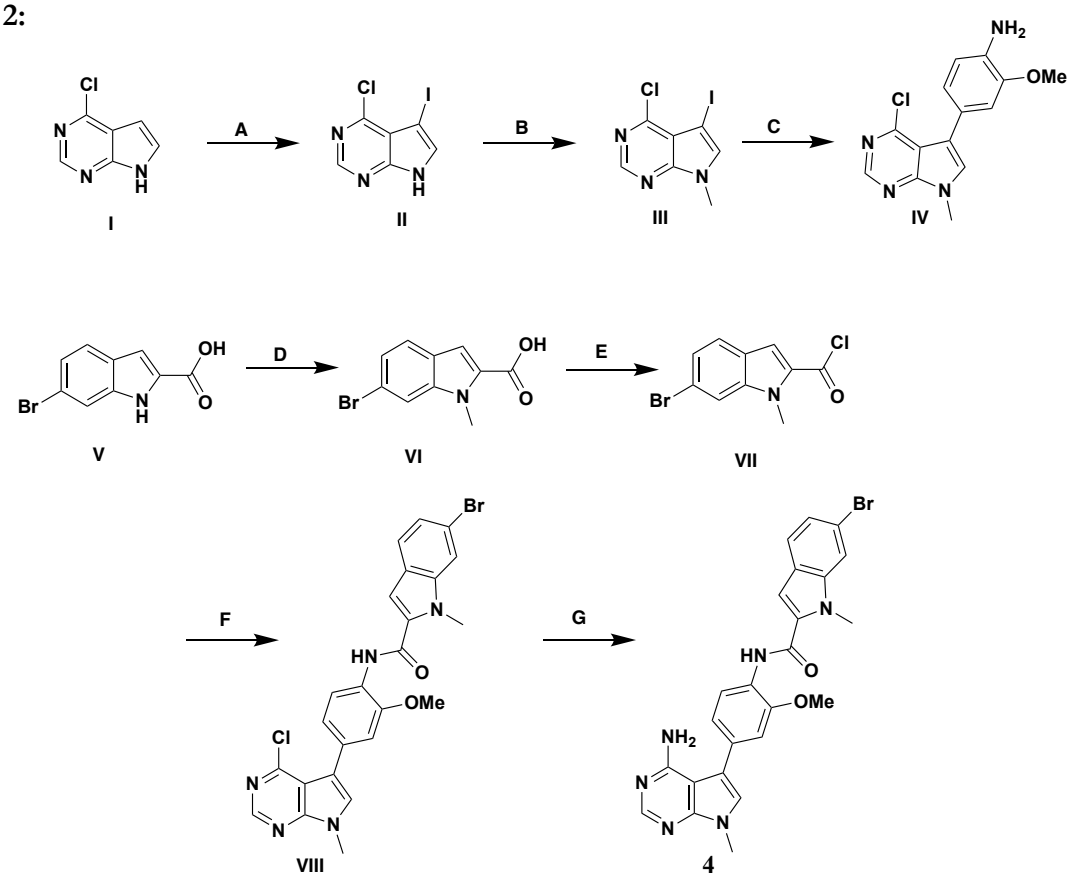
Scheme 1: *A) Tetrahydrofuran, 0 °C, 1 h; B) $PdCl_2$(dppf) DCM complex. K_3PO_4, Dioxane-H_2O 110 °C (microwave). Synthesis of* ***III*** *was done according to previous protocol.*[42]

A solution of I (100 mg, 0.40 mmol, 1.0 equiv.) in THF (7 mL) was cooled to 0 °C. To it was added a solution of (68 mg, 0.36 mmol, 0.9 equiv.) in THF (3 mL) dropwise. The reaction solution was stirred at 0°C for 1 h and then quenched with water (1 mL). The resulting mixture was diluted with ethyl

acetate (50 mL) and the organic phase was washed with 1 M HCl (5 mL), aqueous sodium bicarbonate solution (5 mL), brine (5 mL), dried over anhydrous sodium sulfate and concentrate in *vacuo*. Purification by flash chromatography on silica gel afforded **II** as a yellow solid (142 mg, 81%). ^{1}H-NMR (300 MHz, CDCl3) δ= 8.14 (d, J = 8.0 Hz, 1H), 7.70 (s, 1H), 7.60 (d, J = 7.8 Hz, 1H), 7.51 – 7.38 (m, 2H), 7.37 – 7.28 (m, 3H), 6.77 (s, 1H), 3.91 (s, 3H), 1.35 (s, 12H); MS (ESI, m/z) calculated for $C_{21}H_{24}BF_3N_2O_4$ 436.2, $[M+H]^+$.

A solution of **III** (15 mg, 0.046 mmol, 1.0 equiv.), **II** (30 mg, 0.068 mmol, equiv.), potassium phosphate (24 mg, 0.11 mmol, 2.5 equiv.) and bis(triphenylphosphine palladium (II) dichloride (3.7 mg, 0.0046 mmol, 0.10 equiv.) in 1,4-Dioxane (0.8 mL) and water (0.2 mL) was heated in a microwave reactor at 80 °C for 1 h. The reaction mixture was then diluted with ethyl acetate (50 mL) and washed with aqueous ammonium chloride solution (5 mL), sodium bicarbonate solution (5 mL), brine (5 mL), dried over anhydrous sodium sulfate and concentrate in *vacuo*. Purification by flash chromatography on silica gel rendered **3** as a brown solid (22 mg, 94%) ^{1}H-NMR (300 MHz, DMSO) δ= 9.75 (s, 1H), 8.47 (s, 1H), 8.32 (d, J = 8.3 Hz, 1H), 8.23 (s, 1H), 8.06 (s, 1H), 7.63 – 7.51 (m, 2H), 7.38 – 7.19 (m, 3H), 5.31 – 5.16 (m, 1H), 3.98 (s, 3H), 2.19 – 2.01 (m, 4H), 1.99 – 1.85 (m, 2H), 1.79 – 1.64 (m, 2H); MS (ESI, m/z) calculated for $C_{25}H_{24}F_3N_7O_2$ 511.2, $[M+H]^+$.

Scheme 2:



Scheme 2: *A) N-iodosuccinimide, Dichlomethane (DCM); B) Methyl iodide, Cs_2CO_3, Dimethyl formamide (DMF); C), 2-methoxy-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)aniline Pd(PPh_3)$_4$, K_3PO_4, Dioxane-H_2O, 80 °C; D) Methyl iodide, Sodium hydride, DMF; E) Oxalyl chloride, DMF, DCM; F)* ***IV****, N,N-Diisopropylethylamine (DIPEA), THF; G) aqueous ammonia, Dioxane 130 °C(microwave).*

To a solution of 4-chloro pyrrolopyrimidine (**I**) (5.0 g, 32.6 mmol) in dichloromethane(DCM) (125 mL), N-iodosuccinimide (8.8 g, 30.1 mmol) were added, and the mixture was stirred at room temperature for 12 h. Obtained precipitate was filtered and washed with water (200 mL), methanol (30 mL) and dichloromethane (30 mL), respectively, and dried to obtain 1(H)-3-chloro-4-iodo pyrrolopyrimidine (intermediate **II**, 8.6 g, yield: 95%) as a solid. TLC *Rf*: 0.73 (ethyl acetate), 1H NMR (300 MHz, $CDCl_3$) δ 9.63 (s, 1H), 8.68 (s, 1H), 7.54 (d, J = 2.4 Hz, 1H).

Compound **II** (5.0 g, 17.9 mmol) was dissolved in N,N-dimethylformamide(DMF) (15 mL). To the solution, Cs_2CO_3 (8.7 g, 26.8 mmol) was added, and cooled in ice bath. Methyl iodide (5.1 g, 35.8

mmol) was then added to this and mixture was stirred at ambient temperature for 48 h. Obtained precipitate was filtered, washed with water and vacuum dried to obtain 1-methyl-3-chloro-4-iodo pyrrolopyrimidine (intermediate **III,** 4.1 g, yield: 78%) as a solid. TLC *Rf*: 0.82 (n-hexane/ethyl acetate, 1/2), 1H NMR (300 MHz, $CDCl_3$) δ 8.66 (s, 1H), 7.39 (s, 1H), 3.91 (s, 3H).

Compound **III** (1.6 g, 5.5 mmol) and 2-methoxy-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)aniline (1.6 g, 6.5 mmol) were dissolved in 1,4-dioxane (50 mL). To the solution, K_3PO_4 (3.5g,16.4mmol),water(15mL)and trakis(triphenylphosphine)palladium(0) (0.32 g, 0.27 mmol) were added, and then stirred at 80 °C for 3 h under a nitrogen atmosphere. Water was added and followed by extraction with ethyl acetate. The organic layer was dried over anhydrous $MgSO_4$, filtered, concentrated and vacuum dried. The solid residue obtained was purified by flash chromatography using hexane/ethyl acetate gradient to obtain 4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-aniline (intermediate **IV**, 0.95 g, yield: 60%) as a solid. TLC *Rf*: 0.44 (n-hexane/ethyl acetate, 2/3), 1H NMR (300 MHz, $CDCl_3$) δ 8.65 (s, 1H), 7.17 (s, 1H), 6.97 (d, J = 1.8 Hz, 1H), 6.92 (dd, J = 7.9, 1.8 Hz, 1H), 6.76 (d, J =7.9Hz,1H), 3.90(m, 8H).

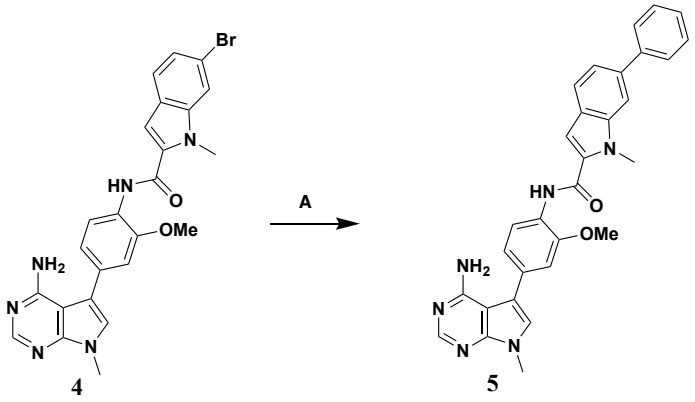
6-Bromoindole-2-carboxylic acid (**V**) (4.4 g, 18.3 mmol) was dissolved in DMF (100 mL). The solution was cooled in ice bath. Sodium hydride (60 wt/% in oil) (2.9 g, 73.2 mmol) was added to this solution and stirred at 5 min. Methyl iodide (10.7 g, 54.9 mmol) was then added and mixture was stirred at ambient temperature for 24 h. After the reaction was quenched by adding water, the precipitate was removed by filtration. 1 mol/L hydrochloric acid was added to the filtrate obtained. The precipitate was filtered and washed with mixture of n-hexane and ethyl acetate and vacuum dried to obtain 6-Bromo-1-methyl-1H-indole-2-carboxylic acid (intermediate **VI,** 1.9 g, yield: 41%) as a solid. TLC *Rf*: 0.30 (ethyl acetate), 1H NMR (300 MHz, DMSO) δ 13.04 (s, 1H), 7.85 (s, 1H), 7.61 (d, J = 8.5 Hz, 1H), 7.23 (d, J = 8.5 Hz, 1H), 7.21(s, 1H), 3.99 (s, 3H).

To a solution of Compound **VI** (440 mg, 1.75 mmol) in DCM (20 mL), oxalyl chloride (160 µL, 2.10 mmol) was added and then a few drops of DMF by using pasteur pipette. The reaction mixture was stirred at ambient temperature for 30 min and then concentrated under reduced pressure. The residue obtained was conducted azeotropic distillation with toluene twice and vacuum dried to obtain the crude product containing 6-bromo-1-methyl-indole-2-carbonyl chloride (intermediate **VII**).

To a solution of Compound **IV** (33 mg, 0.12 mmol) in THF (4 mL), the crude product containing Compound **VII** (38 mg, ~0.14 mmol) and N,N-diisopropylethylamine (DIPEA) (40 υΛ, 0.23 mmol) were added and stirred at ambient temperature for 12 h. Water was added and followed by extraction with mixture of ethyl acetate and THF. The organic layer was collected, dried over anhydrous Na_2SO_4, filtered and then concentrated. The solid residue was washed with mixture of n-hexane and ethyl acetate and vacuum dried to obtain 6-bromo-N-[4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-indole-2-carboxamide (intermediate **VIII**, 50 mg, yield: 83%) as a solid. TLC *Rf*: 0.50 (DCM/ethyl acetate, 2/3), 1H NMR (300 MHz, $CDCl_3$) δ 8.69 (s, 1H), 8.61 (s, 1H), 8.51 (d, J = 8.7 Hz, 1H), 7.60 (s, 1H), 7.56 (d, J = 8.5 Hz, 1H), 7.32 – 7.27 (m, 2H), 7.19 – 7.12 (m, 2H), 7.01 (s, 1H), 4.09 (s, 3H), 4.00 (s, 3H), 3.95 (s, 3H).

In a sealed tube, Compound **VIII** (42 mg, 0.079 mmol) was dissolved in 1,4-dioxane (1 mL). Saturated aqueous ammonia (1 mL) was added to this and microwaved at 130 °C for 8 h. The reaction mixture was concentrated and the residue obtained was purified by flash chromatography using hexane/ethyl acetate and DCM/methanol gradient to obtain N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-6-bromo-1-methyl-indole-2-carboxamid (Compound **4**, 29 mg, yield: 74%) as a solid. TLC *Rf*: 0.46 (ethyl acetate/methanol, 2/3), 1H NMR (300 MHz, $CDCl_3$) δ 8.58 (s, 1H), 8.53 (d, J = 8.2 Hz, 1H), 8.37 (s, 1H), 7.60 (s, 1H), 7.56 (d, J= 8.5 Hz, 1H), 7.29 (dd, J = 8.5, 1.6 Hz, 1H), 7.14 (dd, J = 8.2, 1.7 Hz, 1H), 7.05 (d, J = 1.7 Hz, 1H), 7.02 (s, 1H), 6.97 (s, 1H), 5.16 (s, 2H), 4.09 (s, 3H), 3.99 (s, 3H), 3.86 (s, 3H). MS m/z: 506.4 $[(M+H)^+]$.

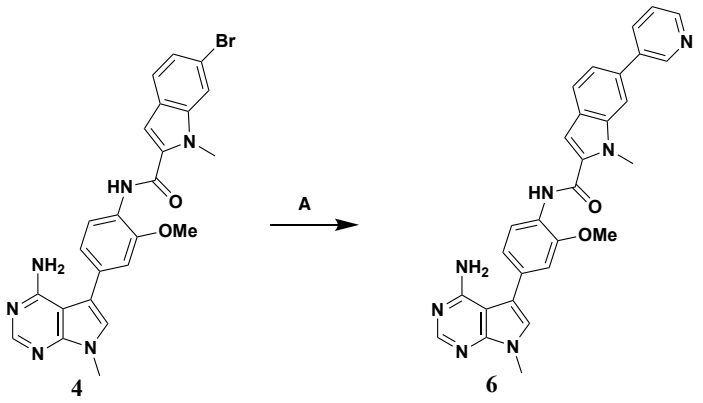
Scheme 3:



Scheme 3: *A) Phenylboronic acid, $PdCl_2(dppf)$ DCM complex. K_3PO_4, Dioxane-H_2O 110 °C (microwave).*

In a sealed tube, Compound **4** (15 mg, 0.030 mmol) and phenylboronic acid (5.5 mg, 0.045 mmol) were dissolved in 1,4-dioxane (2 mL). To the solution, K_3PO_4 (19 mg, 0.090 mmol), water (0.5 mL) and $PdCl_2(dppf)_2$ DCM complex (146 mg, 0.18 mmol) were added and then microwaved at 110 °C for 3 h under a nitrogen atmosphere. The reaction mixture was concentrated under reduced pressure. The residue was purified by flash chromatography using hexane/ethyl acetate and DCM/methanol gradient to obtain N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-6-phenyl-indole-2-carboxamide (Compound **5**, 11 mg, yield: 70%) as a solid. TLC *Rf*: 0.39 ($CHCl_3$/methanol, 9/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.62 (s, 1H), 8.56 (d, J = 8.2 Hz, 1H), 8.37 (s, 1H), 7.76 (d, J = 8.0 Hz, 1H), 7.71 (d, J = 1.4 Hz, 1H), 7.69 (s, 1H), 7.61 (s, 1H), 7.53 – 7.43 (m, 3H), 7.37 (t, J = 7.3 Hz, 1H), 7.15 (dd, J = 8.2, 1.8 Hz, 1H), 7.08 (s, 1H), 7.05 (d, J = 1.8 Hz, 1H), 6.98 (s, 1H), 5.20 (s, 2H), 4.18 (s, 3H), 4.01 (s, 3H), 3.86 (s, 3H). MS m/z: 502.6 [$(M+H)^+$].

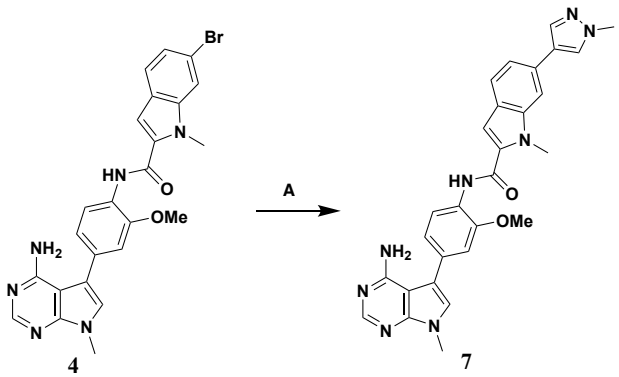
Scheme 4:



Scheme 4: *A) Pyridine-3-boronic acid, $PdCl_2$(dppf) DCM complex. K_3PO_4, Dioxane-H_2O 110 °C (microwave).*

N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-6-(3-pyridyl)indole-2-carboxamide (Compound **6**) was obtained as a solid by using pyridine-3-boronic acid instead of phenylboronic acid and in the procedure for Compound **5** and carrying out the subsequent reaction by the same operation. TLC *Rf*: 0.39 (ethyl acetate/methanol, 9/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.96 (s, 1H), 8.63 (s, 2H), 8.57 (d, J = 8.2 Hz, 1H), 8.37 (s, 1H), 7.98 (d, J= 7.8 Hz, 1H), 7.80 (d, J = 8.2 Hz, 1H), 7.61 (s, 1H), 7.48 – 7.35 (m, 2H), 7.15 (dd, J = 8.2, 1.7 Hz, 1H), 7.10 (s, 1H), 7.05 (d, J = 1.7 Hz, 1H), 6.99 (s, 1H), 5.28 (s, 2H), 4.19 (s, 3H), 4.01 (s, 3H), 3.87 (s, 3H). MS m/z: 503.6 $[(M+H)^+]$.

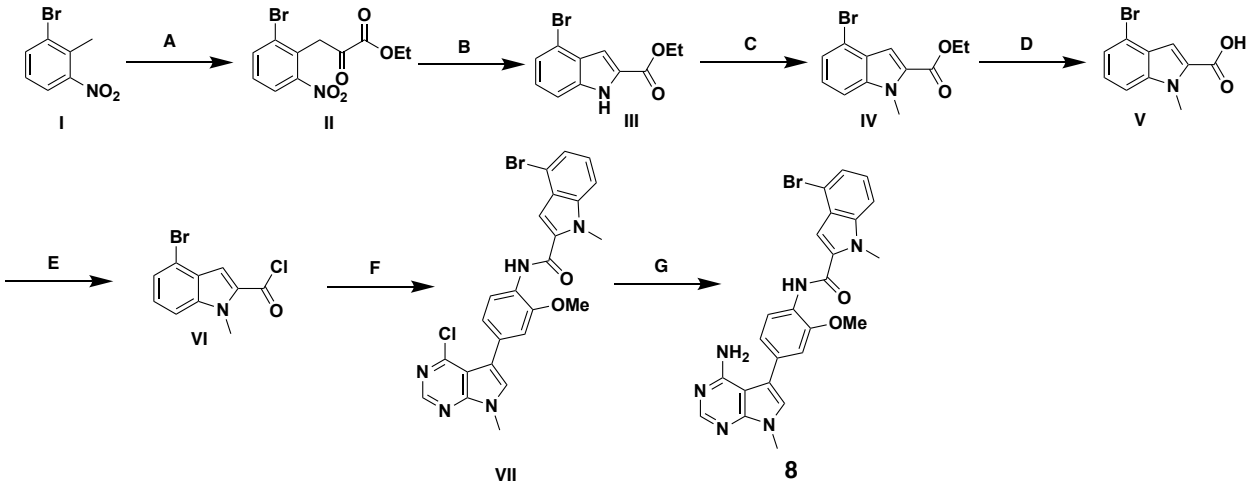
Scheme 5:



Scheme 5: *A) 1-Methyl-4-(4,4,5,5-tetrametyl-1,3,2-dioxaborolan-2-yl)-1H-pyrazole, $PdCl_2$(dppf) DCM complex. K_3PO_4, Dioxane-H_2O 110 °C (microwave).*

N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-6-(1-methylpyrazol-4-yl)indole-2-carboxamide (Compound **7**) was obtained as a solid by using 1-methyl-4-(4,4,5,5-tetrametyl-1,3,2-dioxaborolan-2-yl)-1H-pyrazole instead of phenylboronic acid and in the procedure for Compound **5** and carrying out the subsequent reaction by the same operation. TLC *Rf*: 0.41 (ethyl acetate/methanol, 9/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.60 (s, 1H), 8.55 (d, J = 8.2 Hz, 1H), 8.37 (s, 1H), 7.85 (s, 1H), 7.70 (s, 1H), 7.68 (d, J = 8.3 Hz, 1H), 7.49 (s, 1H), 7.33 (d, J = 8.3 Hz, 1H), 7.14 (d, J = 8.2 Hz, 1H), 7.04 (s, 2H), 6.98 (s, 1H), 5.17 (s, 2H), 4.15 (s, 3H), 4.00 (s, 3H), 3.98 (s, 3H), 3.86 (s, 3H). MS m/z: 506.6 [$(M+H)^+$].

Scheme 6:



Scheme 6: *A) Diethyl oxalate, Sodium hydride, THF, B) Zinc, Ammonium Chloride, THF- MeOH-H_2O; C) Methyl Iodide, Cs_2CO_3, DMF; D) NaOHaq, MeOH-THF; E) Oxalyl chloride, DMF, DCM; F)* ***VII*** *in Scheme 1, N,N-Diisopropylethylamine (DIPEA), THF; G) aqueous ammonia, Dioxane 130 °C(microwave).*

To a solution of 2-bromo-6-nitrotoluene (**I**) (3.1 g, 14.4 mmol) in THF (100 mL), Sodium hydride (60wt% in oil) (0.69 g, 17.3 mmol) was added and stirred at 5 min. Diethyl oxalate (2.5 g, 17.3 mmol) was then added and mixture was stirred at ambient temperature for 24 h. 1 mol/L hydrochloric acid was added, followed by extraction with ethyl acetate. The organic layer was collected and washed again with brine and dried over anhydrous Na_2SO_4, filtered, concentrated and vacuum dried. The residue

was purified by flash chromatography using hexane/ethyl acetate gradient to obtain ethyl 3-(2-bromo-6-nitro-phenyl)-2-oxo-propanoate (intermediate **II**, 1.1 g, yield: 24%) as an oil. TLC *Rf*: 0.19 (n-hexane/ethyl acetate, 9/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.02 (dd, J = 8.2, 1.1 Hz, 1H), 7.91 (dd, J = 8.1, 1.1 Hz, 1H), 7.37 (t, J = 8.1 Hz, 1H), 4.74 (s, 2H), 4.41 (q, J = 7.1 Hz, 2H), 1.42 (t, J = 7.1 Hz, 3H).

Compound **II** (1.1 g, 3.32 mmol) was dissolved in mixture of THF (50 mL) and methanol (50 mL). Water (20 mL), Zinc (0.65 g, 10.0 mmol) and ammonium chloride (1.1 g, 19.9 mmol) were added to this solution. The mixture was stirred at ambient temperature for 12 h. The precipitate was removed by filtration with diatomaceous earth. The filtrate obtained was concentrated and water were added, followed by extraction with ethyl acetate. The organic layer was dried over anhydrous $MgSO_4$, filtered and concentrated. The residue was washed with mixture of n-hexane and ethyl acetate and vacuum dried to obtain ethyl 4-bromo-1H-indole-2-carboxylate (intermediate **III**, 0.74 g, yield: 83%) as a solid. TLC *Rf*: 0.29 (n-hexane/ethyl acetate, 5/1), 1H NMR (300 MHz, $CDCl_3$) δ 9.01 (s, 1H), 7.38 – 7.30 (m, 2H), 7.24 (s, 1H), 7.17 (t, J = 7.9 Hz, 1H), 4.43 (q, J = 7.1 Hz, 2H), 1.43 (t, J = 7.1 Hz, 3H).

Compound **III** (0.64 g, 2.39 mmol) was dissolved in N,N-dimethylformamide(DMF) (10 mL). To the solution, Cs_2CO_3 (1.2 g, 3.56 mmol) and methyl iodide (0.68 g, 4.77 mmol) was added and the mixture was stirred at ambient temperature for 1.5 h. Water were added, followed by extraction with ethyl acetate. The organic layer was dried over anhydrous $MgSO_4$, filtered and concentrated. The residue was purified by flash chromatography using hexane/ethyl acetate gradient to obtain ethyl 4-bromo-1-methyl-indole-2-carboxylate (Compound **IV**, 0.58 g, yield: 87%) as a solid. TLC *Rf*: 0.57 (n-hexane/ethyl acetate, 5/1), 1H NMR (300 MHz, $CDCl_3$) δ 7.35 – 7.31 (m, 3H), 7.19 (t, J = 9.0 Hz, 1H), 4.39 (q, J = 7.1 Hz, 2H), 4.08 (s, 3H), 1.43 (t, J = 7.1 Hz, 3H).

To a solution of Compound **IV** (480 mg, 1.71 mmol) in mixture of THF (8 mL) and methanol (8 mL), 1mol/L Sodium hydroxide aqueous solution (3.4 mL, 3.41 mmoL) was added to this and stirred at

ambient temperature for 12 h. After 1 mol/L hydrochloric acid was added, the obtained precipitate was filtered and dissolved in ethyl acetate. The solution was dried over anhydrous $MgSO_4$, filtered, concentrated and vacuum dried to obtain 4-bromo-1-methyl-indole-2-carboxylic acid (intermediate **V**, 420 mg, 1.64 mmol, yield: 96%) as a solid. TLC *Rf*: 0.14 (n-hexane/ethyl acetate, 3/1), 1H NMR (300 MHz, $CDCl_3$) δ 7.46 (s, 1H), 7.36 – 7.20 (m, 2H), 7.23 (s, 1H), 4.09(s, 3H).

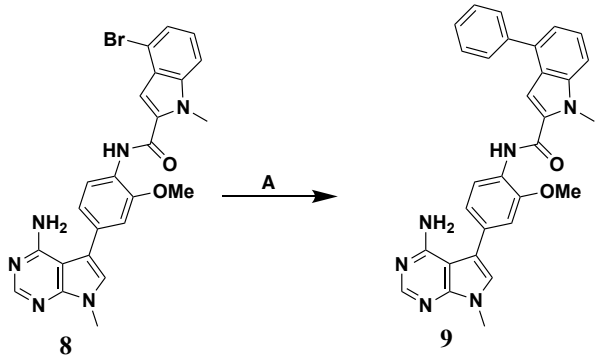
To a suspension of Compound **V** (260 mg, 1.02 mmol) in DCM (5 mL), oxalyl chloride (96 μL, 1.22 mmol) was added and then a drop of DMF by using pasteur pipette. The reaction mixture was stirred at ambient temperature for 30 min. The reaction mixture became clear solution. This was concentrated under reduced pressure and the residue obtained was conducted azeotropic distillation with toluene twice and vacuum dried to obtain the crude product of 4-bromo-1-methyl-indole-2-carbonyl chloride (intermediate **VI**). 1H NMR (300 MHz, $CDCl_3$) δ 7.70 (s, 1H), 7.39 – 7.29 (m, 3H), 4.01 (s, 3H).

To a solution of 4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-aniline (Compound **IV** in scheme 1) (32 mg, 0.11 mmol) in THF (4 mL), the crude product containing Compound **VI** (36 mg, ~0.13 mmol) and N,N-diisopropylethylamine (DIPEA) (39 μL, 0.22 mmol) were added and stirred at ambient temperature for 24 h. The water was added, and the precipitate was filtered, washed with mixture of n-hexane and ethyl acetate and vacuum dried to obtain 4-bromo-N-[4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-indole-2-carboxamide (intermediate **VII**, 54 mg, yield: 93%) as a solid. TLC *Rf*: 0.68 (n-hexane/ethyl acetate, 1/4), 1H NMR (300 MHz, $CDCl_3$) δ 8.69 (s, 1H), 8.65 (s, 1H), 8.51 (d, J = 8.8 Hz, 1H), 7.40 – 7.36 (m, 2H), 7.27 (s, 1H), 7.23 (d, J = 8.0 Hz, 1H), 7.20 – 7.12 (m, 2H), 7.07 (s, 1H), 4.12 (s, 3H), 4.02 (s, 3H), 3.96 (s, 3H).

In a sealed tube, Compound **VII** (42 mg, 0.080 mmol) was dissolved in 1,4-dioxane (1 mL). Saturated aqueous ammonia (1 mL) was added to this and microwaved at 130 °C for 8 h. The reaction mixture was concentrated and the residue obtained was purified by flash chromatography using hexane/ethyl acetate and DCM/methanol gradient to obtain N-[4-(4-amino-7-methyl-pyrrolo[2,3-d] pyrimidin-5-

yl)-2-methoxy-phenyl]-4-bromo-1-methyl-indole-2-carboxamide (Compound **8**, 23 mg, yield: 58%) as a solid. TLC *Rf*: 0.57 (ethyl acetate/methanol, 9/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.62 (s, 1H), 8.53 (d, J = 8.2 Hz, 1H), 8.36 (s, 1H), 7.40 – 7.36 (m, 2H), 7.26 – 7.18 (m, 1H), 7.14 (dd, J = 8.2, 1.7 Hz, 1H), 7.10 – 7.03 (m, 2H), 6.98 (s, 1H), 5.24 (s, 2H), 4.12 (s, 3H), 4.02 (s, 3H), 3.86 (s, 3H). MS m/z: 505.4 $[(M+H)^+]$.

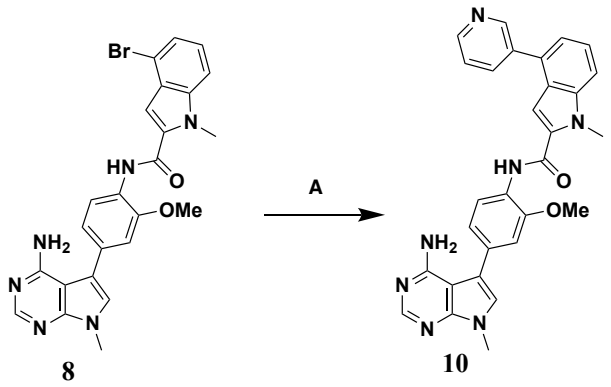
Scheme 7:



Scheme 7: *A) Phenylboronic acid, $PdCl_2(dppf)$ DCM complex. K_3PO_4, Dioxane-H_2O 110 °C (microwave).*

In a sealed tube, Compound **8** (16 mg, 0.032 mmol) and phenylboronic acid (5.9 mg, 0.048 mmol) were dissolved in 1,4-dioxane (2 mL). To the solution, K_3PO_4 (20 mg, 0.096 mmol), water (0.5 mL) and $PdCl_2(dppf)_2$ DCM complex (2.6 mg, 0.0032 mmol) were added and then microwaved at 110°C for 2 h under a nitrogen atmosphere. The reaction mixture was concentrated under reduced pressure. The residue was purified by flash chromatography using hexane/ethyl acetate and DCM/methanol gradient to obtain N-[4-(4-amino-7-methyl-pyrrolo [2,3-d] pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-4-phenyl-indole-2-carboxamide (Compound **9**, 12 mg, yield: 75%) as a solid. 1H NMR (300 MHz, DMSO) δ 11.96 (s, 1H), 9.64 (s, 1H), 8.18 (s, 1H), 7.77 – 7.72 (m, 2H), 7.62 – 7.38 (m, 7H), 7.24 (d, J = 7.2 Hz, 1H), 7.16 (s, 1H), 7.07 (d, J = 8.0 Hz, 1H), 6.16 (s, 2H), 4.07 (s, 3H), 3.88 (s, 3H), 3.76 (s, 3H). MS m/z: 505.4 $[(M+H)^+]$.

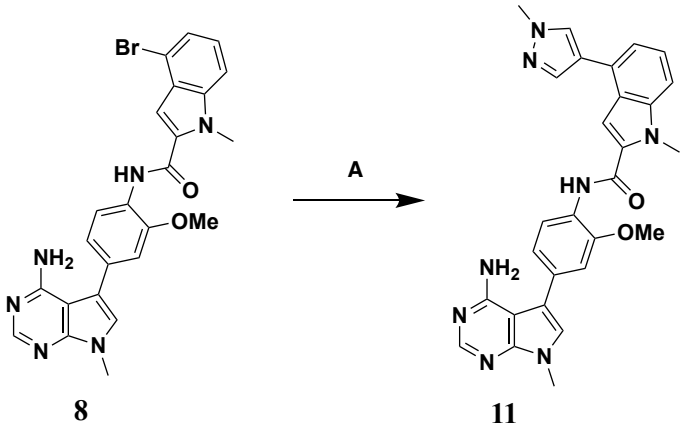
Scheme 8:



Scheme 8: *A) Pyridine-3-boronic acid, $PdCl_2(dppf)$ DCM complex. K_3PO_4, Dioxane-H_2O 110 °C (microwave).*

N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-4-(3-pyridyl)indole-2-carboxamide (Compound **10**) was obtained as a solid by using pyridine-3-boronic acid instead of phenylboronic acid and in the procedure for Compound **9** and carrying out the subsequent reaction by the same operation. TLC *Rf*: 0.20 (ethyl acetate/methanol, 19/1), 1H NMR (300 MHz, $CDCl_3$) δ 9.00 (d, J = 1.6 Hz, 1H), 8.71 – 8.63 (m, 1H), 8.55 – 8.46 (m, 2H), 8.33 (s, 1H), 8.00 (d, J = 7.8 Hz, 1H), 7.51 – 7.45 (m, 3H), 7.25 (d, J = 6.5 Hz, 1H), 7.17 – 7.08 (m, 2H), 7.02 (d, J = 1.7 Hz, 1H), 6.97 (s, 1H), 5.49 (s, 2H), 4.17 (s, 3H), 3.95 (s, 3H), 3.86 (s, 3H). MS m/z : 503.6 [(M+H)+].

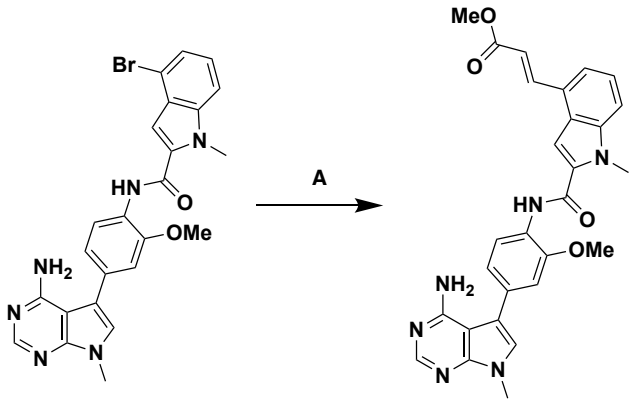
Scheme 9:



Scheme 9: *A) 1-Methyl-4-(4,4,5,5-tetrametyl-1,3,2-dioxaborolan-2-yl)-1H-pyrazole, $PdCl_2(dppf)$ DCM complex. K_3PO_4, Dioxane-H_2O 110 °C (microwave).*

N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-4-(1-methylpyrazol-4-yl)indole-2-carboxamide (Compound **11**) was obtained as a solid by using 1-methyl-4-(4,4,5,5-tetrametyl-1,3,2-dioxaborolan-2-yl)-1H-pyrazole instead of phenylboronic acid and in the procedure for Compound **9** and carrying out the subsequent reaction by the same operation. TLC *Rf*: 0.14 (ethyl acetate/methanol, 19/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.61 – 8.48 (m, 2H), 8.28 (s, 1H), 7.94 (s, 1H), 7.75 (s, 1H), 7.43 – 7.31 (m, 2H), 7.27 – 7.19 (m, 2H), 7.13 (dd, J = 8.2, 1.7 Hz, 1H), 7.02 (d, J = 1.7 Hz, 1H), 6.98 (s, 1H), 5.99 (s, 2H), 4.14 (s, 3H), 4.03 (s, 3H), 3.98 (s, 3H), 3.86 (s, 3H). MS m/z: 506.6 [$(M+H)^+$].

Scheme 10:



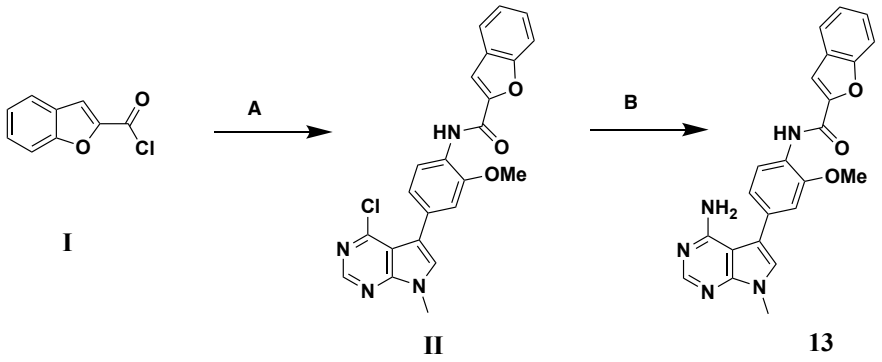
8 **12**

Scheme 10: *A) Methyl acrylate, $Pd(OAc)_2$, (ortho-MePh)$_3$P, Et_3N, DMF, 130 °C (microwave).*

In a sealed tube, Compound **8** (16 mg, 0.032 mmol) was dissolved in DMF (2 mL). To the solution, triethylamine (0.5 mL), methyl acrylate (8.3 mg, 0.097 mmol), $Pd(OAc)_2$ (0.7 mg, 0.0032 mmol) and tri-o-tolylphosphine (1.0 mg, 0.032 mmol) were added and the mixture was microwaved at 130 °C for 2 h under a nitrogen atmosphere. Water was added, followed by extraction with ethyl acetate. The organic layer was collected, dried over anhydrous Na_2SO_4, filtered and concentrated. The residue was purified by flash chromatography using hexane/ethyl acetate and DCM/methanol gradient to obtain methyl (E)-3-[2-[[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]carbamoyl]-

1-methyl-indol-4-yl]prop-2-enoate (Compound **12**, 4.5 mg, yield: 28%) as a solid. TLC *Rf*: 0.48 (ethyl acetate/methanol, 19/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.63 (s, 1H), 8.57 (d, J = 8.2 Hz, 1H), 8.29 (s, 1H), 8.16 (d, J = 16.0 Hz, 1H), 7.50 – 7.47 (m, 2H), 7.39 (dd, J = 8.9, 6.6 Hz, 1H), 7.29 (s, 1H), 7.12 (dd, J = 8.3, 1.8 Hz, 1H), 7.07 – 7.00 (m, 2H), 6.66 (d, J = 16.0 Hz, 1H), 6.30 (s, 2H), 4.14 (s, 3H), 4.03 (s, 3H), 3.89 (s, 3H), 3.86 (s, 3H). MS m/z: 510.6 [$(M+H)^+$].

Scheme 11:

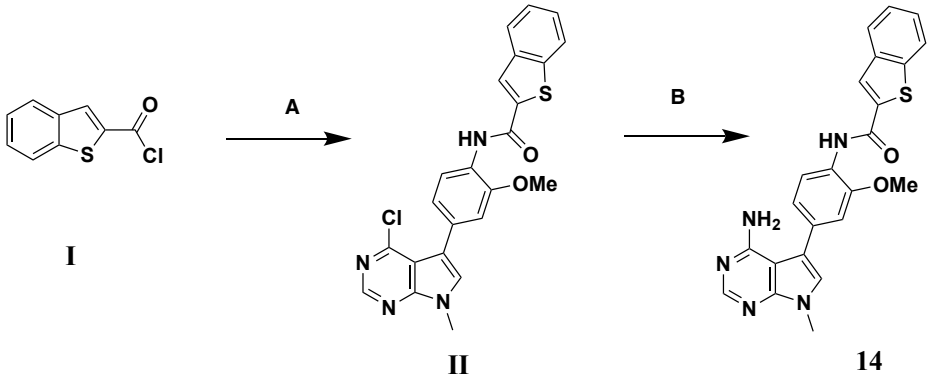


Scheme 11: *A)* ***VII*** *in Scheme 1, N,N-Diisopropylethylamine (DIPEA), THF; B) Aqueous ammonia, Dioxane 130 °C(microwave).*

To a solution of 4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-aniline (Compound **IV** in scheme 1) (26 mg, 0.092 mmol) in THF (5 mL), benzofuran-2-carbonyl chloride (20 mg, 0.11 mmol) and N,N-diisopropylethylamine (DIPEA) (32 μL, 0.18 mmol) were added and stirred at ambient temperature for 2 h. The water was added and the precipitate was filtered, washed with mixture of n-hexane and ethyl acetate and vacuum dried to obtain N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]benzofuran-2-carboxamide (intermediate **II**, 32 mg, yield: 81%) as a solid. TLC *Rf*: 0.65 (DCM/ethyl acetate, 2/1), 1H NMR (300 MHz, $CDCl_3$) δ 9.08 (s, 1H), 8.69 (s, 1H), 8.60 (d, J = 8.7 Hz, 1H), 7.72 (d, J = 7.0 Hz, 1H), 7.63 – 7.61 (m, 2H), 7.46 (t, J = 7.0 Hz, 1H), 7.33 (t, J = 7.0 Hz, 1H), 7.27 (s, 1H), 7.17- 7.13 (m, 2H), 4.05 (s, 3H), 3.96 (s, 3H).

In a sealed tube, Compound **II** (22 mg, 0.051 mmol) was dissolved in 1,4-dioxane (1.5 mL). Saturated aqueous ammonia (1.5 mL) was added to this and microwaved at 130 °C for 5 h. The reaction mixture was concentrated and the residue obtained was purified by flash chromatography using hexane/ethyl acetate and DCM/methanol gradient to obtain N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-4-bromo-1-methyl-indole-2-carboxamide (Compound **13**, 10 mg, yield: 49%) as a solid. TLC *Rf*: 0.50 (ethyl acetate/methanol, 19/1), 1H NMR (300 MHz, $CDCl_3$) δ 9.05 (s, 1H), 8.62 (d, J = 8.2 Hz, 1H), 8.37 (s, 1H), 7.72 (d, J = 7.4 Hz, 1H), 7.64 – 7.60 (m, 2H), 7.47 (t, J = 7.4 Hz, 1H), 7.34 (t, J = 7.4 Hz, 1H), 7.14 (dd, J = 8.2, 1.7 Hz, 1H), 7.06 (d, J = 1.7 Hz, 1H), 6.98 (s, 1H), 5.15 (s, 2H), 4.05 (s, 3H), 3.86 (s, 3H). MS m/z: 413.3 [(M+H)+].

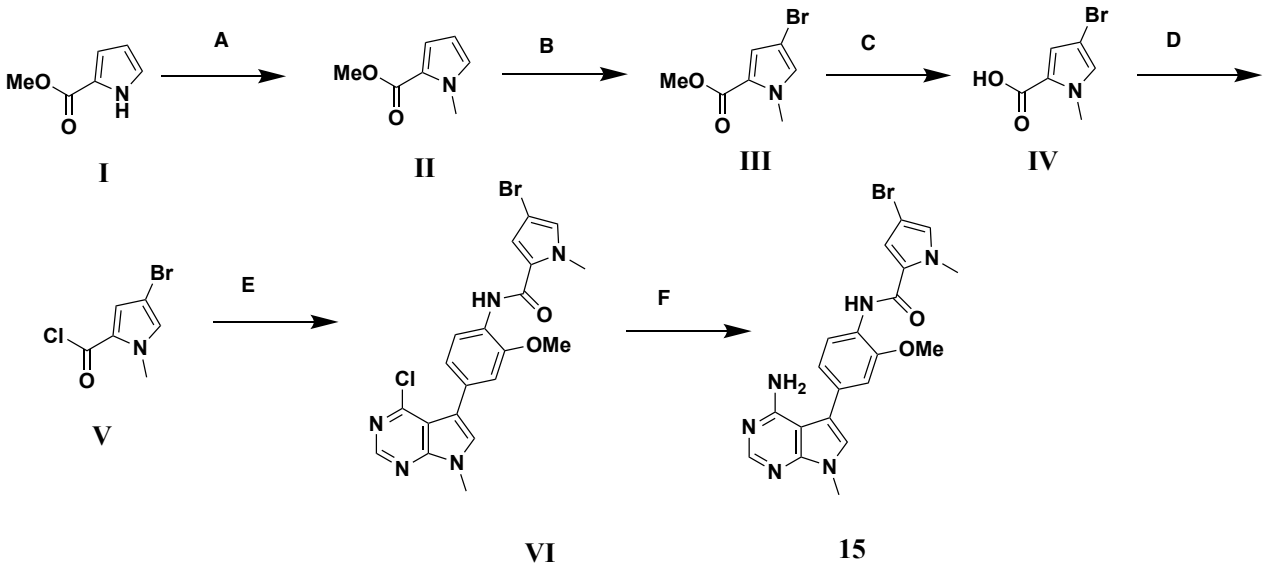
Scheme 12:



Scheme 12*: A)* ***VII*** *in Scheme 1, N,N-Diisopropylethylamine (DIPEA), THF; B) Aqueous ammonia, Dioxane 130 °C(microwave).*

N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]benzothiophene-2-carboxamide (Compound **14**) was obtained as a solid by using benzo[b] thiophene-2-carbonyl chloride instead of benzofuran-2-carbonyl chloride and in the procedure for Compound **13** and carrying out the subsequent reaction by the same operation. TLC *Rf*: 0.39 (ethyl acetate/methanol, 19/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.59 – 8.37 (m, 2H), 8.37 (s, 1H), 7.92 – 7.89 (m, 3H), 7.50 – 7.43 (m, 2H), 7.14 (dd, J = 8.2, 1.7 Hz, 1H) 7.05 (d, 1.7 Hz, 1H), 6.97 (s, 1H), 5.18 (s, 2H), 4.02 (s, 3H), 3.86 (s, 3H). MS m/z: 429.5 $[(M+H)^+]$.

Scheme 13:



Scheme 13: *A) Methyl Iodide, Sodium hydride, DMF; B) N-Bromosuccinimide, DCM; C) NaOHaq, MeOH-THF; D) Oxalyl chloride, DMF, DCM; E)* ***VII*** *in Scheme 1, N,N-Diisopropylethylamine (DIPEA), THF; F) Aqueous ammonia, Dioxane 130 °C(microwave).*

Methyl 1H-pyrrole-2-carboxylate (**I**) (3.9 g, 30.8 mmol) was dissolved in DMF (70 mL). After the solution was cooled in ice bath, Sodium hydride (60wt% in oil) (1.5 g, 36.9 mmol) was added and stirred at 10 min. Methyl Iodide (7.2 g, 36.9 mmol) was then added and the mixture was stirred at ambient temperature for 30 min. Saturated ammonium chloride aqueous solution was added, followed by extraction with ethyl acetate. The organic layer was collected and washed again with brine and dried over anhydrous Na_2SO_4, filtered, concentrated and vacuum dried. The residue was purified by flash chromatography using hexane/ethyl acetate gradient to obtain methyl 1-methylpyrrole-2-carboxylate (intermediate **II,** 4.1 g, 29.5 mmol, yield: 96%) as a solid. TLC *Rf*: 0.68 (n-hexane/ethyl acetate, 2/1), 1H NMR (300 MHz, $CDCl_3$) δ 6.93 (dd, J = 4.0, 1.8 Hz, 1H), 6.78 (dd, J = 1,8, 2.5 Hz, 1H), 6.11 (dd, J = 4.0, 2.5 Hz, 1H), 3.93 (s, 3H), 3.81 (s, 3H).

To a solution of Compound **II** (1.6 g, 11.5 mmol) in DCM (50 mL), N-bromosuccinimide (3.1 g, 17.3 mmol) was added and stirred at 40 °C for 1 hour. Water was added, followed by extraction with DCM.

The organic layer was collected and washed again with brine and dried over anhydrous $MgSO_4$, filtered, concentrated and vacuum dried. The residue was purified by flash chromatography using hexane/ethyl acetate gradient to obtain methyl 4-bromo-1-methyl-pyrrole-2-carboxylate (intermediate **III,** 0.83 g, 3.8 mmol, yield: 33%). TLC *Rf*: 0.42 (n-hexane/ethyl acetate, 9/1), 1H NMR (300 MHz, $CDCl_3$) δ 6.90 (d, J = 1.9 Hz, 1H), 6.77 (d, J = 1.9 Hz, 1H), 3.90 (s, 3H), 3.81 (s, 3H).

To a solution of Compound **III** (480 mg, 2.22 mmol) in mixture of THF (2 mL) and methanol (10 mL), 1mol/L Sodium hydroxide aqueous solution (4.5 mL, 4.43 mmoL) was added to this and stirred at ambient temperature for 96 h. After 1 mol/L hydrochloric acid was added, followed by extraction with DCM. The organic layer was collected and washed again with brine and dried over anhydrous $MgSO_4$, filtered, concentrated and vacuum dried to obtain 4-bromo-1-methyl-pyrrole-2-carboxylic acid (intermediate **IV**, 320 mg, 1.56 mmol, yield: 70%) as a solid. TLC *Rf*: 0.34 (n-hexane/ethyl acetate, 1/1), 1H NMR (300 MHz, $CDCl_3$) δ 7.04 (d, J = 1.9 Hz, 1H), 6.82 (d, J = 1.9 Hz, 1H), 3.90 (s, 3H).

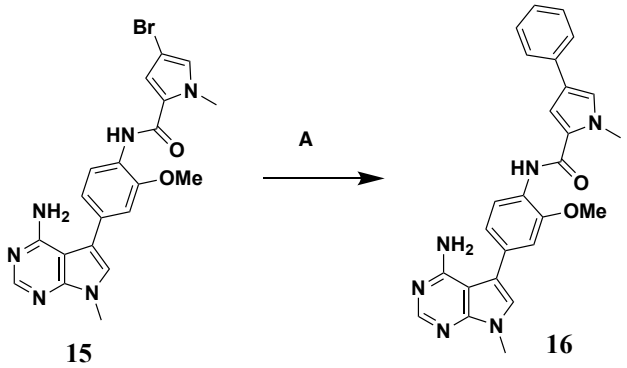
To a solution of Compound **IV** (160 mg, 0.81 mmol) in DCM (4 mL), oxalyl chloride (76 μL, 0.97 mmol) was added and then a drop of DMF by using pasteur pipette. The reaction mixture was stirred at ambient temperature for 20 min. The reaction mixture was concentrated under reduced pressure and the residue obtained was conducted azeotropic distillation with toluene twice and vacuum dried to obtain the crude product of 4-bromo-1-methyl-pyrrole-2-carbonyl chloride (intermediate **V**). 1H NMR (300 MHz, $CDCl_3$) δ 7.26 (d, J = 1.8 Hz, 1H), 6.96 (d, J = 1.8 Hz, 1H), 3.86 (s, 3H).

To a solution of 4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-aniline (Compound **IV** in scheme 1) (64 mg, 0.22 mmol) in THF (4 mL), Compound **V** (60 mg, 0.27 mmol) and N,N-diisopropylethylamine (DIPEA) (78 μL, 0.45 mmol) were added and stirred at ambient temperature for 6 h. The water was added and the precipitate was filtered, washed with mixture of n-hexane and ethyl acetate and vacuum dried to obtain 4-bromo-N-[4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-pyrrole-2-carboxamide (intermediate **VI**, 88 mg, 0.19 mmol, yield:

83%) as a solid. TLC *Rf*: 0.47 (n-hexane/ethyl acetate, 1/2), 1H NMR (300 MHz, $CDCl_3$) δ 8.68 (s, 1H), 8.41 (d, J = 8.8 Hz, 1H), 8.26 (s, 1H), 7.25 (s, 1H), 7.14 – 7.06 (m, 2H), 6.79 (d, J = 1.8 Hz, 1H), 6.72 (d, J = 1.8 Hz, 1H), 3.97 (s, 6H), 3.94 (s, 3H).

In a sealed tube, Compound **VI** (75 mg, 0.16 mmol) was dissolved in 1,4-dioxane (2 mL). Saturated aqueous ammonia (1 mL) was added to this and microwaved at 130 °C for 6 h. The reaction mixture was concentrated and the residue obtained was purified by flash chromatography using hexane/ethyl acetate and DCM/methanol gradient to obtain N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-4-bromo-1-methyl-pyrrole-2-carboxamide (Compound **15**, 57 mg, 0.13 mmol, yield: 80%) as a solid. TLC *Rf*: 0.43 (ethyl acetate/methanol, 19/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.44 (d, J = 8.2 Hz, 1H), 8.36 (s, 1H), 8.23 (s, 1H), 7.10 (dd, J = 8.2, 1.8 Hz, 1H), 7.01 (d, J = 1.8 Hz, 1H), 6.95 (s, 1H), 6.79 (d, J = 1.8 Hz, 1H), 6.72 (d, J = 1.8 Hz, 1H), 5.13 (s, 2H), 3.97 (s, 6H), 3.85 (s, 3H). MS m/z: 455.3 [(M+H)+].

Scheme 14:

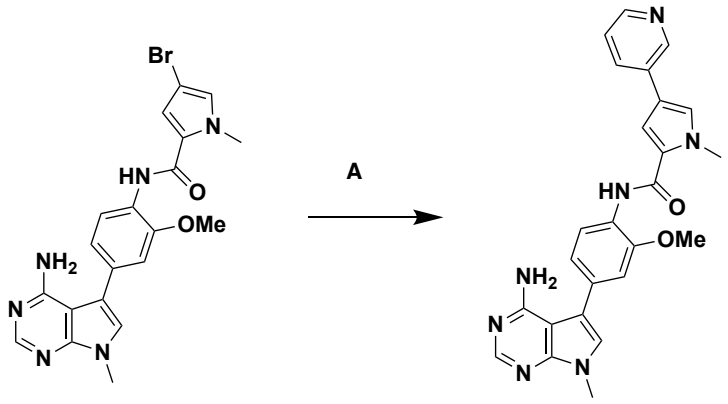


Scheme 14: *A) Phenylboronic acid, $PdCl_2$(dppf) DCM complex. K_3PO_4, Dioxane-H_2O 110 °C (microwave).*

In a sealed tube, Compound **15** (16 mg, 0.036 mmol) and phenylboronic acid (6.5 mg, 0.054 mmol) were dissolved in 1,4-dioxane (2 mL). To the solution, K_3PO_4 (23 mg, 0.11 mmol), water (0.5 mL) and $PdCl_2(dppf)_2$ DCM complex (2.9 mg, 0.0036 mmol) were added and then microwaved at 110 °C for 2

h under a nitrogen atmosphere. The reaction mixture was concentrated under reduced pressure. The residue was purified by flash chromatography using hexane/ethyl acetate and DCM/methanol gradient to obtain N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-4-phenyl-pyrrole-2-carboxamide (Compound **16**, 5.9 mg, 0.013 mmol, yield: 36%) as a solid. TLC *Rf*: 0.30 (ethyl acetate/methanol, 19/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.49 (d, J = 8.2 Hz, 1H), 8.36 (s, 2H), 7.53 (d, J = 7.1 Hz, 2H), 7.38 (t, J = 7.5 Hz, 2H), 7.25 (d, J = 7.5 Hz, 1H), 7.11 (dd, J = 8.2 Hz, 1.8 Hz, 2H), 7.03 (d, 1.8 Hz, 1H), 6.99 (d, J =1.8 Hz, 1H), 6.96 (s, 1H), 5.12 (s, 2H), 4.05 (s, 3H), 3.99 (s, 3H), 3.85 (s, 3H). MS m/z: 452.5 $[(M+H)^+]$.

Scheme 15:

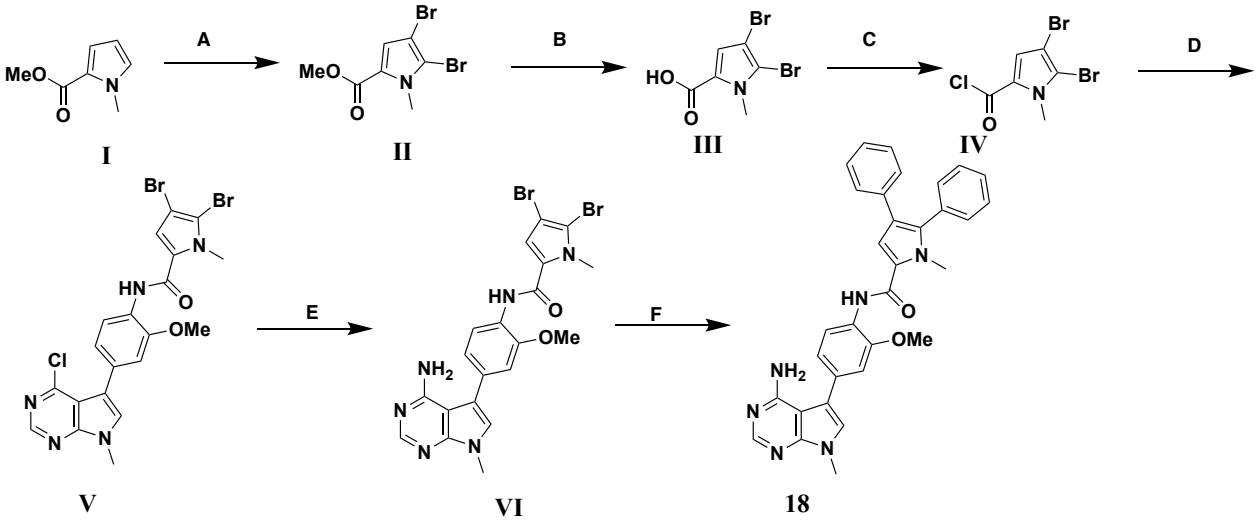


15 **17**

Scheme 15: *A) Pyridine-3-boronic acid, $PdCl_2(dppf)$ DCM complex. K_3PO_4, Dioxane-H_2O 110 °C (microwave).*

N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-4-(3-pyridyl)pyrrole-2-carboxamide (Compound **17**) was obtained as a solid by using 1- pyridine-3-boronic acid instead of phenylboronic acid and in the procedure for Compound **16** and carrying out the subsequent reaction by the same operation. TLC *Rf*: 0.10 (ethyl acetate/methanol, 19/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.81 (s, 1H), 8.50 - 8.47 (m, 2H), 8.35 (s, 2H), 7.79 (d, J = 8.0 Hz, 1H), 7.30 (dd, J = 8.0, 4.8 Hz, 1H), 7.17 – 7.08 (m, 2H), 7.03 (d, J = 1.8 Hz, 1H), 7.00 (d, J = 1.8 Hz, 1H), 6.97 (s, 1H), 5.27 (s, 2H), 4.06 (s, 3H), 4.00 (s, 3H), 3.86 (s, 3H). MS m/z: 453.5 $[(M+H)^+]$.

Scheme 16:



Scheme 16*: A) N-Bromosuccinimide, DCM; B) NaOHaq, MeOH-THF; C) Oxalyl chloride, DMF, DCM; D)* ***VII*** *in Scheme 1, N,N-Diisopropylethylamine (DIPEA), THF; E) Aqueous ammonia, Dioxane 130 °C(microwave); F) Phenylboronic acid, $PdCl_2(dppf)$ DCM complex. K_3PO_4, Dioxane-H_2O 110 °C (microwave).*

To a solution of Compound **I** (= Compound **II** in scheme 12) (2.1 g, 14.9 mmol) in chloroform (40 mL), N-bromosuccinimide (4.2 g, 29.8 mmol) was added and stirred at ambient temperature for 12 h. The reaction mixture was filtrated and the filtrate obtained was concentrated. The residue was purified by flash chromatography using hexane/ethyl acetate gradient to obtain methyl 4,5-dibromo-1-methyl-pyrrole-2-carboxylate (2.9 g, 9.8 mmol, yield: 65%) TLC *Rf*: 0.46 (n-hexane/ethyl acetate, 9/1), 1H NMR (300 MHz, $CDCl_3$) δ 7.01 (s, 1H), 3.97 (s, 3H), 3.82 (s, 3H).

To a solution of Compound **II** (1.6 g, 5.29 mmol) in mixture of THF (10 mL) and methanol (20 mL), 1mol/L Sodium hydroxide aqueous solution (10.6 mL, 10.6 mmoL) was added to this and stirred at ambient temperature for 96 h. After 1 mol/L hydrochloric acid was added and followed by extraction with ethyl acetate. The organic layer was collected and washed again with brine and dried over anhydrous $MgSO_4$, filtered, concentrated and vacuum dried to obtain 4,5-dibromo-1-methyl-pyrrole-2-carboxylic acid (intermediate **III**, 0.49 g, 1.73 mmol, yield: 33%) as a solid. TLC *Rf*: 0.32 (n-hexane/ethyl acetate, 1/1), 1H NMR (300 MHz, $CDCl_3$) δ 7.15 (s, 1H), 3.98 (s, 3H).

To a solution of Compound **III** (240 mg, 0.86 mmol) in DCM (5 mL), oxalyl chloride (81 μL, 1.00 mmol) was added and then a drop of DMF by using pasteur pipette. The reaction mixture was stirred at ambient temperature for 30 min. The reaction mixture was concentrated under reduced pressure and the residue obtained was conducted azeotropic distillation with toluene twice and vacuum dried to obtain the crude product of 4,5-dibromo-1-methyl-pyrrole-2-carbonyl chloride (intermediate **IV**). 1H NMR (300 MHz, CDCl3) δ 7.37 (s, 1H), 3.93 (s, 3H).

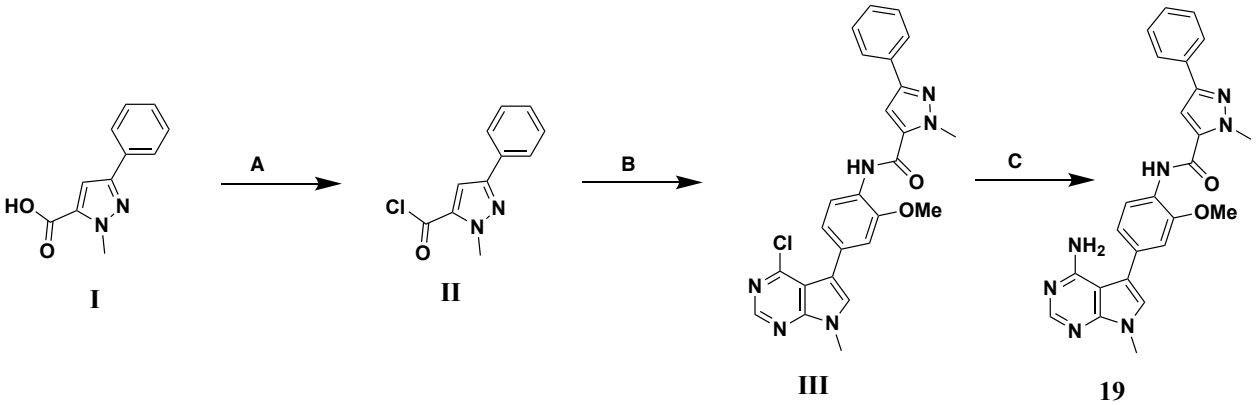
To a solution of 4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-aniline (Compound **IV** in scheme 1) (64 mg, 0.22 mmol) in THF (5 mL), Compound **IV** (80 mg, 0.26 mmol) and N,N-diisopropylethylamine (DIPEA) (58 μL, 0.33 mmol) were added and stirred at ambient temperature for 3 h. The water was added and the precipitate was filtered, washed with mixture of n-hexane and ethyl acetate and vacuum dried to obtain 4,5-dibromo-N-[4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-pyrrole-2-carboxamide (intermediate **V**, 98 mg, 0.18 mmol, yield: 81%) as a solid. TLC *Rf*: 0.47 (n-hexane/ethyl acetate, 1/2), 1H NMR (300 MHz, $CDCl_3$) δ 8.68 (s, 1H), 8.40 (d, J = 8.8 Hz, 1H), 8.24 (s, 1H), 7.25 (s, 1H), 7.15 – 7.06 (m, 2H), 6.81 (s, 1H), 4.02 (s, 3H), 3.97 (s, 3H), 3.95 (s, 3H).

In a sealed tube, Compound **V** (81 mg, 0.15 mmol) was dissolved in 1,4-dioxane (2 mL). Saturated aqueous ammonia (1 mL) was added to this and microwaved at 130 °C for 6 h. The reaction mixture was concentrated and the residue obtained was purified by flash chromatography using hexane/ethyl acetate and DCM/methanol gradient to obtain N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-4,5-dibromo-1-methyl-pyrrole-2-carboxamide (intermediate **VI**, 35 mg, 0.066 mmol, yield: 45%) as a solid. TLC *Rf*: 0.44 (ethyl acetate/methanol, 19/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.43 (d, J = 8.2 Hz, 1H), 8.36 (s, 1H), 8.21 (s, 1H), 7.10 (dd, J = 8.2, 1.8 Hz, 1H), 7.02 (d, J

= 1.8 Hz, 1H), 6.95 (s, 1H), 6.82 (s, 1H), 5.11 (s, 2H), 4.02 (s, 3H), 3.97 (s, 3H), 3.85 (s, 3H). MS m/z: 535 [$(M+H)^+$].

In a sealed tube, Compound **VI** (18 mg, 0.032 mmol) and phenylboronic acid (12 mg, 0.096 mmol) were dissolved in 1,4-dioxane (2 mL). To the solution, K_3PO_4 (41 mg, 0.19 mmol), water (0.5 mL) and $PdCl_2(dppf)_2$ DCM complex (5.2 mg, 0.0064 mmol) were added and then microwaved at 110°C for 2 h under a nitrogen atmosphere. The reaction mixture was concentrated under reduced pressure. The residue was purified by flash chromatography using hexane/ethyl acetate and DCM/methanol gradient to obtain N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-4,5-diphenyl-pyrrole-2-carboxamide (Compound **16**, 6.2 mg, 0.012 mmol, yield: 37%) as a solid. TLC *Rf*: 0.49 (ethyl acetate/methanol, 19/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.52 (d, J = 8.2 Hz, 1H), 8.43 (s, 1H), 8.31 (s, 1H), 7.45 – 7.39 (m, 3H), 7.36 – 7.07 (m, 8H), 7.02 (d, J = 1.7 Hz, 1H), 6.98 (s, 1H), 6.97 (s, 1H), 5.69 (s, 2H), 3.99 (s, 3H), 3.86 (s, 3H), 3.84 (s, 3H). MS m/z: 528.6 [$(M+H)^+$].

Scheme 17:



Scheme 17: *A) Oxalyl chloride, DMF, DCM; B)* ***VII*** *in Scheme 1, N,N-Diisopropylethylamine (DIPEA), THF; C) Aqueous ammonia, Dioxane 130 °C(microwave).*

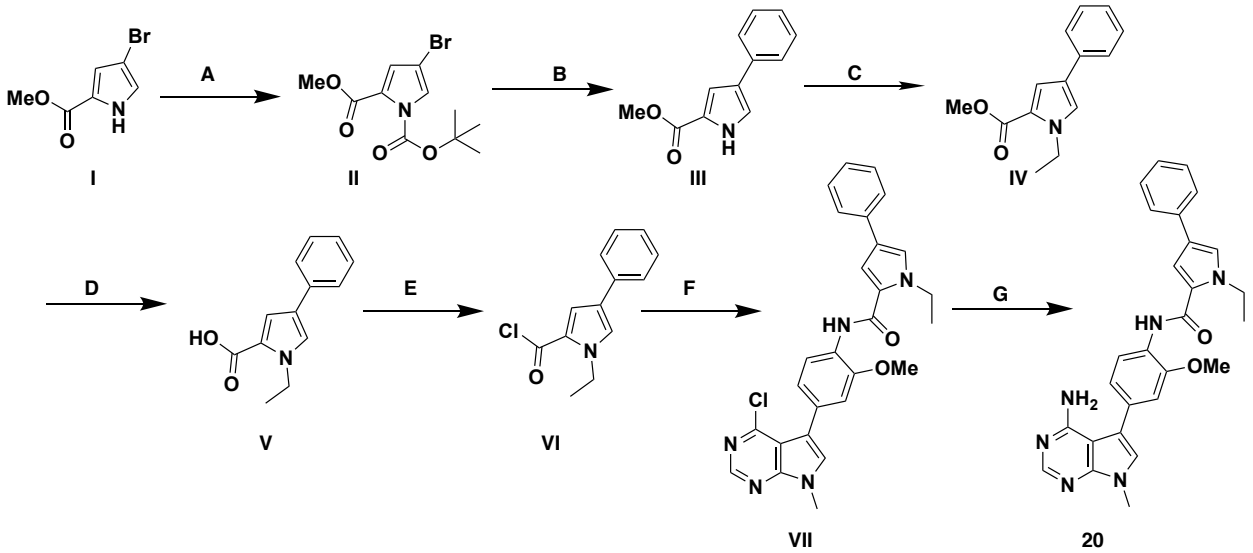
To a solution of 2-methyl-5-phenyl-pyrazole-3-carboxylic acid (**I**) (130 mg, 0.63 mmol) in DCM (5 mL), oxalyl chloride (59 μL, 0.75 mmol) was added and then a drop of DMF by using pasteur pipette.

The reaction mixture was stirred at ambient temperature for 30 min. The reaction mixture was concentrated under reduced pressure and the residue obtained was conducted azeotropic distillation with toluene twice and vacuum dried to obtain the crude product of 2-methyl-5-phenyl-pyrazole-3-carbonyl chloride (intermediate **II**). 1H NMR (300 MHz, $CDCl_3$) δ 7.85 – 7.75 (m, 2H), 7.49 – 7.30 (m, 4H), 4.19 (s, 3H).

To a solution of 4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-aniline (Compound **IV** in scheme 1) (36 mg, 0.13 mmol) in THF (6 mL), Compound **II** (33 mg, 0.15 mmol) and N,N-diisopropylethylamine (DIPEA) (44 µL, 0.25 mmol) were added and stirred at ambient temperature for 1 hour. The water was added and the precipitate was filtered, washed with mixture of n-hexane and ethyl acetate and vacuum dried to obtain N-[4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-2-methyl-5-phenyl-pyrazole-3-carboxamide (intermediate **III**, 52 mg, 0.11 mmol, yield: 88%) as a solid. TLC *Rf*: 0.59 (DCM/ethyl acetate, 3/2), 1H NMR (300 MHz, $CDCl_3$) δ 8.69 (s, 1H), 8.51 – 8.41 (m, 2H), 7.84 (d, J = 7.2 Hz, 2H), 7.47 – 7.40 (m, 2H), 7.35 (t, J = 7.2 Hz, 1H), 7.26 (s, 1H), 7.17 – 7.13 (m, 2H), 6.94 (s, 1H), 4.29 (s, 3H), 4.00 (s, 3H), 3.95 (s, 3H).

In a sealed tube, Compound **III** (39 mg, 0.083 mmol) was dissolved in 1,4-dioxane (2 mL). Saturated aqueous ammonia (2 mL) was added to this and microwaved at 140 °C for 7 h. The reaction mixture was concentrated and the residue obtained was purified by flash chromatography using hexane/ethyl acetate and DCM/methanol gradient to obtain N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-2-methyl-5-phenyl-pyrazole-3-carboxamide (Compound **19**, 31 mg, 0.067 mmol, yield: 81%) as a solid. TLC *Rf*: 0.47 (ethyl acetate/methanol, 9/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.51 (d, J = 8.2 Hz, 1H), 8.41 (s, 1H), 8.36 (s, 1H), 7.87 – 7.80 (m, 2H), 7.49 – 7.40 (m, 2H), 7.37 (dd, J = 5.0, 3.6 Hz, 1H), 7.14 (dd, J = 8.2, 1.8 Hz, 1H), 7.05 (d, J = 1.8 Hz, 1H), 6.97 (s, 1H), 6.94 (s, 1H), 5.18 (s, 2H), 4.29 (s, 3H), 4.00 (s, 3H), 3.86 (s, 3H). MS m/z: 453.5[$(M+H)^+$].

Scheme 18:



Scheme 18: *A) Di-tert-butyl dicarbonate, Et_3N, N,N-dimethyl-4-aminopyridine(DMAP), THF; B) Phenylboronic acid, $Pd(PPh_3)_4$, Na_2CO_3, DMF-H_2O; C) Ethyl Iodide, Sodium hydride, DMF; D) NaOHaq, MeOH-THF; E) Oxalyl chloride, DMF, DCM; F)* ***VII*** *in Scheme 1, N,N-Diisopropylethylamine (DIPEA), THF; G) Aqueous ammonia, Dioxane 130 °C(microwave).*

To a solution of methyl 4-bromo-1H-pyrrole-2-carboxylate (2.5 g, 12.4 mmol) in anhydrous THF (50 mL), di-tert-butyl decarbonate (3.0 g, 13.6 mmol), triethylamine (2.1 mL, 14.8 mmol) and N,N-dimethyl-4-aminopyridine (0.30 g, 2.47 mmol) were added and stirred at ambient temperature for 12 h. Saturated aqueous solution of $NaHCO_3$ was added, followed by extraction with ethyl acetate. The organic layer was dried over anhydrous $MgSO_4$, filtered and concentrated. The residue obtained was purified by flash chromatography using hexane/ethyl acetate gradient to obtain 1-(tert-butyl) 2-methyl 4-bromo-1H-pyrrole-1,2-dicarboxylate (intermediate **II**, 3.7 g, 12.2 mmol, yield: 99%) as a solid. TLC *Rf*: 0.63 (n-hexane/ethyl acetate, 3/1), 1H NMR (300 MHz, $CDCl_3$) δ 7.31 (d, J = 1.9 Hz, 1H), 6.79 (d, J = 1.9 Hz, 1H), 3.84 (s, 3H), 1.57 (s, 9H).

Compound **II** (1.8 g, 5.95 mmol) was dissolved in DMF (50 mL), To this solution, phenylboronic acid (1.8 g, 14.9 mmol), tetrakis(triphenylphosphine)palladium(0) (340 mg, 0.30 mmol) were added and

then added aqueous solution of Na_2CO_3 (5.7 g, 53.5 mmol, water; 15 mL). The mixture was stirred at 110°C for 12 h under a nitrogen atmosphere.

Water was added and followed by extraction with ethyl acetate. The organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated. The residue obtained was purified by flash chromatography using hexane/ethyl acetate gradient to obtain methyl 4-phenyl-1H-pyrrole-2-carboxylate (intermediate **III**, 0.48 g, 2.37 mmol, yield: 40%) as a solid. TLC *Rf*: 0.34 (n-hexane/ethyl acetate, 3/1), 1H NMR (300 MHz, $CDCl_3$) δ 9.14 (s, 1H), 7.55 – 7.48 (m, 2H), 7.36 (dd, J = 10.8, 4.5 Hz, 2H), 7.27 – 7.18 (m, 3H), 3.89 (s, 3H).

Compound **III** (75 mg, 0.37 mmol) was dissolved in DMF (2 mL). Sodium hydride (60wt% in oil) (18 mg, 0.44 mmol) and ethyl iodide (30 μL, 0.37 mmol) were added and stirred at ambient temperature for 30 min. Water was added and followed by extraction with ethyl acetate. The organic layer was dried over anhydrous $MgSO_4$, filtered and concentrated. The residue was purified by flash chromatography using hexane/ethyl acetate gradient to obtain methyl 1-ethyl-4-phenyl-pyrrole-2-carboxylate (intermediate **IV,** 42 mg, 0.19 mmol, yield: 52%) as a solid. TLC *Rf*: 0.52 (n-hexane/ethyl acetate, 4/1), 1H NMR (300 MHz, $CDCl_3$) δ 7.49 (dd, J = 8.2, 1.1 Hz, 2H), 7.35 (dd, J = 10.5, 4.8 Hz, 2H), 7.25 – 7.13 (m, 3H), 4.39 (q, J = 7.2 Hz, 2H), 3.84 (s, 3H), 1.44 (t, J = 7.2 Hz, 3H).

To a solution of Compound **IV** (38 mg, 0.16 mmol) in methanol (3 mL), 1mol/L Sodium hydroxide aqueous solution (0.50 mL, 0.49 mmoL) was added to this and stirred at 40 °C for 96 h. 2 mol/L hydrochloric acid was added and followed by extraction with ethyl acetate. The organic layer was dried over anhydrous $MgSO_4$, filtered, concentrated and vacuum dried to obtain 1-ethyl-4-phenyl-pyrrole-2-carboxylic acid (intermediate **V**, 38 mg, quant) as a solid. TLC *Rf*: 0.56 (n-hexane/ethyl acetate, 1/1),

1H NMR (300 MHz, $CDCl_3$) δ 7.55 – 7.46 (m, 2H), 7.41 – 7.31 (m, 3H), 7.26 – 7.17 (m, 2H), 4.40 (q, J = 7.2 Hz, 2H), 1.46 (t, J = 7.2 Hz, 3H).

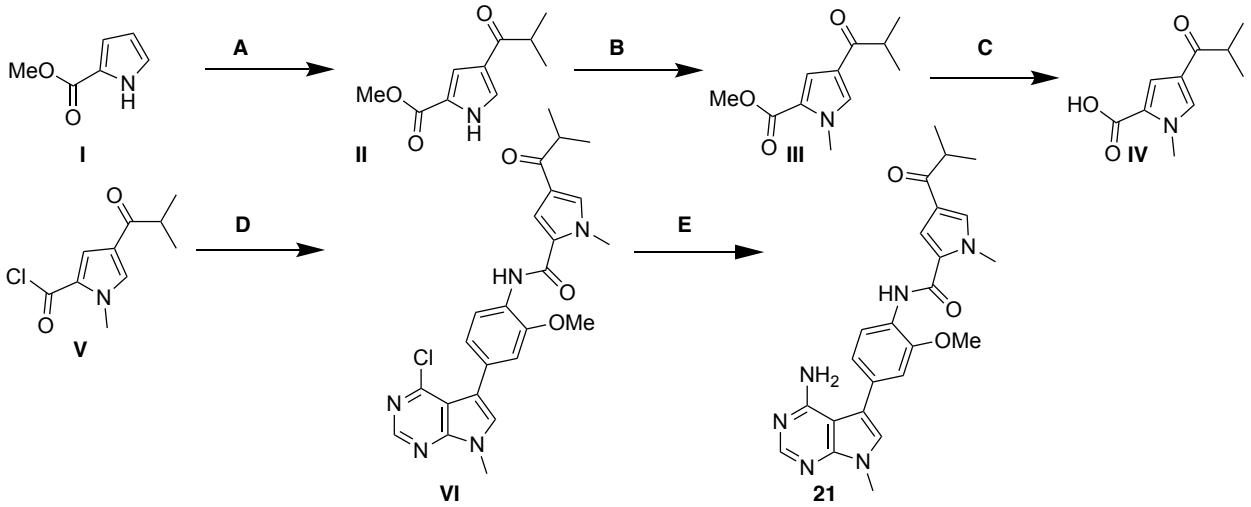
To a solution of Compound **V** (37 mg, 0.17 mmol) in DCM (4 mL), oxalyl chloride (16 µL, 0.21 mmol) was added and then a drop of DMF by using pasteur pipette. The mixture was stirred at ambient temperature for 30 min. The reaction mixture was concentrated under reduced pressure and the residue obtained was conducted azeotropic distillation with toluene twice and vacuum dried to obtain the crude product of 1-ethyl-4-phenyl-pyrrole-2-carbonyl chloride (intermediate **VI**).1H NMR (300 MHz, $CDCl_3$) δ 7.57 – 7.47 (m, 3H), 7.42 – 7.31 (m, 3H), 7.29 – 7.18 (m, 1H), 4.30 (q, J = 7.2 Hz, 2H), 1.44 (t, J = 7.2 Hz, 3H).

To a solution of 4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-aniline (Compound **IV** in scheme 1) (33 mg, 0.11 mmol) in THF (4 mL), Compound **VI** (40 mg, 0.17 mmol) and N,N-diisopropylethylamine (DIPEA) (40 µL, 0.23 mmol) were added and stirred at ambient temperature for 48 hour. The water was added and the precipitate was filtered, washed with mixture of n-hexane and ethyl acetate and vacuum dried to obtain N-[4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-ethyl-4-phenyl-pyrrole-2-carboxamide (intermediate **VII**, 33 mg, 0.067 mmol, yield: 59%) as a solid. TLC *Rf*: 0.58 (n-hexane/ethyl acetate, 1/2), 1H NMR (300 MHz, $CDCl_3$) δ 8.68 (s, 1H), 8.47 (d, J = 8.7 Hz, 1H), 8.40 (s, 1H), 7.54 (d, J = 7.1 Hz, 2H), 7.38 (t, J = 7.7 Hz, 2H), 7.27 – 7.22 (m, 2H), 7.17 (d, J = 1.9 Hz, 1H), 7.14 – 7.09 (m, 2H), 6.99 (d, J = 1.9 Hz, 1H), 4.48 (q, J = 7.1 Hz, 2H), 3.99 (s, 3H), 3.95 (s, 3H), 1.50 (t, J = 7.2 Hz, 3H).

In a sealed tube, Compound **VII** (33 mg, 0.067 mmol) was dissolved in 1,4-dioxane (2 mL). Saturated aqueous ammonia (1 mL) was added to this and microwaved at 140 °C for 6 h. The reaction mixture

was concentrated and the residue obtained was purified by flash chromatography using hexane/ethyl acetate and DCM/methanol gradient to obtain N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-ethyl-4-phenyl-pyrrole-2-carboxamide (Compound **20**, 25 mg, 0.053 mmol, yield: 79%) as a solid. TLC *Rf*: 0.43 (ethyl acetate/methanol, 9/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.49 (d, J = 8.2 Hz, 1H), 8.38 – 3.33 (m, 2H), 7.57 – 7.50 (m, 2H), 7.38 (t, J = 7.6 Hz, 2H), 7.22 (t, J = 7.4 Hz, 1H), 7.17 (d, J = 1.8 Hz, 1H), 7.11 (dd, J = 8.2, 1.8 Hz, 1H), 7.02 (d, J = 1.7 Hz, 1H), 6.99 (d, J = 1.9 Hz, 1H), 6.96 (s, 1H), 5.24 (s, 2H), 4.48 (q, J = 7.2 Hz, 2H), 3.99 (s, 3H), 3.85 (s, 3H), 1.50 (t, J = 7.2 Hz, 3H). MS m/z: 466.6 [$(M+H)^+$].

Scheme 19:



Scheme 19: *A) Isobutyl chloride; B) Methyl Iodide, Cs_2CO_3, DMF; C) NaOHaq, MeOH-THF; D) Oxalyl chloride, DMF, DCM; E)* ***VII*** *in Scheme 1, N,N-Diisopropylethylamine (DIPEA), THF; F) Aqueous ammonia, Dioxane 130 °C(microwave).*

To a solution of methyl 1H-pyrrole-2-carboxylate (**I**) (520 mg, 4.13 mmol) in DCM (5 mL), isobutyryl chloride (570 μL, 5.36 mmol) was added and the mixture was cooled in ice bath. To this solution, aluminium(III) chloride (1.7 g, 12.4 mmol) was added and stirred at ambient temperature for 12 h. Saturated aqueous solution of $NaHCO_3$ was added and followed by extraction with ethyl acetate. The organic layer was dried over anhydrous $MgSO_4$, filtered, concentrated and vacuum dried to obtain the

crude product containing methyl 4-(2-methylpropanoyl)-1H-pyrrole-2-carboxylate (intermediate **II**) as a solid.

The crude product containing Compound **II** (190 mg, ~0.96 mmol) was dissolved in DMF (6 mL). To the solution, Cs_2CO_3 (620 mg, 1.92 mmol) and methyl iodide (120 μL, 1.44 mmol) was added and the mixture was stirred at ambient temperature for 12 h. Water was added and the precipitate was filtered and vacuum dried to obtain methyl 1-methyl-4-(2-methylpropanoyl)pyrrole-2-carboxylate (intermediate **III**, 160 mg, 0.76 mmol, yield: 80%) as a solid. TLC *Rf*: 0.41 (n-hexane/ethyl acetate, 3/1), 1H NMR (300 MHz, $CDCl_3$) δ 7.39 (d, J = 1.7 Hz, 1H), 7.33 (d, J = 1.7 Hz, 1H), 3.96 (s, 3H), 3.84 (s, 3H), 3.24 – 3.11 (m, 1H), 1.18 (d, J = 6.9 Hz, 6H).

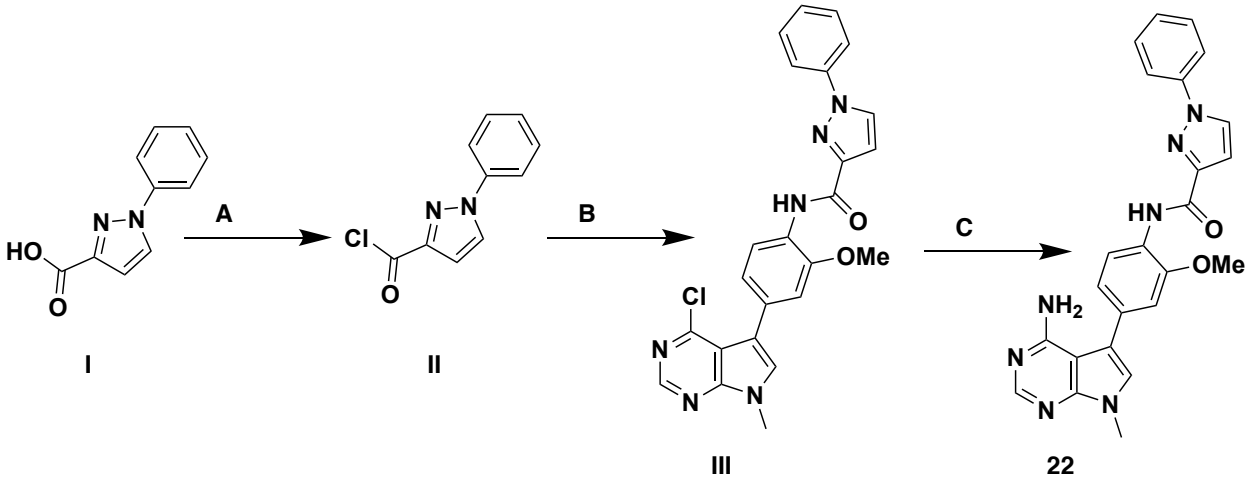
To a solution of Compound **III** (140 mg, 0.65 mmol) in mixture of THF (2 mL) and methanol (2 mL), 1mol/L Sodium hydroxide aqueous solution (2.0 mL, 2.0 mmoL) was added to this and stirred at 40 °C for 96 h. 2 mol/L hydrochloric acid was added and followed by extraction with ethyl acetate. The organic layer was dried over anhydrous $MgSO_4$, filtered, concentrated and vacuum dried to obtain 1-methyl-4-(2-methylpropanoyl)pyrrole-2-carboxylic acid (intermediate **IV**, 120 mg, 0.60 mmol, yield; 91%) as a solid. TLC *Rf*: 0.19 (n-hexane/ethyl acetate, 1/1), 1H NMR (300 MHz, $CDCl_3$) δ 7.48 – 7.43 (m, 2H), 3.97 (s, 3H), 3.26 – 3.13 (m, 1H), 1.19 (d, J = 6.9 Hz, 6H).

To a solution of Compound **IV** (56 mg, 0.29 mmol) in DCM (4 mL), oxalyl chloride (27 μL, 0.34 mmol) was added and then a drop of DMF by using pasteur pipette. The mixture was stirred at ambient temperature for 30 min. The reaction mixture was concentrated under reduced pressure and the residue obtained was conducted azeotropic distillation with toluene twice and vacuum dried to obtain the crude product of 1-methyl-4-(2-methylpropanoyl)pyrrole-2-carbonyl chloride (intermediate **V**). 1H NMR (300 MHz, $CDCl_3$) δ 7.65 (d, J = 1.9 Hz, 1H), 7.54 (d, J = 1.9 Hz, 1H), 3.92 (s, 3H), 3.28 – 3.11 (m, 1H), 1.20 (d, J = 6.9 Hz, 6H).

To a solution of 4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-aniline (Compound **IV** in scheme 1) (55 mg, 0.19 mmol) in THF (3 mL), Compound **V** (62 mg, 0.29 mmol) and N,N-diisopropylethylamine (DIPEA) (66 µL, 0.38 mmol) were added and stirred at ambient temperature for 12 hour. water was added and followed by extraction with mixture of ethyl acetate and THF. The organic layer was dried over anhydrous $MgSO_4$, filtered and concentrated. The solid residue obtained was wash with mixture of n-hexane and ethyl acetate and vacuum dried to obtain N-[4-(4-chloro-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-4-(2-methylpropanoyl)pyrrole-2-carboxamide (intermediate **VI**, 84 mg, 0.181 mmol, yield: 95%) as a solid. TLC *Rf*: 0.38 (n-hexane/ethyl acetate, 2/3), 1H NMR (300 MHz, $CDCl_3$) δ 8.68 (s, 1H), 8.47 – 8.35 (m, 2H), 7.38 (d, J = 1.7 Hz, 1H), 7.25 (s, 1H), 7.19 (d, J = 1.7 Hz, 1H), 7.15 – 7.07 (m, 2H), 4.04 (s, 3H), 3.98 (s, 3H), 3.94 (s, 3H), 3.29 – 3.11 (m, 1H), 1.22 (d, J = 6.8 Hz, 6H).

In a sealed tube, Compound **VI** (67 mg, 0.14 mmol) was dissolved in 1,4-dioxane (3 mL). Saturated aqueous ammonia (2 mL) was added to this and microwaved at 140 °C for 7 h. The reaction mixture was concentrated and the residue obtained was purified by flash chromatography using hexane/ethyl acetate and DCM/methanol gradient to obtain N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-methyl-4-(2-methylpropanoyl)pyrrole-2-carboxamide (Compound **21**, 45 mg, 0.10 mmol, yield: 71%) as a solid. TLC *Rf*: 0.39 (ethyl acetate/methanol, 19/1), 1H NMR (300 MHz, $CDCl_3$) δ 8.45 (d, J = 8.2 Hz, 1H), 8.36 (s, 2H), 7.39 (d, J = 1.8 Hz, 1H), 7.20 (d, J = 1.8 Hz, 1H), 7.11 (dd, J = 8.2, 1.8 Hz, 1H), 7.02 (d, J = 1.8 Hz, 1H), 6.96 (s, 1H), 5.14 (s, 2H), 4.04 (s, 3H), 3.98 (s, 3H), 3.85 (s, 3H), 3.29 – 3.09 (m, 1H), 1.22 (d, J = 6.9 Hz, 6H). MS m/z: 446.5 [$(M+H)^+$].

Scheme 20:



Scheme 20: *A) Oxalyl chloride, DMF, DCM; B)* ***VII*** *in Scheme 1, N,N-Diisopropylethylamine (DIPEA), THF; C) Aqueous ammonia, Dioxane 130 °C(microwave).*

N-[4-(4-amino-7-methyl-pyrrolo[2,3-d]pyrimidin-5-yl)-2-methoxy-phenyl]-1-phenyl-pyrazole-3-carboxamide (Compound **22**) was obtained as a solid by using 1-phenyl-1H-pyrrole-3-carboxylic acid instead of 2-methyl-5-phenyl-pyrazole-3-carboxylic acid and in the procedure for Compound **19** and carrying out the subsequent reaction by the same operation. TLC *Rf*: 0.47 (ethyl acetate/methanol, 19/1), 1H NMR (300 MHz, $CDCl_3$) δ 9.49 (s, 1H), 8.63 (d, J = 8.2 Hz, 1H), 8.37 (s, 1H), 7.99 (d, J = 1.7 Hz, 1H), 7.78 (d, J = 7.4 Hz, 2H), 7.54 (t, J = 7.8 Hz, 2H), 7.39 (t, J = 7.4 Hz, 1H), 7.15 - 7.11 (m, 2H), 7.04 (d, J = 1.7 Hz, 1H), 6.97 (s, 1H), 5.16 (s, 2H), 4.01 (s, 3H), 3.86 (s, 3H). MS m/z: 439.5 $[(M+H)^+]$.

1.6 References

[1] Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002) The protein kinase complement of the human genome, *Science 298*, 1912-1934.

[2] Lahiry, P., Torkamani, A., Schork, N. J., and Hegele, R. A. (2010) Kinase mutations in human disease: interpreting genotype-phenotype relationships, *Nat Rev Genet 11*, 60-74.

[3] Ferguson, F. M., and Gray, N. S. (2018) Kinase inhibitors: the road ahead, *Nat Rev Drug Discov 17*, 353-377.

[4] Brandvold, K. R., Steffey, M. E., Fox, C. C., and Soellner, M. B. (2012) Development of a highly selective c-Src kinase inhibitor, *ACS Chem Biol 7*, 1393-1398.

[5] Telliez, J. B., Dowty, M. E., Wang, L., Jussif, J., Lin, T., Li, L., Moy, E., Balbo, P., Li, W., Zhao, Y., Crouse, K., Dickinson, C., Symanowicz, P., Hegen, M., Banker, M. E., Vincent, F., Unwalla, R., Liang, S., Gilbert, A. M., Brown, M. F., Hayward, M., Montgomery, J., Yang, X., Bauman, J., Trujillo, J. I., Casimiro-Garcia, A., Vajdos, F. F., Leung, L., Geoghegan, K. F., Quazi, A., Xuan, D., Jones, L., Hett, E., Wright, K., Clark, J. D., and Thorarensen, A. (2016) Discovery of a JAK3-Selective Inhibitor: Functional Differentiation of JAK3-Selective Inhibition over pan-JAK or JAK1-Selective Inhibition, *ACS Chem Biol 11*, 3442-3451.

[6] Xia, W., Mullin, R. J., Keith, B. R., Liu, L. H., Ma, H., Rusnak, D. W., Owens, G., Alligood, K. J., and Spector, N. L. (2002) Anti-tumor activity of GW572016: a dual tyrosine kinase inhibitor blocks EGF activation of EGFR/erbB2 and downstream Erk1/2 and AKT pathways, *Oncogene 21*, 6255-6263.

[7] Roskoski, R., Jr. (2016) Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes, *Pharmacol Res 103*, 26-48.

[8] Zhao, Z., Wu, H., Wang, L., Liu, Y., Knapp, S., Liu, Q., and Gray, N. S. (2014) Exploration of type II binding mode: A privileged approach for kinase inhibitor focused drug discovery?, *ACS Chem Biol 9*, 1230-1241.

[9] Liu, Y., and Gray, N. S. (2006) Rational design of inhibitors that bind to inactive kinase conformations, *Nat Chem Biol 2*, 358-364.

[10] Georghiou, G., Kleiner, R. E., Pulkoski-Gross, M., Liu, D. R., and Seeliger, M. A. (2012) Highly specific, bisubstrate-competitive Src inhibitors from DNA-templated macrocycles, *Nat Chem Biol 8*, 366-374.

[11] Kwarcinski, F. E., Brandvold, K. R., Phadke, S., Beleh, O. M., Johnson, T. K., Meagher, J. L., Seeliger, M. A., Stuckey, J. A., and Soellner, M. B. (2016) Conformation-Selective Analogues of Dasatinib Reveal Insight into Kinase Inhibitor Binding and Selectivity, *ACS Chem Biol 11*, 1296-1304.

[12] Palmieri, L., and Rastelli, G. (2013) alphaC helix displacement as a general approach for allosteric modulation of protein kinases, *Drug Discov Today 18*, 407-414.

[13] Ingley, E. (2008) Src family kinases: regulation of their activities, levels and identification of new pathways, *Biochim Biophys Acta 1784*, 56-65.

[14] Williams, J. C., Wierenga, R. K., and Saraste, M. (1998) Insights into Src kinase functions: structural comparisons, *Trends Biochem Sci 23*, 179-184.

[15] Fraser, C., Dawson, J. C., Dowling, R., Houston, D. R., Weiss, J. T., Munro, A. F., Muir, M., Harrington, L., Webster, S. P., Frame, M. C., Brunton, V. G., Patton, E. E., Carragher, N. O., and Unciti-Broceta, A. (2016) Rapid Discovery and Structure-Activity Relationships of Pyrazolopyrimidines That Potently Suppress Breast Cancer Cell Growth via SRC Kinase Inhibition with Exceptional Selectivity over ABL Kinase, *J Med Chem 59*, 4697-4710.

[16] Brandvold, K. R., Santos, S. M., Breen, M. E., Lachacz, E. J., Steffey, M. E., and Soellner, M. B. (2015) Exquisitely specific bisubstrate inhibitors of c-Src kinase, *ACS Chem Biol 10*, 1387-1391.

[17] Aleem, S., Georghiou, G., Kleiner, R. E., Guja, K., Craddock, B. P., Lyczek, A., Chan, A. I., Garcia-Diaz, M., Miller, W. T., Liu, D. R., and Seeliger, M. A. (2016) Structural and Biochemical Basis for Intracellular Kinase Inhibition by Src-specific Peptidic Macrocycles, *Cell Chem Biol 23*, 1103-1112.

[18] Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions, *Protein Eng 8*, 127-134.

[19] Leonard, S. E., Register, A. C., Krishnamurty, R., Brighty, G. J., and Maly, D. J. (2014) Divergent modulation of Src-family kinase regulatory interactions with ATP-competitive inhibitors, *ACS Chem Biol 9*, 1894-1905.

[20] Register, A. C., Leonard, S. E., and Maly, D. J. (2014) SH2-catalytic domain linker heterogeneity influences allosteric coupling across the SFK family, *Biochemistry 53*, 6910-6923.

[21] Stachlewitz, R. F., Hart, M. A., Bettencourt, B., Kebede, T., Schwartz, A., Ratnofsky, S. E., Calderwood, D. J., Waegell, W. O., and Hirst, G. C. (2005) A-770041, a novel and selective small-molecule inhibitor of Lck, prevents heart allograft rejection, *J Pharmacol Exp Ther 315*, 36-41.

[22] Wang, Q., Cahill, S. M., Blumenstein, M., and Lawrence, D. S. (2006) Self-reporting fluorescent substrates of protein tyrosine kinases, *J Am Chem Soc 128*, 1808-1809.

[23] Register, A. C., Chakraborty, S., and Maly, D. J. (2017) Allosteric Modulation of Src Family Kinases with ATP-Competitive Inhibitors, *Methods Mol Biol 1636*, 79-89.

[24] Bantscheff, M., Eberhard, D., Abraham, Y., Bastuck, S., Boesche, M., Hobson, S., Mathieson, T., Perrin, J., Raida, M., Rau, C., Reader, V., Sweetman, G., Bauer, A., Bouwmeester, T., Hopf, C., Kruse, U., Neubauer, G., Ramsden, N., Rick, J., Kuster, B., and Drewes, G. (2007)

Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors, *Nat Biotechnol 25*, 1035-1044.

[25] Golkowski, M., Brigham, J. L., Perera, G. K., Romano, G. E., Maly, D. J., and Ong, S. E. (2014) Rapid profiling of protein kinase inhibitors by quantitative proteomics, *Medchemcomm 5*, 363-369.

[26] Golkowski, M., Perera, G. K., Vidadala, V. N., Ojo, K. K., Van Voorhis, W. C., Maly, D. J., and Ong, S. E. (2018) Kinome chemoproteomics characterization of pyrrolo[3,4-c]pyrazoles as potent and selective inhibitors of glycogen synthase kinase 3, *Mol Omics 14*, 26-36.

[27] Golkowski, M., Vidadala, R. S., Lombard, C. K., Suh, H. W., Maly, D. J., and Ong, S. E. (2017) Kinobead and Single-Shot LC-MS Profiling Identifies Selective PKD Inhibitors, *J Proteome Res 16*, 1216-1227.

[28] Eid, S., Turk, S., Volkamer, A., Rippmann, F., and Fulle, S. (2017) KinMap: a web-based tool for interactive navigation through human kinome data, *BMC Bioinformatics 18*, 16.

[29] Hari, S. B., Perera, B. G., Ranjitkar, P., Seeliger, M. A., and Maly, D. J. (2013) Conformation-selective inhibitors reveal differences in the activation and phosphate-binding loops of the tyrosine kinases Abl and Src, *ACS Chem Biol 8*, 2734-2743.

[30] Wodicka, L. M., Ciceri, P., Davis, M. I., Hunt, J. P., Floyd, M., Salerno, S., Hua, X. H., Ford, J. M., Armstrong, R. C., Zarrinkar, P. P., and Treiber, D. K. (2010) Activation state-dependent binding of small molecule kinase inhibitors: structural insights from biochemistry, *Chem Biol 17*, 1241-1249.

[31] Lombard, C. K., Davis, A. L., Inukai, T., and Maly, D. J. (2018) Allosteric Modulation of JNK Docking Site Interactions with ATP-Competitive Inhibitors, *Biochemistry 57*, 5897-5909.

[32] Krishnamurty, R., Brigham, J. L., Leonard, S. E., Ranjitkar, P., Larson, E. T., Dale, E. J., Merritt, E. A., and Maly, D. J. (2013) Active site profiling reveals coupling between domains in SRC-family kinases, *Nat Chem Biol 9*, 43-50.

[33] Young, M. A., Gonfloni, S., Superti-Furga, G., Roux, B., and Kuriyan, J. (2001) Dynamic coupling between the SH2 and SH3 domains of c-Src and Hck underlies their inactivation by C-terminal tyrosine phosphorylation, *Cell 105*, 115-126.

[34] Theile, C. S., Witte, M. D., Blom, A. E., Kundrat, L., Ploegh, H. L., and Guimaraes, C. P. (2013) Site-specific N-terminal labeling of proteins using sortase-mediated reactions, *Nat Protoc 8*, 1800-1807.

[35] Ayrapetov, M. K., Wang, Y. H., Lin, X., Gu, X., Parang, K., and Sun, G. (2006) Conformational basis for SH2-Tyr(P)527 binding in Src inactivation, *J Biol Chem 281*, 23776-23784.

[36] Okada, M., Nada, S., Yamanashi, Y., Yamamoto, T., and Nakagawa, H. (1991) CSK: a protein-tyrosine kinase involved in regulation of src family kinases, *J Biol Chem 266*, 24249-24252.

[37] Levinson, N. M., Seeliger, M. A., Cole, P. A., and Kuriyan, J. (2008) Structural basis for the recognition of c-Src by its inactivator Csk, *Cell 134*, 124-134.

[38] Okada, M. (2012) Regulation of the SRC family kinases by Csk, *Int J Biol Sci 8*, 1385-1397.

[39] Staben, S. T., Feng, J. A., Lyle, K., Belvin, M., Boggs, J., Burch, J. D., Chua, C. C., Cui, H., DiPasquale, A. G., Friedman, L. S., Heise, C., Koeppen, H., Kotey, A., Mintzer, R., Oh, A., Roberts, D. A., Rouge, L., Rudolph, J., Tam, C., Wang, W., Xiao, Y., Young, A., Zhang, Y., and Hoeflich, K. P. (2014) Back pocket flexibility provides group II p21-activated kinase (PAK) selectivity for type I 1/2 kinase inhibitors, *J Med Chem 57*, 1033-1045.

[40] Sun, G., and Budde, R. J. (2001) Affinity purification of Csk protein tyrosine kinase based on its catalytic requirement for divalent metal cations, *Protein Expr Purif 21*, 8-12.

[41] Wissing, J., Jansch, L., Nimtz, M., Dieterich, G., Hornberger, R., Keri, G., Wehland, J., and Daub, H. (2007) Proteomics analysis of protein kinases by target class-selective prefractionation and tandem mass spectrometry, *Mol Cell Proteomics 6*, 537-547.

[42] Ahler, E., Register, A. C., Chakraborty, S., Fang, L., Dieter, E. M., Sitko, K. A., Vidadala, R. S. R., Trevillian, B. M., Golkowski, M., Gelman, H., Stephany, J. J., Rubin, A. F., Merritt, E. A., Fowler, D. M., and Maly, D. J. (2019) A Combined Approach Reveals a Regulatory Mechanism Coupling Src's Kinase Activity, Localization, and Phosphotransferase-Independent Functions, *Mol Cell*.

CHAPTER 2[2]: Massively parallel assessment of Src drug resistant variants reveals an unexpected hotspot for resistance

2.1 Abstract

Kinases are implicated in diseases and are one of the targets in cancer therapeutics. However, an outstanding challenge in the field of kinase drug discovery is the emergence of resistant variants in the target kinase long-term efficacy of the drug. Resistant variants can be varied, and even for the same kinase, multiple sites of resistance may appear. One way to address this challenge is to identify the mutational routes to drug resistance. Here, we employ a massively parallel yeast growth assay on the prototypical kinase Src and test a library of over 3500 Src variants in the presence of a panel of kinase inhibitors that bias the kinase towards divergent active site conformations. We observe that despite the divergent binding modes, most resistant hotspots are shared between inhibitors with few unique sites recognized for specific inhibitors. Counter to our expectation, the ligand-binding interface was not the most critical site of resistance, and we discovered hotspot of resistance at the β1/2 sheet where most residues make no contact with inhibitors. Biochemical and structural investigation of variants at the β1/2 sheet indicates that this site is regulatory and communicates with Src's regulatory apparatus through an allosteric network involving the dynamic phosphate-binding loop and the flexible helix αC.

[2] This chapter is adapted from a manuscript currently in preparation.

2.2 Introduction

Protein kinases are therapeutic targets for the treatment of various diseases, including cancer. Over the last twenty years there has been great success in developing small-molecule, ATP-competitive inhibitors that potently block kinase activity.[1-3] This had resulted in several Food and Drug Administration (FDA)-approved drugs that hold great promise for the treatment of cancer. However, many targeted kinases acquire resistance mutations that diminish the efficacy of small-molecule, ATP-competitive inhibitors. Drug resistance mechanisms can be diverse, unpredictable, and often multiple resistant mutants can appear in the same kinase. One recurring site of resistance in kinases is the 'gatekeeper' residue.[3, 4] Here, mutations obstruct drug binding by creating a steric clash with small molecules that interact with the ATP-binding site. For example, in BCR-ABL, mutation of the gatekeeper residue from a threonine to a bulkier isoleucine results in a variant with greatly diminished affinity for imatinib.[5] Studying how gatekeeper mutations confer resistance and how resistance can be overcome through the structural modification of inhibitors has been very informative for drug development. Therefore, identifying additional sites of resistance mutations, and understanding the mechanistic basis of resistance will enable the rational design of new kinase inhibitors that are more likely to overcome common mechanisms of drug resistance. However, the identification of sites that confer broad inhibitor resistance in kinases has so far been a cumbersome and time-consuming endeavor.[6]

Here, we have used a massively parallel Deep Mutational Scanning (DMS) technique paired with a well-characterized, yeast-based screen to gain insight into the drug-resistance of thousands of Src variants in a single high-throughput experiment. This technique allowed us to identify hundreds of resistant variants and pinpoint hotspots of resistance in Src. We used this screen to study how different modes of ATP-competitive inhibition impacted the emergence of resistant mutations in Src. [7-11] Using DMS, we were able to identify positions in Src's catalytic domain that confer broad inhibitor resistance

as well as positions that conferred resistance only against an inhibitor with a specific binding mode. Surprisingly, resistant mutations appeared throughout Src's catalytic domain and were not isolated to the ATP-binding site. We extensively characterized one hotspot of resistance that is located on the N-terminal lobe of Src's catalytic domain.

2.3 Results and discussion

Parallel analysis of dasatinib-resistant Src variants. To comprehensively profile mechanisms of Src drug resistance, we used an assay that relies on the correlation between Src-mediated phospho-tyrosine levels and *S. cerevisiae* growth inhibition[12, 13] to perform a deep mutational scan (DMS).[14, 15] We reasoned that a comparison of how mutations affect Src-mediated inhibition of yeast growth in the presence and absence of ATP-competitive inhibitors would provide a quantitative analysis of levels of drug resistance (**Figure 2-1 A**). Prior to performing a parallel analysis of Src mutants, we validated that ATP-competitive inhibitors, including the highly potent Src inhibitor dasatinib (**Figure 2-1 B**), were able to dose-dependently rescue the growth inhibition that resulted from expression of a full-length Src (Src^{FL}) construct in yeast (**Figure 2-2 A-C**). Next, we transformed a mutagenized library of Src^{FL}'s catalytic domain[12] into yeast, performed outgrowth in the presence (25 or 100 μM) or absence (DMSO) of dasatinib and collected samples at various time points. At each time point, we extracted plasmid DNA, amplified barcodes, and then deeply sequenced the barcode amplicons to quantify the frequency of each Src variant. From these data, residual activity scores were calculated for ~3,500 single amino acid Src mutants in yeast treated with DMSO or two different dasatinib concentrations (**Figures 2-1 C-F** and **2-2 D, E**).

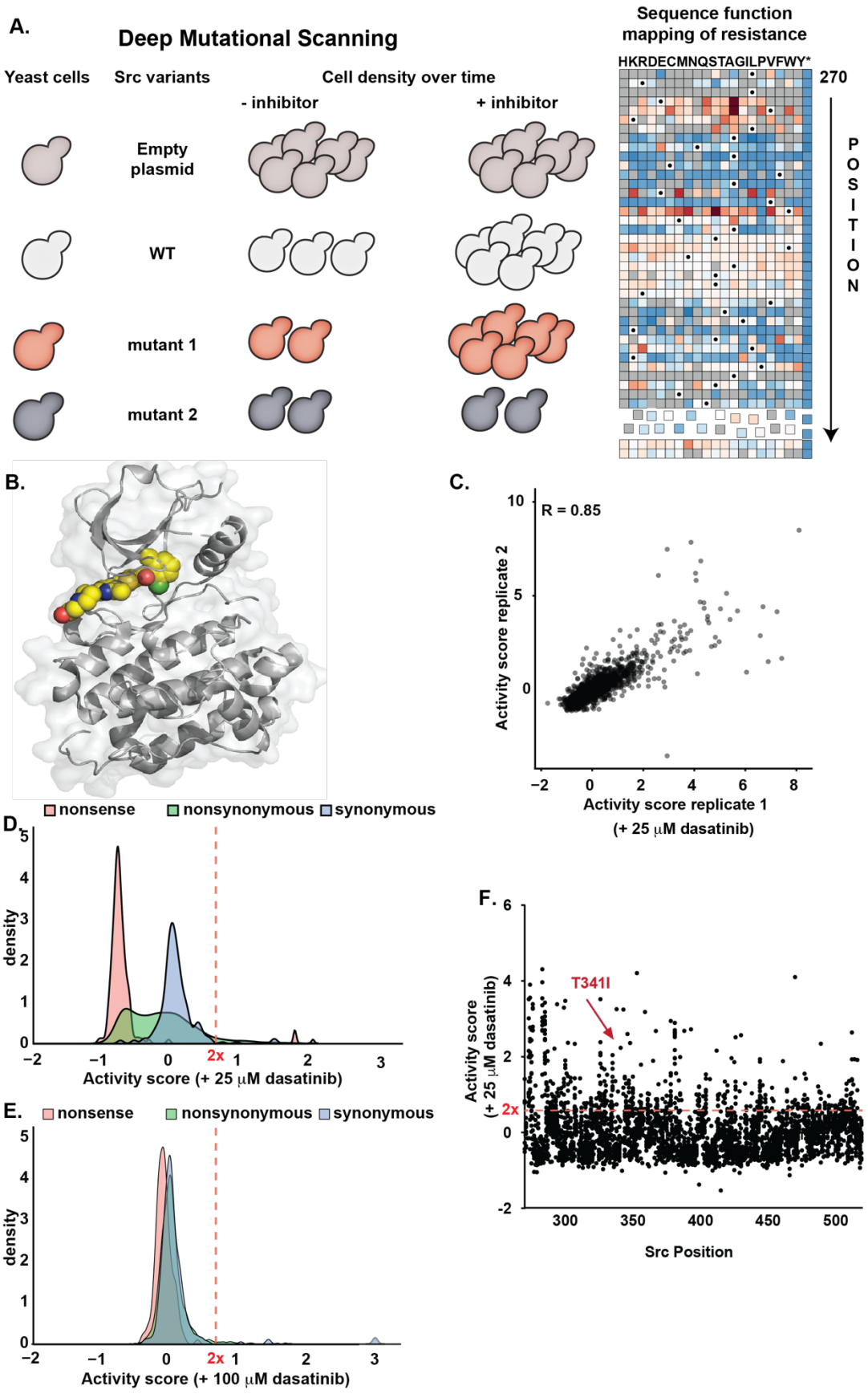

A.
Deep Mutational Scanning
Yeast cells
Src variants
Cell density over time
- inhibitor
+ inhibitor
Empty plasmid
WT
mutant 1
mutant 2
Sequence function mapping of resistance
HKRDECMNQSTAGILPVFWY*
270
POSITION
B.
C.
R = 0.85
Activity score replicate 2
Activity score replicate 1 (+ 25 μM dasatinib)
D.
nonsense
nonsynonymous
synonymous
density
Activity score (+ 25 μM dasatinib)
2x
E.
Activity score (+ 100 μM dasatinib)
F.
T341I
Activity score (+ 25 μM dasatinib)
Src Position

Figure 2-1: Interrogation of Src drug resistance using deep mutational scanning (DMS).

A. Schematic of the assay used to quantify the drug resistance of thousands of Src variants. **B.** Co-crystal structure of Src's catalytic domain (PDB id: 3G5D) bound to dasatinib (yellow sphere). **C.** Scatterplot showing the correlation (Pearson's R = 0.85) between the activity scores obtained in 25 μM dasatinib-treated yeast for two independent transformations of the Src^{FL} variant library. **D.** Activity score distributions of all synonymous and nonsynonymous (missense and nonsense) Src mutants for the Src^{FL} library treated with 25 μM dasatinib. **E.** Activity score distributions of all synonymous and nonsynonymous (missense and nonsense) Src mutants for the Src^{FL} library treated with 100 μM dasatinib. **F.** Scatterplot showing activity scores in yeast treated with 25 μM dasatinib for every single amino acid substitution at every position in Src's catalytic domain (each black dot represents an average of n=2 replicates). The crimson dashed line indicates the threshold beyond which variants were labeled as drug resistant. The gatekeeper mutant T341I is indicated by a red arrow.

Analysis of our DMS data revealed, as expected, synonymous Src mutations demonstrated uniform and dose-dependent decreases in activity scores in dasatinib-treated yeast (**Figure 2-1 C-D**). Furthermore, the activity scores of almost all catalytically active missense Src mutants were reduced, although to varying levels, in dasatinib-treated yeast compared to the DMSO control (**Figure 2-1 C-D, 2-2 E**). For each concentration of dasatinib treatment, we defined a Src mutant as being drug resistant if it showed an activity score greater than the 2 times the standard deviation of the synonymous distribution. We found that patterns of dasatinib resistance looked similar for both dasatinib concentrations but with activity scores generally being more compressed for 100 μM dasatinib treatment (**Figure 2-1 E**, **2-2 G**). Therefore, we selected the data obtained from 25 μM dasatinib treatment for the functional mapping of dasatinib resistance.

Functional mapping of dasatinib resistance. We assembled our large-scale DMS data into a sequence-function map covering more than 75% of all possible single mutants of Src's catalytic domain (**Figure 2-3 A**). About 12% of these single mutants are defined as dasatinib resistant based on our classification criteria. We observe that a majority (~60%) of dasatinib-resistant mutations are also classified as gain-of-function in DMSO-treated yeast (**Figure 2-4 A**), which implies that a common mechanism of resistance is through elevation of basal phosphotransferase activity.

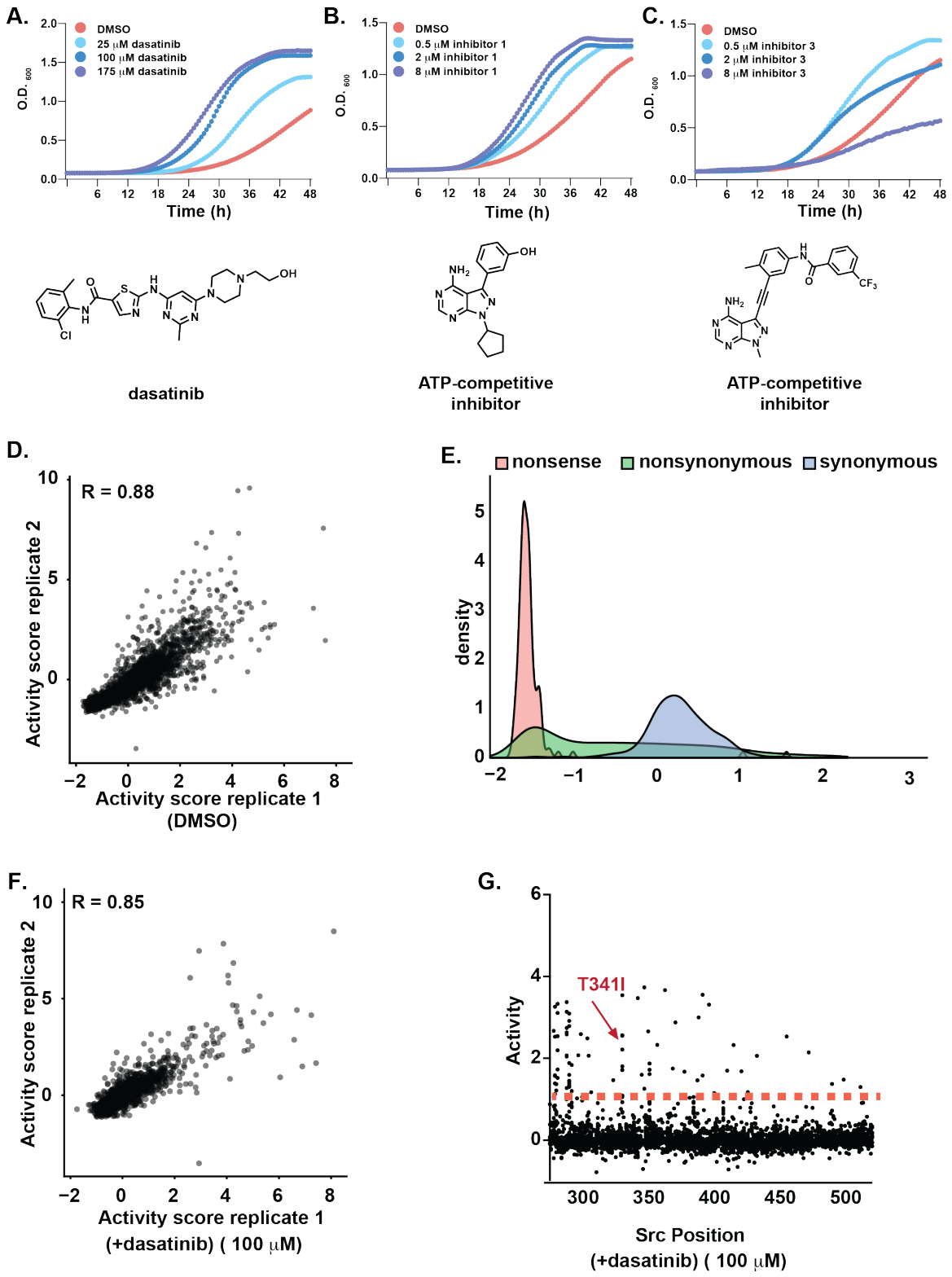


Figure 2-2: Dasatinib resistance.

A. (top) Growth curves for yeast transformed with wild-type Src in the presence of different concentrations of dasatinib. (bottom) Structure of dasatinib. **B.** (top) Growth curves for yeast transformed with wild-type Src in the presence of different concentrations of an ATP-competitive inhibitor. (bottom) Structure of the inhibitor. **C.** (top) Growth curves for yeast transformed with wild-type Src in the presence of different concentrations of an ATP-competitive inhibitor. (bottom) Structure of the inhibitor. **D.** Scatterplot showing the correlation (Pearson's R = 0.85) between the activity scores obtained in yeast treated with DMSO for two independent transformations of the Src^{FL} variant library. **E.** Activity score distributions of all synonymous and nonsynonymous (missense and nonsense) Src mutants in the presence of DMSO. **F.** Scatterplot showing the correlation (Pearson's R = 0.85) between the activity scores obtained in yeast treated with 100 μM dasatinib for two independent transformations of the Src^{FL} variant library. **G.** Scatterplot showing activity scores for every single amino acid substitution at every position in Src's catalytic domain in yeast treated with 100 μM dasatinib (each black dot represents an average of n=2 replicates). The crimson dashed line indicates the threshold beyond which variants were labeled as drug resistant. The gatekeeper mutant T341I is indicated by a red arrow.

However, there is not a strict correlation between activity scores and levels of dasatinib resistance (**Figure 2-4 B**), suggesting that mutations can confer resistance through other mechanisms. Our sequence-function map reveals that dasatinib-resistant mutations are distributed throughout the catalytic domain, with residues in the β1 and β2 strands of the N-terminal lobe and the αD and αE helices of the C-lobe being hotspots of resistance (**Figure 2-3 B**). Interestingly, for a number of positions relatively few substitutions lead to dasatinib resistance, while for other positions multiple mutations confer resistance. Together, these observations suggest that Src is capable of obtaining resistance to dasatinib inhibition through a number of different mechanisms.

A common mechanism of resistance to kinase inhibitors is through mutations that perturb the ATP-binding pocket.[6] We explored whether this mechanism of dasatinib resistance was captured in our DMS data. To determine this, we mapped maximum and position-averaged activity scores onto the fifteen residues in Src's ATP-binding site that interact with dasatinib (**Figure 2-3 C-D**). As expected, a number of mutations at positions that line the ATP-binding pocket led to substantial resistance to dasatinib inhibition, including the well-characterized T341I gatekeeper mutation. Residues Leu276

(located on the β1 strand), Val326 (located between the αC helix and β4 strand) and Tyr343 (located

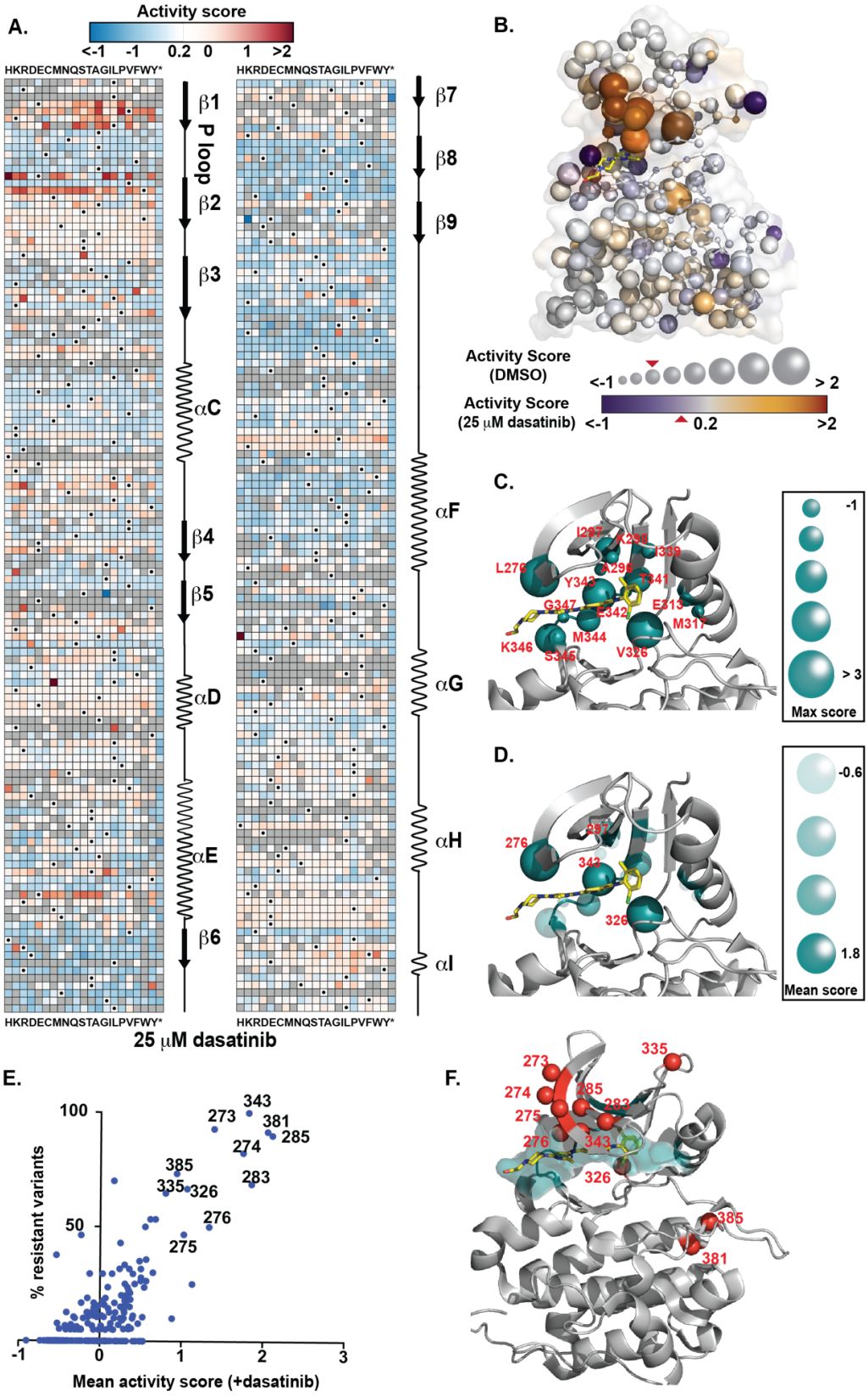


Figure 2-3. Comprehensive mapping of Src inhibitor resistant mutations.
A. Sequence-activity score map for all residues in Src's catalytic domain in the presence of 25 μM dasatinib. Black dots in the map represent the wild-type amino acid and gray tiles indicate missing data. Secondary structure and functional motif annotations were obtained from the ProKinO database. **B.** Position-averaged activity scores mapped on Src's catalytic domain (PDB id: 3G5D) for the Src^{FL} variant library in yeast treated with DMSO (activity scores represented by relative sphere size) and 25 μM dasatinib (activity scores represented by color). **C.** Co-crystal structure (PDB id: 3G5D) of Src's catalytic domain bound to dasatinib (yellow stick) with all fifteen residues that interact with dasatinib shown as green spheres. Sphere size corresponds to the maximum activity score for a substitution at this position. **D.** Co-crystal structure (PDB id: 3G5D) of Src's catalytic domain bound to dasatinib (yellow stick) with all fifteen residues that interact with dasatinib shown as green spheres. Sphere opaqueness corresponds to the average activity score of all substitutions at this position. **E.** Plot of the mean activity score for all substitutions at a position versus the percentage of all substitutions at this position that are classified as resistant for the Src^{FL} variant library in yeast treated with 25 μM dasatinib. **F.** Src CD (PDB id: 3G5D) showing all the residues that comprise the dasatinib resistant hotspots.

on the hinge region) all showed high position-averaged activity scores in the presence of dasatinib (**Figure 2-3 D-E**) and appear to be hotspots for resistance in this region. In total, there are ten dasatinib-interacting positions where specific substitutions resulted in resistance. Thus, ~70% of the residues that interact with dasatinib can be replaced with substitutions that maintain sufficient activity to confer resistance.

A large percentage of highly resistant mutations are located at positions that do not directly contact dasatinib (**Figure 2-3 E-F**). One site of notable resistance is Glu381, which forms part of the regulatory αF-pocket in the C-terminal lobe and possesses the most hyperactivating mutations in Src's catalytic domain. Four residues (Glu273, Val274, Glu283, and Trp285) that displayed some of the highest position-averaged activity scores in presence of dastainib form a cluster on the β1 and β2 strands of the N-terminal lobe that line the ATP-binding pocket but do not interact with dasatinib. Substitutions at these positions also tend to be highly hyperactivating in the absence of dasatinib.

Recently, Persky et. al. analyzed pablociclib binding interface in cyclin dependent kinase-6 (CDK6) and observed a similar pattern where despite there being specific resistant variants in the drug binding interface, important positions that lead to broad drug resistance appear all throughout the kinase

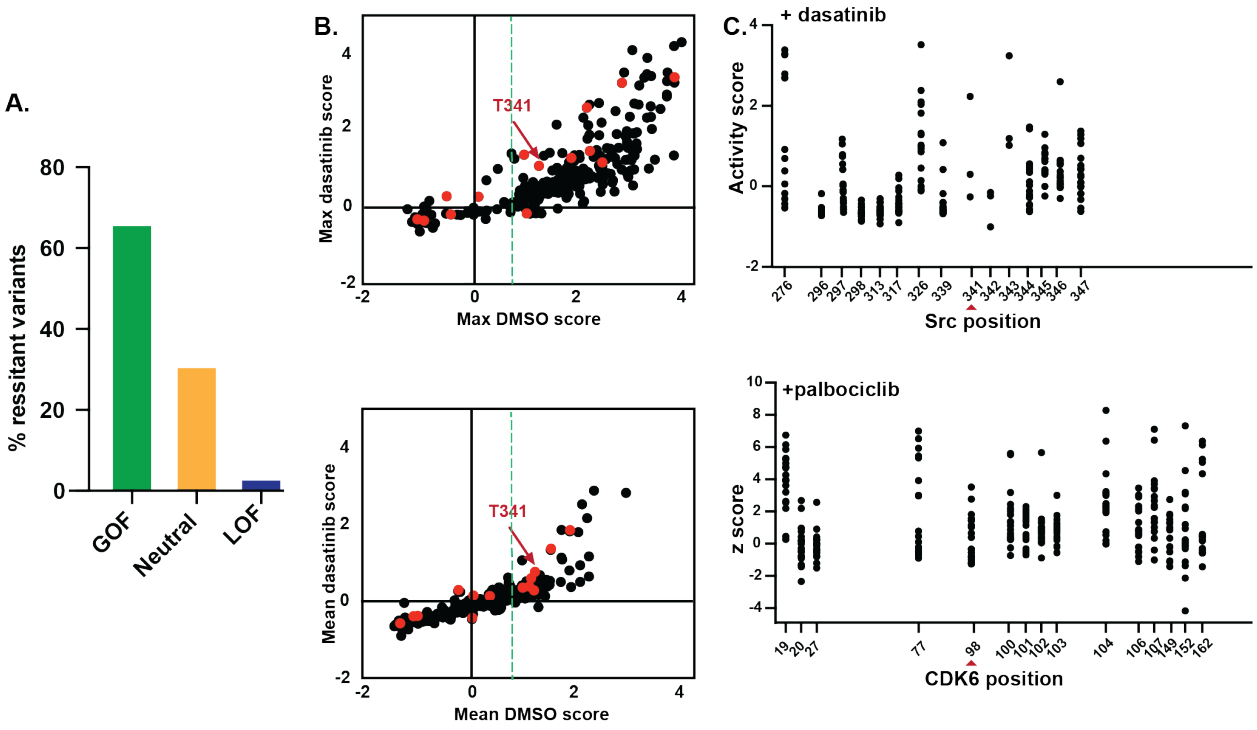
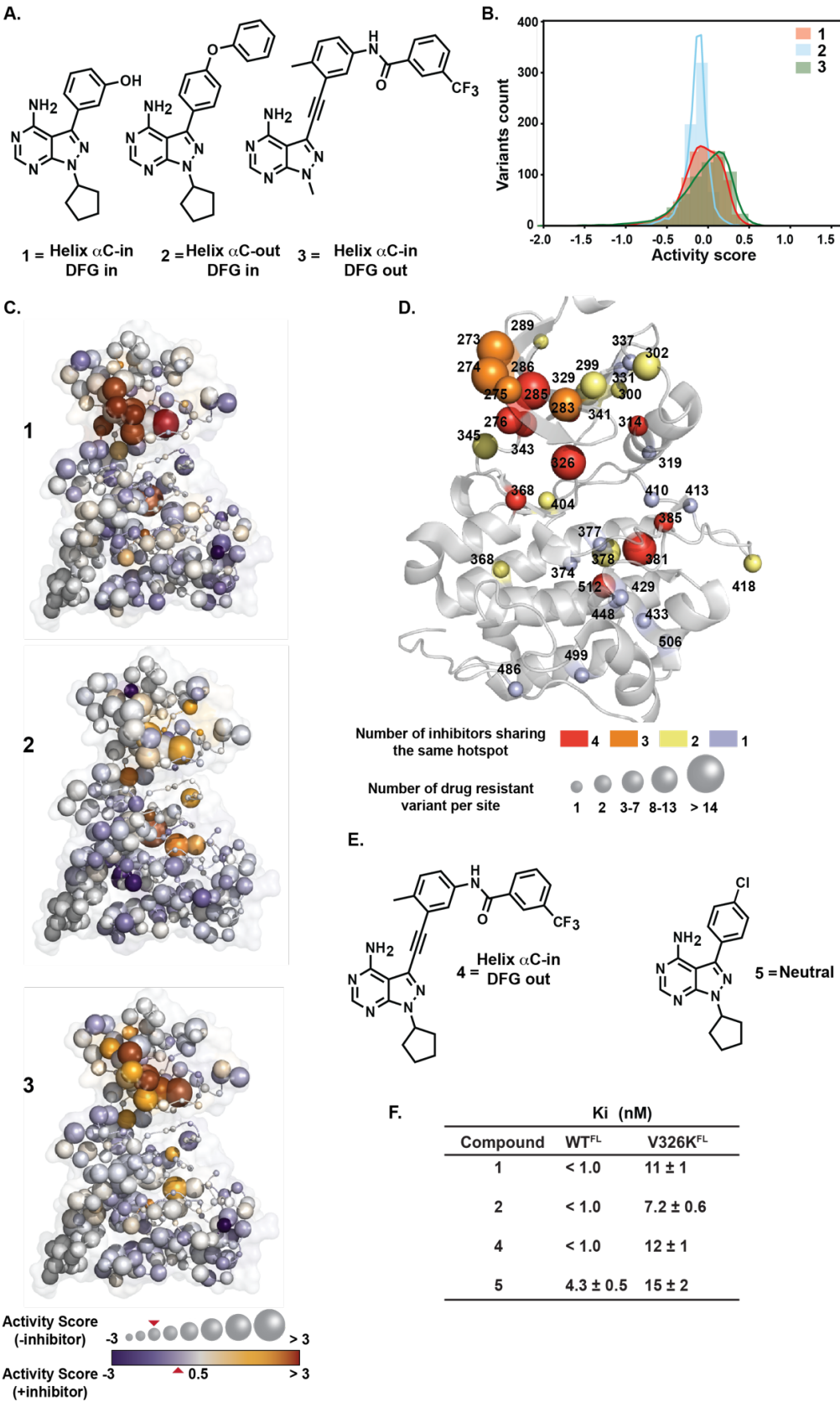


Figure 2-4: Comparison between inhibitor resistance in CDK6 and Src.
A. Percentage of dasatinib resistance mutants that are intrinsically gain of function, WT-like or loss of function with respect to inhibitor. Classifications were derived previously (Ahler et. al. 2019). **B.** (top) The maximum activity score for each Src residue in yeast treated with DMSO or 25 μM dasatinib. Red dots indicate residues that are classified as dasatinib interacting. Black dots indicate residues that do not directly contact dasatinib. (bottom) Comparison of mean activity scores in DMSO versus dasatinib for the Src^{FL} variant library. Red dots indicate residues that makes direct interaction with dasatinib and black dots indicate residues distal to inhibitor binding site. The green dotted line indicates the threshold beyond which variants were labeled as drug resistant. **C.** (top) Scatter plots showing the resistance score achieved for each substitution (y-axis, black dots) in Src for all sites that make contact with dasatinib. (bottom) Corresponding positions in CDK6 with calculated z-score achieved for each substitution (y-axis, black dots) at the palbociclib binding site. Data derived from (Persky et. al. 2020). Red triangle indicates the respective gatekeeper positions in Src (341) and CDK6 (98).

catalytic domain and not just at the drug binding site (**Figure 2-4 C**). Together, our DMS data suggest that Src is capable of obtaining resistance to dasatinib inhibition through a number of different mechanisms.

Resistance to conformation selective inhibitors. We next used our Src variant library to determine how resistance arises in response to different modes of ATP-competitive inhibition. To do this, we generated a set of pyrazolopyrimidine-based ATP-competitive small molecule kinase



Compound	Ki (nM) WT^{FL}	Ki (nM) V326K^{FL}
1	< 1.0	11 ± 1
2	< 1.0	7.2 ± 0.6
4	< 1.0	12 ± 1
5	4.3 ± 0.5	15 ± 2

Figure 2-5. Drug resistance with conformation selective inhibitors.
A. Structure of the ATP-competitive conformation selective inhibitors used in the yeast growth assay. **B.** Activity score distribution of all non-synonymous Src mutants under the treatment of inhibitors **1**, **2** or **3**. **C.** Src CD (PDB id: 3G5D) depicting the distribution of position-averaged activity scores in the presence of inhibitor **1** (top), **2** (center) or **3** (bottom) colored from a shade of orange indicating high score to purple indicating low score. Sphere size represents the position-averaged activity score in the absence of inhibitor. **D.** Src CD (PDB id: 3G5D) highlighting various positions that are shared between all four inhibitors, a sub-set of inhibitors or are unique to a specific inhibitor. The size of spheres represents the number of inhibitor-resistant variants observed at each position. **E.** Structures of two other conformation-selective inhibitors **4** and **5** that were used in the biochemical assays. **F**. Ki values (n = 3, mean ± SEM), determined in the presence of 1 mM ATP for purified Src^{FL} for both wild-type (WT) and the V326K mutant.

inhibitors (**1-3**, **Figure 2-5 A**), that we predicted would stabilize structurally distinct conformations of Src's ATP-binding pocket (**Figure 2-6 A**). Inhibitors **2** and **3** contain pharmacophores at the C-3 position of the pyrazolopyrimidine scaffold that promote the helix αC-out and DFG-out inactive conformations, respectively, while inhibitor **1** contains a 3-phenol at the C-3 position that can potentially form a hydrogen-bond with the side-chain of Glu313 in the helix αC, leading to stabilization of the active αC helix-in conformation (**Figure 2-6 A**). Although we attempted to only vary the C-3 substituent of the pyrazolopyrimidine scaffold, it was necessary to replace the N-1 group of inhibitor **3** with a smaller methyl group to achieve satisfactory penetration in yeast.

Next, we transformed the Src^{FL} variant library into yeast, performed outgrowth in the presence of each conformation-selective inhibitor and collected samples at various time points. For each inhibitor, we calculated residual activity scores for ~3,500 Src mutants, which were then quantile normalized (**Figure 2-5 B**). As for dasatinib, a Src mutant was defined as being resistant to an inhibitor if its residual activity score was greater than the 2 times the standard deviation of the synonymous distribution. Using this definition, we identified 107, 86, and 114 Src mutants that were resistant to inhibitors **1**, **2**, and **3**, respectively. In comparing the overlap between resistance mutations, we observed that 22 mutations that occur at nine positions in Src were resistant to all three inhibitors and ~25-45% of all resistance mutations for a particular inhibitor were unique. Inhibitors **1** and **3** shared

the highest overlap (~60%) in resistance mutations, suggesting that their shared ability to stabilize an active conformation of the αC helix makes them susceptible to similar mechanisms of resistance.

To compare patterns of resistance to **1**-**3**, we mapped the position-averaged residual activity scores for each inhibitor onto the catalytic domain of Src (**Figure 2-5 C**). We found that resistance mutations that were unique for inhibitor **1** or **3** were mainly clustered in the C-terminal lobe, while unique resistance mutations for inhibitor **2** spanned the entirety of Src's kinase domain (**Figure 2-6 D**). There are three positions (Leu300, Ile337, and Leu413) that contain multiple substitutions that uniquely confer strong resistance to the helix αC-out-stabilizing inhibitor **2**. Leu300 is located on the linker that connects the helix αC to the β-sheet of the N-terminal lobe and Ile337 is on the face of the β4 strand that is directed towards helix αC. A plausible explanation for how mutations at these positions confer resistance is that they negatively influence the ability of helix αC to adopt the inactive "out" conformation that is required for inhibitor **2** to be accommodated in Src's ATP-binding pocket.

There are nine positions that contain mutations that are resistant to all three inhibitors. Mutations at these positions are also capable of conferring resistance to dasatinib (**Figure 2-5 D**). We biochemically characterized one (Val326) of these nine positions. Val326 is on the loop connecting the helix αC to the β4 sheet and projects into the ATP-binding site. Substitutions at V326 are tolerated by Src and results in no change in activity (**Figure 2-5 E**, DMSO) according to a previously reported large scale mutagenesis data. We purified the full-length Src (Src^{FL}, **Figure 2-8 A**) wildtype (WT) and the V326K mutantsand measured their respective phosphotransferase activities. Consistent with mutagenesis data[12] we see no difference in kinase activity or Km[ATP] between WT or the mutant (**Figure 2-6 F**). V326 is a universal drug resistance site also found in other kinases including Src's close relative Abl. Based on our DMS study, most V326 variants confer resistance to inhibitors irrespective of their

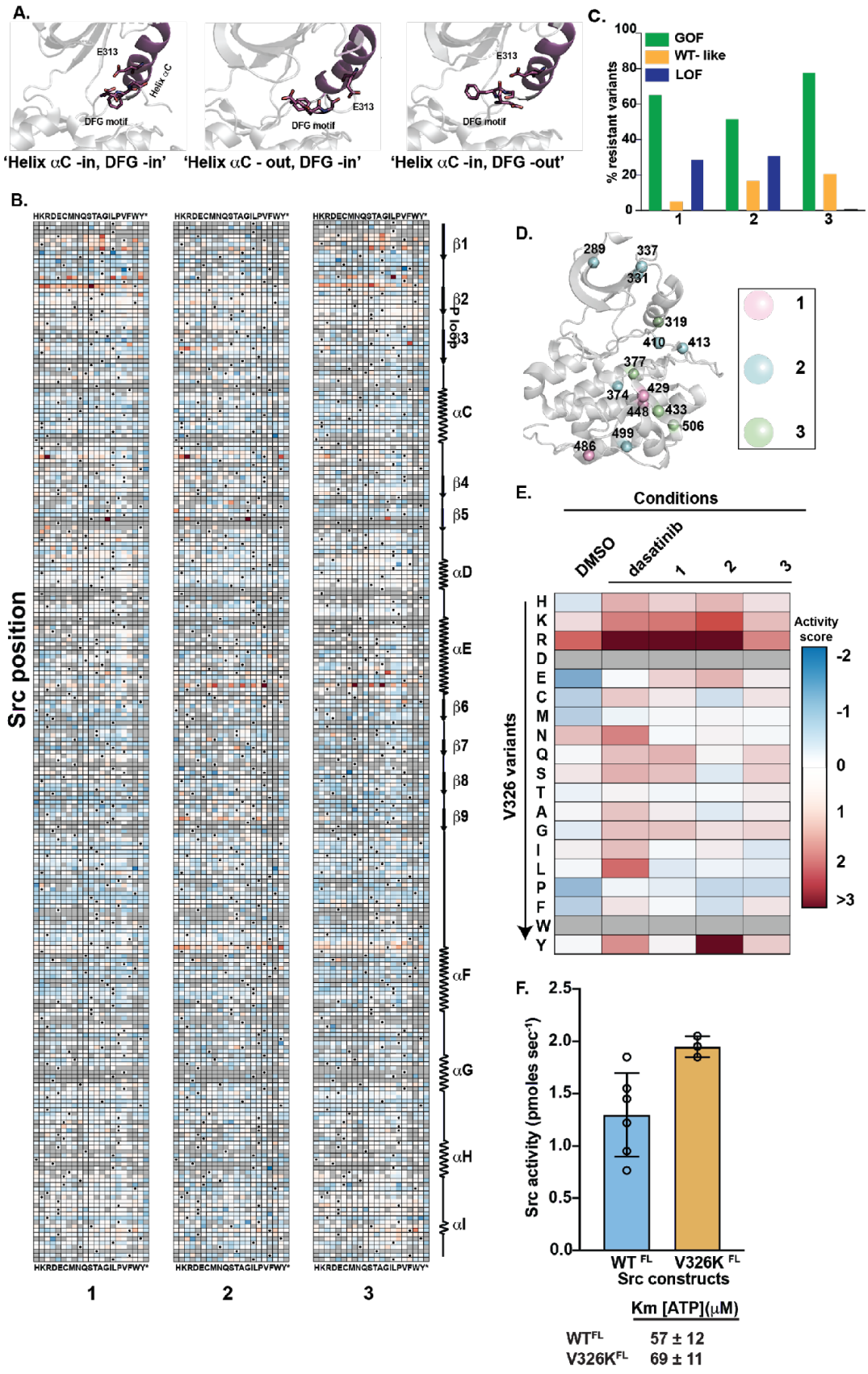
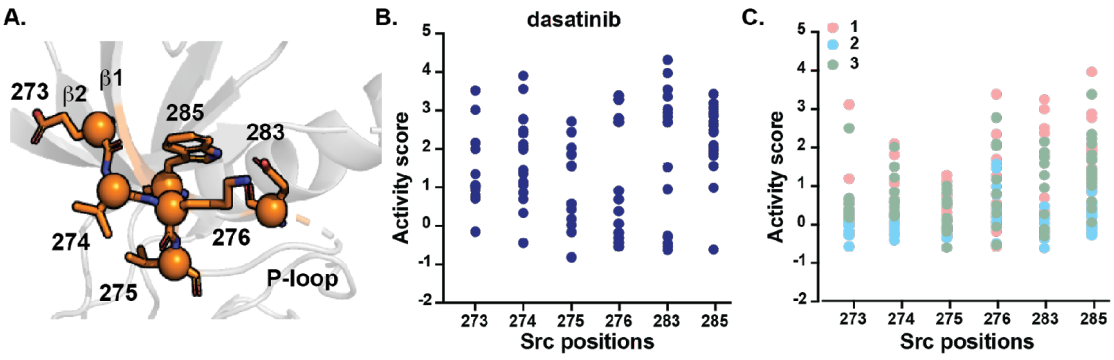

A.
E313
Helix αC
DFG motif
E313
DFG motif
E313
DFG motif
'Helix αC -in, DFG -in'
'Helix αC - out, DFG -in'
'Helix αC -in, DFG -out'
C.
GOF
WT- like
LOF
% resistant variants
100
80
60
40
20
0
1
2
3
B.
HKRDECMNQSTAGILPVFWY*
Src position
β1
β2
β3
αC
β4
β5
αD
αE
β6
β7
β8
β9
αF
αG
αH
αI
1
2
3
D.
289
337
331
319
410
413
377
429
374
448
433
506
499
486
1
2
3
E.
Conditions
DMSO
dasatinib
1
2
3
V326 variants
H
K
R
D
E
C
M
N
Q
S
T
A
G
I
L
P
F
W
Y
Activity score
-2
-1
0
1
2
>3
F.
Src activity (pmoles sec-1)
2.5
2.0
1.5
1.0
0.5
0.0
WT FL
V326K FL
Src constructs
Km [ATP](μM)
WTFL 57 ± 12
V326KFL 69 ± 11

Figure 2-6: Determination of resistant hotspots.

A. Src ATP-binding site conformations stabilized by different conformation-selective inhibitors. (left) the 'active' conformation with helix αC-in and DFG-in (PDB id: 3G5D). (center) the inactive conformation with helix αC-out and DFG-in (PDB id: 4YBK). (right) the inactive conformation with helix αC-in and DFG-out (PDB id: 4YBJ). **B.** Sequence-activity score map for all residues in Src's catalytic domain in the presence of inhibitors **1**, **2** or **3**. Nonsense mutants were excluded from the average score. Black dots in the map represent the WT amino acid, and gray tiles indicate missing data. Red indicates higher activity while blue indicates lower activity. **C.** Percentage of resistance mutants that are intrinsically gain-of-function (GOF), WT-like, or loss-of-function (LOF) with respect to inhibitor. Classifications were derived previously.[12] **D.** Src crystal (PDB id: 3G5D) structure showing positions in Src where inhibitor-specific resistant variants appear. **E.** Activity scores for every substitution in V326 in the absence of any inhibitor (DMSO) or the presence of dasatinib **1**, **2** or **3**. Gray tiles indicate missing data, red indicates higher activity in presence of inhibitor and blue indicate lower activity in presence of inhibitor. **F.** (top) Phosphotransferase activity of purified full-length Src^{FL} with either the WT or V326K sequence, (n = 3-6). (bottom) Michaelis-Menten constant (Km [ATP]) for Src^{FL} WT and V326K constructs (n= 3, mean ± SEM).

binding modes (**Figure 2-6 E**). This is in contrast to other resistant hotspots uncovered in our study where resistant sites are also primarily hyperactivating. Based on our comprehensive resistance data, we speculate that substitution of the wild-type valine residue to other residues, particularly positively charged lysine or arginine, results in repulsive interactions with most inhibitor but does not affect



D.

Compound	Ki (nM) WT FL	W285T FL	WT CD	W285T CD
1	< 1.0	5.2 ± 0.5	2.6 ± 0.1	4.7 ± 0.9
2	< 1.0	7.6 ± 0.6	2.5 ± 0.5	4.4 ± 1.1
4	< 1.0	11 ± 1	2.9 ±0. 6	3.2 ± 0.3
5	4.3 ± 0.5	9.2 ± 0.8	3.5 ± 1.1	3.8 ± 0.4

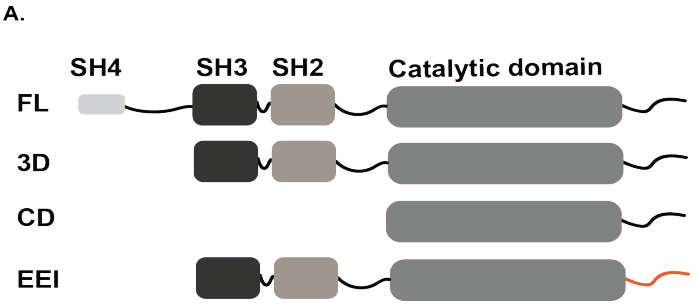
E.

Compound	Ki (nM) WT 3D	E283D 3D	E283D CD
1	5.2 ± 0.2	< 1.0	2.3 ± 0.5
2	3.7 ± 0.1	< 1.0	< 1.0
4	< 1.0	< 1.0	< 1.0
5	3.5 ± 0.7	< 1.0	3.2 ± 0.8

Figure 2-7. β1/2 hotspot is an unexpected drug resistant hotspot.

A. Crystal structure of Src (PDB id: 1Y57) showing every β1/2 hotspot residue in orange. **B.** Activity scores for every substitution at each position in the β1/2 hotspot in presence of dasatinib. **C.** Activity scores for every

substitution at each position in the β1/2 hotspot in presence of conformation-selective inhibitors **1**, **2** or **3**. **D.** Ki values (n = 3, mean ± SEM), determined in the presence of 1 mM ATP, for **1**, **2**, **4** and **5** against Src WT^{FL}, $W285T^{FL}$, Src WT^{CD} and $W285T^{CD}$. **E.** Ki values (n = 3, mean ± SEM), determined in the presence of 1 mM ATP, for **1**, **2**, **4** and **5** against Src WT^{3D}, $E283D^{FL}$ and $E283D^{CD}$.



B.

	Km [ATP](μM)
WT^{FL}	57 ± 12
$W285T^{FL}$	53 ± 9
$E283M^{FL}$	51 ± 1
WT^{3D}	85 ± 6
$W285T^{3D}$	62 ± 7
$E283D^{3D}$	90 ± 10
WT^{CD}	55 ± 5
$W285T^{CD}$	35 ± 7
$E283M^{CD}$	28 ± 2
$E283D^{CD}$	72 ± 12

C.

	Ki (nM)			
Compound	WT^{FL}	$E283M^{FL}$	WT^{CD}	$E283M^{CD}$
1	< 1.0	6.2 ± 0.5	2.6 ± 0.1	2.9 ± 0.5
2	< 1.0	9.6 ± 0.8	2.5 ± 0.5	2.4 ± 1.1
4	< 1.0	13 ± 3	2.9 ±0. 6	2.8 ± 0.7
5	4.3 ± 0.5	12 ± 1	3.5 ± 1.1	4.1 ± 0.9

Figure 2-8. *In vitro* Km and Ki determination of β1/2 hotspot resistance mutants.
A. Src constructs used for biochemical characterization of the β1/2 hotspot. **B.** Michaelis-Menten constant (Km [ATP]) of all Src constructs, (n= 3, mean ± SEM). **C.** Ki values (n= 3, mean ± SEM), determined in the presence of 1mM ATP, for **1**, **2**, **4** and **5** against Src WT^{FL}, $E283M^{FL}$, WT^{CD} and $E283M^{CD}$.

ATP-binding. We determined the Ki values of inhibitors **1**, **2** and DFG-out-stabilizing inhibitor **4** (**Figure 2-5 F,2-6 F**) against purified Src^{FL} WT and V326K. Consistent with our data in yeast, all three inhibitors potently inhibited WT Src^{FL} but had reduced potency against the V326K mutant of Src^{FL}. The affinity of dasatinib for Src is in the low picomolar range and, therefore, we could not determine its Ki value.

An unexpected drug resistance hotspot. A major hotspot uncovered by our DMS study was in the β1/2 stands, hereafter referred to as the β1/2 hotspot, where 6 positions shared the biggest clusters of inhibitor resistance mutations (**Figure 2-7 A**). Each of the 6 positions had 8 or more resistant mutations conferring resistance to dasatinib (**Figure 2-7 B**) and all of these positions are shared

amongst the conformation-selective inhibitor panel (**Figure 2-7 C**), with the exception of inhibitor **2**. The β1/2 hotspot is located at the beginning of Src's catalytic domain and forms the base of the phosphate-binding loop (commonly referred to as the P-loop). The P-loop is a structurally important feature in kinases that coordinates the β and γ phosphates of ATP, thereby aiding in catalysis. In BCR-ABL, multiple imatinib-resistant mutations occur at the P-loop. We were intrigued to see that in Src it is in the ordered β1 and β2 strands and not the dynamic P-loop that inhibitor resistance arises. Interestingly, residues L276 and E283, which are located at either ends of the P-loop, are conserved and sites of drug resistance in Abl. Position 285, which is part of the ordered β2 strand, was the most resistant position in our dataset. Almost every amino acid substitution at W285 was resistant to dasatinib, **1** and **3**, and at least 3 resistant variants were identified for inhibitor **2**.

We next tested whether W285 variants are drug resistant in *vitro* by performing activity assays with purified Src constructs. Consistent with our yeast data, Src^{FL} W285T was at least 5-fold less sensitive to all four inhibitors than WT (**Figure 2-7 D, 2-8 A** and **2-8 B**). Src has a regulatory apparatus that allosterically modulates its kinase activity. We wondered whether Src's regulatory apparatus plays a role in drug resistance. Therefore, we purified a construct consisting of only Src's catalytic domain (Src^{CD}) and determined Ki values for WT and W285T variants. Surprisingly, W285T and WT Src^{CD} have similar sensitivities towards all four inhibitors. WT Src^{CD} was in general less sensitive towards all four inhibitors compared to Src^{FL}, which explains why the W285T Src^{CD} variant did not show resistance. Similar drug resistance was observed for a second β1/2 hotspot variant–E283M (**Figure 2-8 C**). Like W285T, Src^{FL} E283M was at least ~5-fold less sensitive than Src^{FL} WT Src against inhibitors **1**, **2**, and **4**. We saw no significant difference in inhibitor sensitivity between WT and the E283M mutant in the context of Src^{CD}.

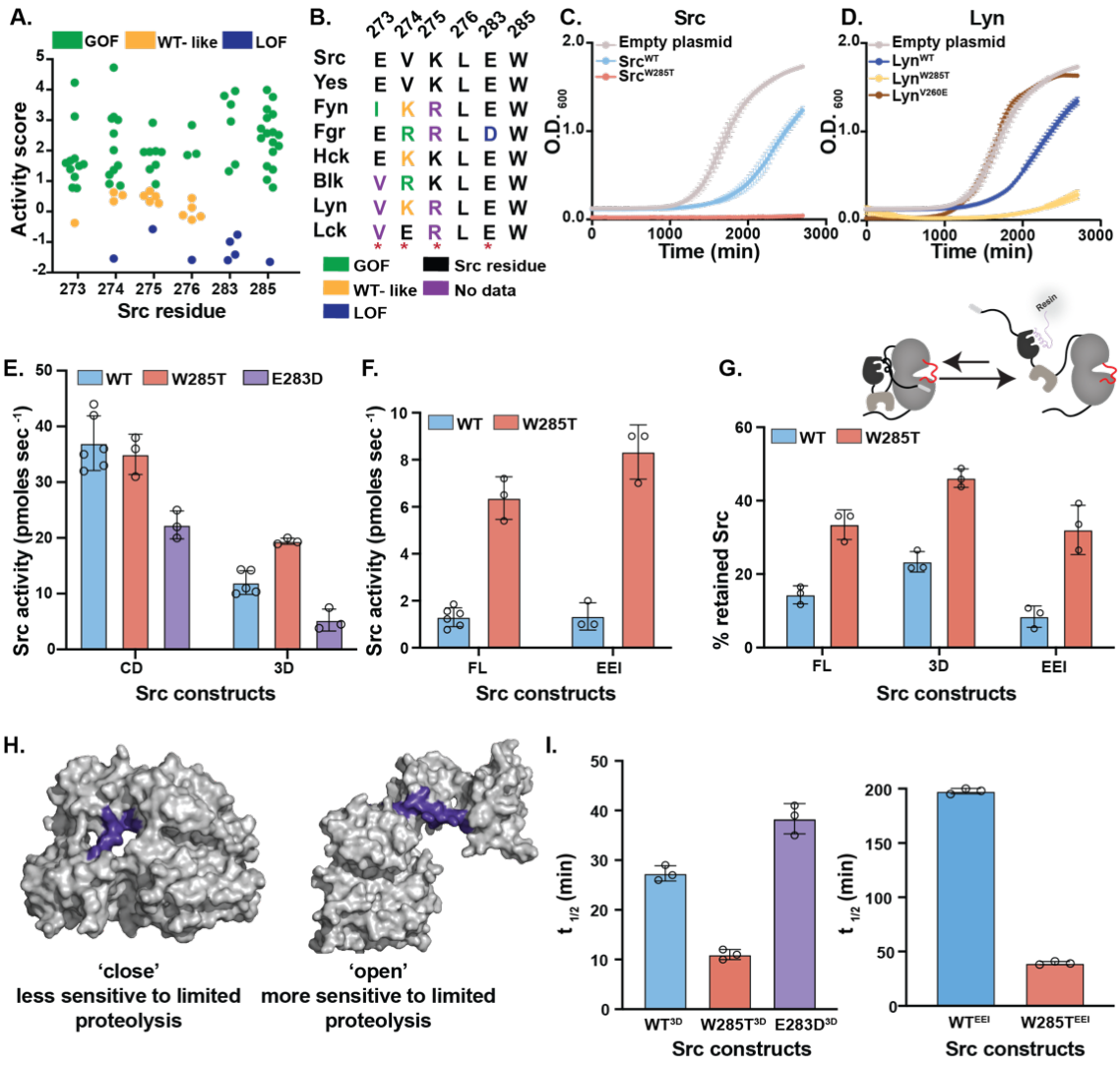


Figure 2-9. Characterization of the β1/2 hotspot.

A. Activity scores for every substitution at each position of the β1/2 hotspot in the absence of inhibitors. **B.** Sequence alignment of the β1/2 hotspot residues for Src family kinases (SFKs). Colors indicate classification of the analogous mutation from the Src DMS as reported in (Ahler et. al., Mol. Cell, 2019). **C.** Yeast growth curve over time comparing empty plasmid, Src^{WT} and Src^{W285T} (n=3-6). **D.** Yeast growth curve over time comparing empty plasmid, Lyn^{WT}, $Lyn^{W285T}$and Lyn^{V273E} (n=3-6). **E.** Phosphotransferase activity of purified Src^{CD} and Src^{3D} with either the WT, W285T or E283D sequence (n= 3-6). **F.** Phosphotransferase activity of purified Src^{FL} and Src^{EEI} with either the WT or W285T sequence (n= 3-6). **G.** (top) Schematic for the SH3 domain pulldown assay. To detect global conformation, Src is incubated with an immobilized SH3 domain ligand. Closed, SH3 domain-engaged Src is unable to bind to the resin, whereas open, SH3 domain-disengaged Src binds. After washing, retained Src is eluted and quantified by western blot. (bottom) Quantification of percent retained Src with purified Src WT^{FL}, $W285T^{FL}$, WT^{3D}, $W285T^{3D}$, WT^{EEI}, and $W285T^{EEI}$. **H.** Crystal structure of 3-domain Src (Src^{3D}) in the 'closed' conformation (PDB id: 2SRC) and global 'open' conformation (PDB id: 1Y57). The open conformation is more sensitive to proteolysis by the metallo-protease thermolysin at the SH2-CD linker (violet). **I.** Quantification of decay kinetics by thermolysin with purified Src^{3D} (top) Src WT^{3D}, $W285T^{3D}$, E283D, WT^{EEI} and $W285T^{EEI}$.

Previously, Barouch-Bentov et. al. reported that a catalytically important salt bridge forms between E283 and K275, which is located at either end of the P-loop. [16] This salt-bridge was speculated to limit the conformational flexibility of the P-loop, thus, influencing ATP coordination as well as inhibition by ATP-competitive inhibitors. The mutant E283D is deactivating according to our DMS and we speculated that it may enhance the salt bridge with K275. We reasoned that if the hyperactivating W285T mutant is resistant to inhibitors, then the deactivating E283D mutant should be sensitized to inhibition. To test this, we purified WT and E283D variants of Src^{3D} and determined Ki values. The results (**Figure 2-7 E, 2-8 B**) corroborated our prediction, as Src^{3D} E283D was highly sensitized to all four inhibitors. Src^{CD} E283D also appeared to be sensitized to inhibitors **2** and **4**. The observed increases in activity scores for β1/2 hotspot mutants suggests that this region is more than just a site for inhibitor resistance. We speculate that the β1/2 hotspot is allosterically coupled to the regulatory apparatus. We set out to test this hypothesis.

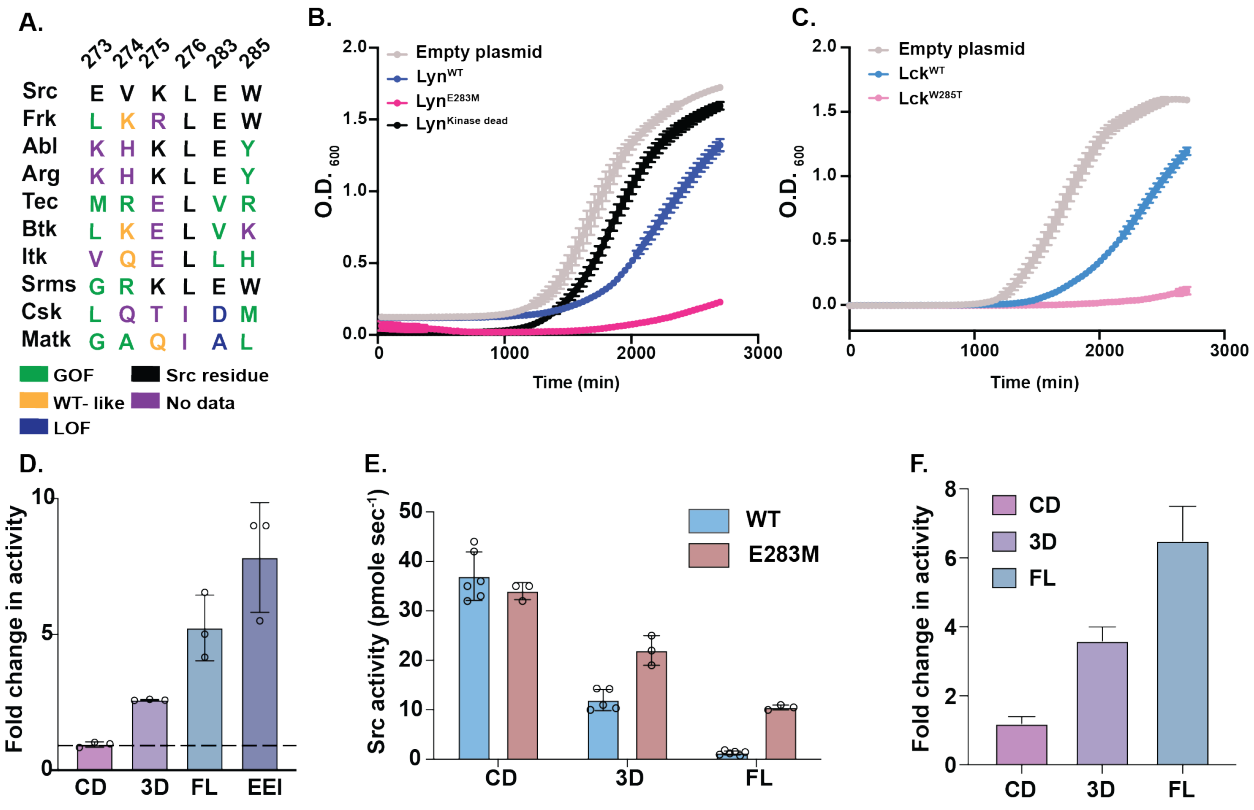


Figure 2-10: β1/2 hotspot regulatory in other SFKs.

A. Sequence alignment of Src β1/2 hotspot to related tyrosine kinases. Color indicate classification of analogous mutation from Src DMS as reported in (Mol. Cell. 2019). **B.** Yeast growth curve over time comparing empty plasmid, Lyn^{WT}, Lyn^{E283M}, and $Lyn^{Kinase\ dead}$ mutants. **C.** Yeast growth curve over time comparing empty plasmid, Lck^{WT}, and Lck^{W285T}. **D.** Fold change in Src activity between WT and W285T for catalytic domain (CD), 3-domain (3D), full-length (FL) and autoinhibited (EEI) construct. **E.** Phosphotransferase activity of purified Src^{CD} and Src^{3D} with either WT, W285T or E283D sequence (n= 3-6). **F.** Fold change in Src activity between WT and E283M for catalytic domain (CD), 3-domain (3D), full-length (FL) and autoinhibited (EEI) construct.

Characterization of the β1/2 hotspot. Previously, residues from the β1/2 hotspot have been defined as forming a regulatory interface[12] based on hierarchical clustering of hyperactivating mutations. We looked at activity scores for every substitution corresponding to the residues that make the β1/2 hotspot (**Figure 2-9 A**) and found that every residue in this region has 3 or more gain of function mutations. Residues in the β1/2 hotspot are mainly conserved in SFKs (**Figure 2-9 B**). We were curious how β1/2 hotspot mutations affect the activities of the SFKs Lyn and Lck. Therefore, we determined the effects of these mutations on Lyn and Lck in the yeast growth assay. WT Lyn and Lck

inhibit the growth of yeast like WT Src (**Figure 2-9, C-D 2-10 B-C**). As in Src, introducing the W285T mutation into Lyn and Lck enhances their abilities to inhibit yeast growth (**Figure 2-9, C-D 2-10 B-C**). Therefore, the regulatory mechanism of the β1/2 hotspot appears to be conserved in the SFKs.

We next tested the phosphotransferase activities of the β1/2 hotspot mutants W285T (**Figure 2-9 E**) and E283M (**Figure 2-10 E**) *in vitro*. We found that these mutations had a minimal effect on the phosphotransferase activity of Src^{CD} (**Figure 2-10 D**). However, the W285T and E283M mutations increased the phosphotransferase activity of Src^{3D} (**Figure 2-8 A**), suggesting that β1/2 hotspot region allosterically communicates with the regulatory SH2 and SH3 domains. We found that the fold increase in phosphotransferase activity conferred by the W285T (**Figure 2-9 F** and **2-10 D**) and E283M (**Figure 2-10 F**) mutations was even greater in the context of Src^{FL} than in Src^{3D}. Both mutations showed a comparable increase in phosphotransferase activity in the context of a 3-domain Src construct (Src^{EEI}) that contains mutations that enhance engagement between the SH2 domain and the C-terminal tail (**Figure 2-9 F** and **2-10 D**). Together, these results suggest that the β1/2 hotspot is allosterically coupled to the regulatory SH2 and SH3 domain apparatus but does not depend on the N-terminus of Src.

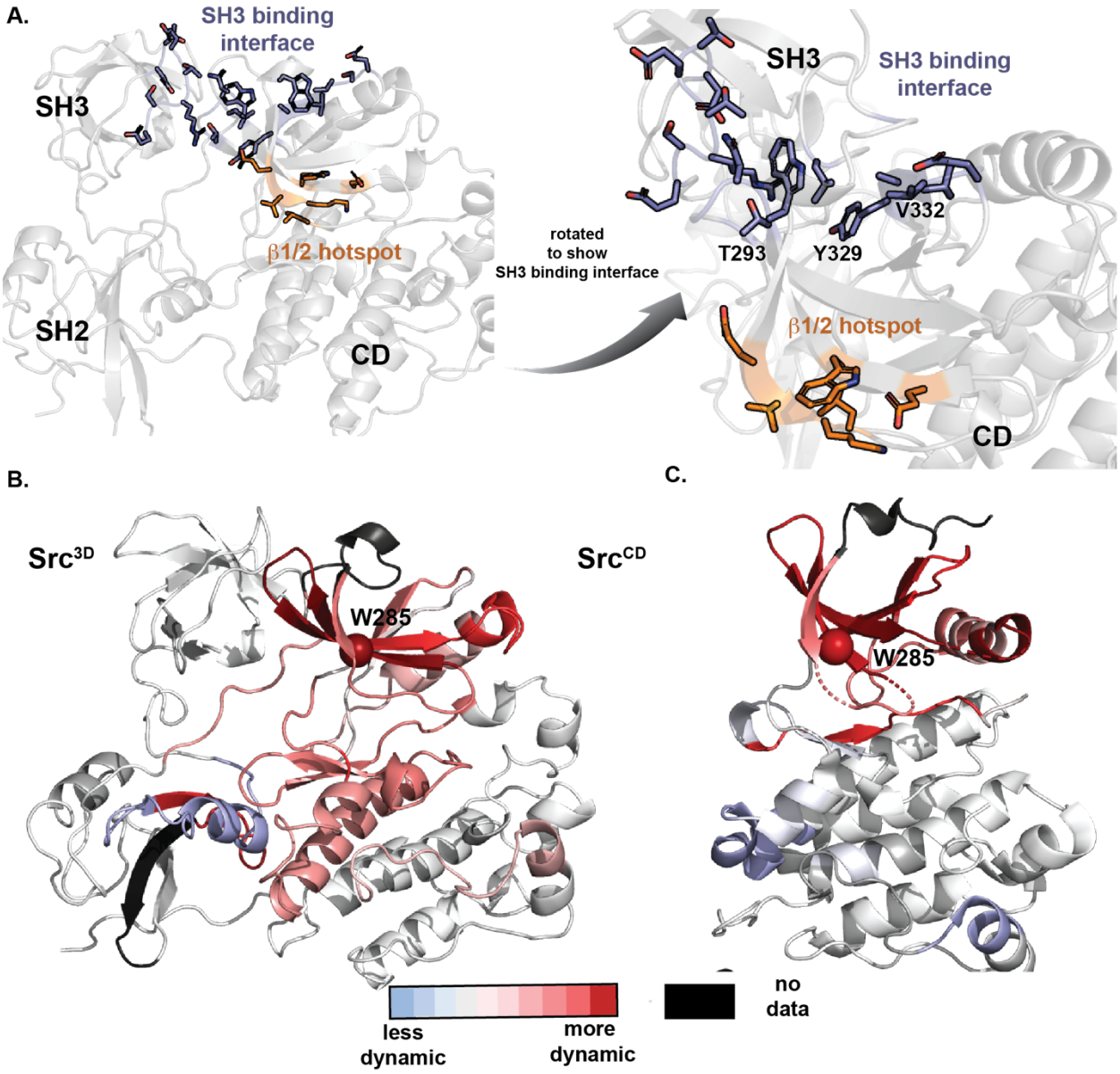


Figure 2-11. Structural characterization of the β1/2 hotspot.
A. (Right) Structural detail of residues that comprise SH3-CD regulatory interface (blue) and β1/2 hotspot (orange). (Left) the same structure rotated to focus on the three nearest residues on the SH3-CD interface. **B.** HDX-MS analysis with purified Src^{3D} comparing deuteration pattern of WT and W285T is plotted on Src crystal structure (PDB id: 2SRC, n = 3-4), sphere represents W285. **C.** HDX-MS analysis with purified Src^{CD} comparing deuteration pattern of WT and W285T is plotted on Src crystal structure (PDB id: 3G5D, n = 3-4), sphere represents W285.

The W285T mutation promotes an 'open' global conformation of Src. We investigated the effects of β1/2 hotspot mutations on the SH2/SH3 domain regulatory apparatus in two separate biochemical assays. First, we used an immobilized, SNAP-tagged SH3 domain ligand as a bait to assess

intermolecular SH3 domain accessibility (**Figure 2-9 G**). If β1/2 hotspot residues communicate with the SH2/SH3 domain regulatory apparatus, mutations at the β1/2 hotspot should result in abrogation of intramolecular SH3-CD interactions, resulting in higher intermolecular accessibility of the SH3 domain to the immobilized SH3 domain ligand. Consistent with β1/2 hotspot mutations disrupting the regulatory apparatus, we observed a 2-fold increase in SH3 domain ligand pulldown for the W285T mutant compared to WT Src^{FL}. A similar fold-increase in the SH3 domain ligand pulldown assay was observed in the context of Src^{3D}, which was further enhanced in the Src^{EEI} construct.

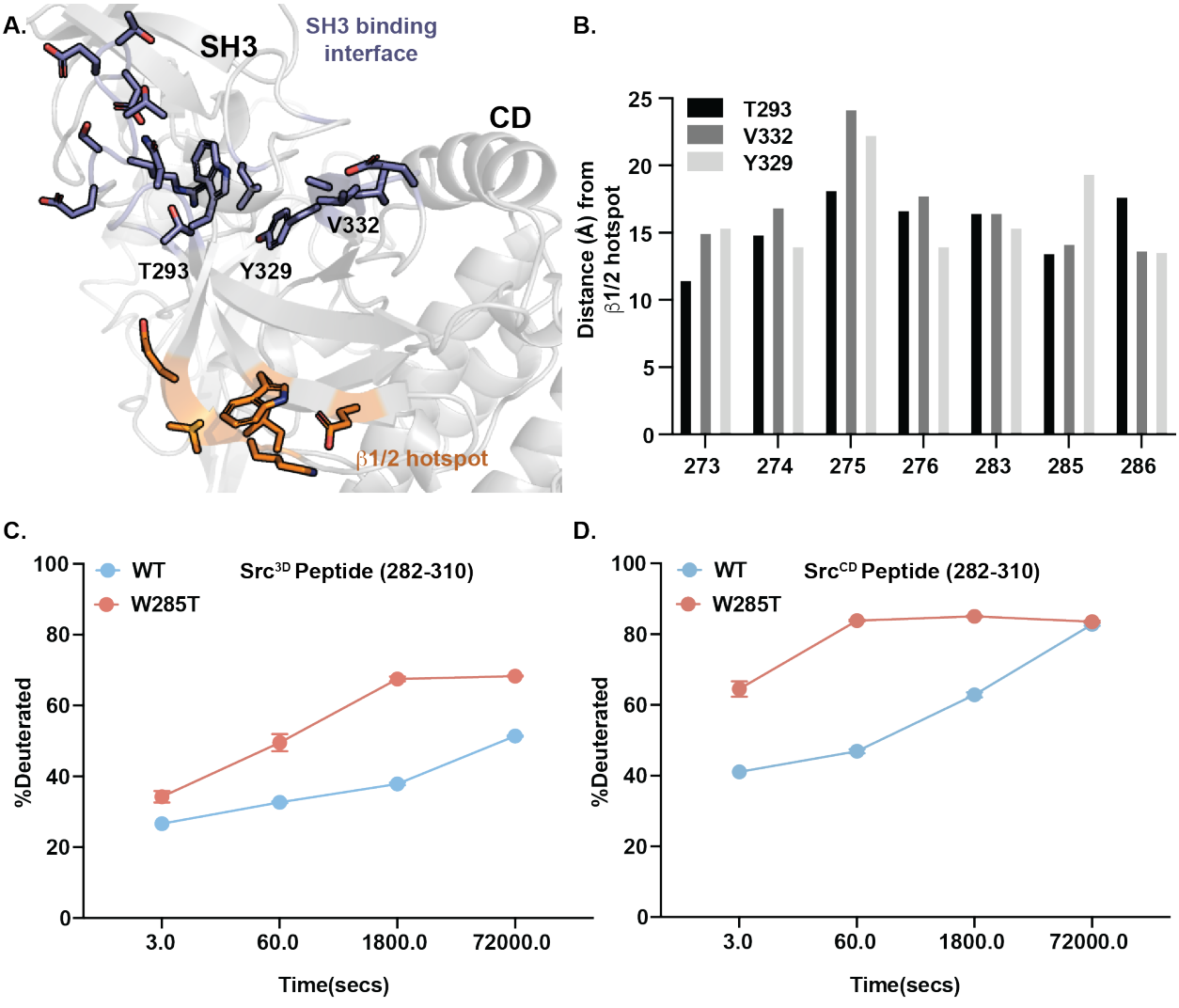


Figure 2-12. β1/2 hotspot and SH3 binding interface.

A. (Right) Structural detail of residues that comprise SH3-CD regulatory interface (blue) and β1/2 hotspot (orange). (Left) the same structure rotated to focus on the three nearest residues on the SH3-CD interface. **B.**

HDX-MS analysis with purified Src^{3D} comparing deuteration pattern of WT and W285T is plotted on Src crystal structure (PDB id: 2SRC, n = 3-4), sphere represents W285. **C.** HDX-MS analysis with purified Src^{CD} comparing deuteration pattern of WT and W285T is plotted on Src crystal structure (PDB id: 3G5D, n = 3-4), sphere represents W285.

Previously Agius et. al. has demonstrated that the metalloprotease thermolysin can selectively cleave the flexible SH2-CD linker [17] (shown in purple in **Figure 2-9 H**) and that the rate of cleavage correlates to Src's global conformational state (**Figure 2-9 H**). We used this assay to determine the half-lives ($t_{1/2}$) of proteolytic cleavage by thermolysin for the WT and W285T variants of Src^{3D}. The calculated $t_{1/2}$ for WT Src^{3D} (28 ± 4 min) was in agreement with the previously reported value. W285T Src^{3D}, on the other hand, had a $t_{1/2}$ of 11 min, which is consistent with the W285T mutation promoting a more 'open' global conformational state. Furthermore, thermolysin's rate of proteolytic cleavage of the deactivating E283D mutant was 2-fold slower than for WT Src^{3D}. The proteolytic cleavage of WT Src^{EEI} was slower than for WT Src^{3D}, with a $t_{1/2}$ of ~200 min. Introduction of the W285T mutation into the Src^{EEI} construct increased the rate of proteolytic cleavage by 5-fold. These results are consistent with mutations in the β1/2 hotspot leading to allosteric disruption of the SH2/SH3 regulatory apparatus despite their spatial separation from any regulatory interfaces.

β1/2 hotspot residues regulate Src activity by modulating P-loop flexibility. To better understand how β1/2 hotspot residues are able to influence the regulatory SH2/SH3 domain apparatus, we mapped Src's SH3 domain-binding interface onto the inactive 'closed' structure of Src (**Figure 2-11 A**). This structure shows that the three residues (T293, Y329 and V332) in the SH3 domain-binding interface that are in closest proximity to the β1/2 hotspot are over 10 angstroms away in distance (**Figure 2-12 B**). Therefore, there does not appear to be any direct interactions between these regions.

In order to better understand the communication between the β1/2 hotspot and Src's regulatory apparatus, we performed Hydrogen-Deuterium exchange Mass Spectrometry (HDX-MS) on WT and W285T variants of purified Src^{3D}. Deuterium exchange was initiated by incubating each Src variant in buffered D_2O (containing 90% deuterated buffer, pH 7.8). In order to quench the deuterium exchange, the pH of the buffer was dropped to 2.5 at various time points. Samples were then processed using standard protocols (see "Methods"). After filtering out peptides showing weak signal or partial overlap, 32 unique peptic peptides remained, which covered 85% of Src^{3D}'s sequence. In determining which peptic peptides demonstrate a difference between the WT and W285T Src variants, we excluded peptides with overlapping sequences that demonstrate different levels of deuteration in order to reduce the possibility of interpreting noise. This yielded a relatively conservative set of peptides for further analysis. A summary of the deuterium exchange kinetics for WT and W285T Src^{3D} are represented on the crystal structure shown in **Figure 2-11 B**. There are numerous differences in deuterium exchange kinetics between WT and W285T Src^{3D}, including evidence for a disengaged regulatory domain apparatus, which is consistent with with our biochemical analyses. A striking feature of the HDX-MS results was the extent to with the W285T mutation increased the rate of deuteration for all the peptides in the N-lobe of the kinase. Nearly all identified peptides in this region appeared highly dynamic in the W285T mutant compared to WT. Interestingly, these differences in deuteration kinetics in the N-terminal lobe were also observed in Src^{CD} (**Figure 2-11 C**). We hypothesize that β1/2 hotspot mutations affect Src's catalytic activity by influencing the conformation of the phosphate-binding loop (P-loop) and helix αC. Previously, NMR studies and MD simulations have revealed similar pattern of allosteric communication between the N-lobe and C-lobe of the kinase that involve the P-loop and helix αC.[18, 19]

2.4 Conclusion

Kinase drug resistance is a major challenge in the field of drug discovery. Our large-scale mutagenesis data enabled us to identify sequence-function relationships in conferring inhibitor resistance in the model tyrosine protein kinase Src. Upon testing a Src variant library against a panel of different conformation-selective inhibitors, we were able to identify both unique and common sites of drug resistance. Interestingly, our data revealed that there are multiple factors contributing to the emergence of drug resistance and that drug resistant mutations occur throughout kinase catalytic domain. Mutation of residues that line the kinase active site can confer resistance by disrupting drug binding through perturbation of the ATP-binding pocket or by enhancing basal phosphotransferase activity. We found that only a subset of resistance mutations lines the ATP-binding pocket and that many mutations appear to confer resistance by other mechanisms. Many drug-resistant mutants that are physically separated from the ATP-binding pocket confer resistance through hyperactivation of phosphotransferase activity. The β1/2 hotspot is one such site of resistance, where a majority of mutations lead to hyperactivation of the kinase. However, β1/2 hotspot mutations also appear to diminish drug binding by an unknown mechanism. Further analysis revealed that the β1/2 hotspot is allosterically coupled to the regulatory apparatus of Src and appears to communicate through a network that includes the P-loop and helix αC.

2.5 Methods

***S. cerevisiae* genetics and cell culture.** Yeast experiments were performed in BY4741 Green Monster (a generous gift from Dr. Fritz Roth, DMS experiments) or BY4741 (MATa his3D1 leu2D0 met15D0 ura3D0, all other experiments). All Src, Fyn and Lck constructs were cloned into the p415GAL1 plasmid and using standard LiAc transformation protocols and plated on C-Leu media to

select for successful transformants. All growth experiments were performed in C-Leu media to maintain plasmid.

Cloning, protein purification and expression. All SFK mutants were synthesized using either QuikChange Site Directed Mutagenesis (Agilent) or IVA cloning following standard protocols. Mutations were verified by Sanger sequencing. All subcloning was achieved using Gibson assembly or directional cloning following standard protocols and validated by Sanger sequencing.

For *in vitro* biochemical assays, full-length Src (WT/ V326K/ W285T/ E283M^{FL}, residues 2-536), 3-domain Src (WT/ W285T/ E283M/ E283D^{3D}, residues 87-536), catalytic domain Src (WT/ W285T/ E283M/ E283D^{CD}, residues 261-536) and autoinhibited Src (WT/W285T^{EEI},residues 87-536; mutations Q531E, P532E and G533I) were cloned into the bacterial expression plasmid pMCSG7 as an N-terminal His6-SUMO tagged constructs. Src constructs were co-transformed in *E.coli* with YopH/GroEL and plated on triple selective plates (Ampicilin/Chloramphenicol/Streptomycin). A single colony was picked and grown in an overnight culture of 15 mL of Terrific broth containing all three antibiotics. A 1 L culture was then inoculated with the starter culture, grown to an O.D.$_{600}$ of 1.1, the temperature was then dropped to 18^{o}C and protein expression was induced with 0.4 M IPTG. Protein was expressed overnight at 18^{o}C. Ni-NTA was used to purify His$_6$-SUMO-Src after lysing cells in lysis buffer (50 mM HEPES, pH 8.0, 300 mM NaCl, 1 mM PMSF, 0.1% Triton-X, 20 mM imidazole) and eluted using purification buffer (50 mM HEPES, pH 8.0, 300 mM NaCl, 1 mM PMSF, 0.1% Triton-X, 10% glycerol, 0.2% BME, 100 mM imidazole). A 2 h dialysis in dialysis buffer (50 mM HEPES, pH 8.0, 150 mM NaCl, 1 mM DTT, 10% glycerol) was performed prior to adding the SUMO protease His6-ULP1 (1:25 protease:eluted protein, wt/wt) to the protein. The protein protease mixture was then transferred to fresh dialysis buffer and was allowed to cleave overnight at 4^{o}C. Following cleavage, a second Ni-NTA purification was carried out to remove any non-cleaved SFK

and His6-ULP1. Finally, an anion exchange column (Pierce, 90011) was used to remove YopH and GroEL to yield Src at > 95% purity.

Src phosphotransferase activity. Phosphotransferase activity of purified Src variants were measured using a self-reporting fluorescent SFK peptide (EEEIYGE-(DAP- Pyrene)-EA) (16464077) in an *in vitro* kinase assay. Briefly, 20 mL of purified full-length Src (FL,6.5 nM), 3-domain Src (3D, 5 nM), catalytic-domain Src (CD, 5.5 nM) or autoinhibited Src (EEI, 8 nM) constructs of each variant (WT, V326K, W285T,E238m and E283D) were diluted in kinase reaction buffer (76 mM HEPES, pH 7.5, 5 mM $MgCl_2$, 150 mM NaCl, 3.8 mM EGTA, 0.2 mg/mL BSA, 150 mM Sodium orthovanadate (Na_3VO_4)) and incubated with 4 mL 1 mM ATP at room temperature for 30 min. Next, 5 mL of 60 nmoles of SFK peptide was pipetted to respective wells and raw fluorescence units were measured immediately on an Envision fluorometer (Perkin Elmer) with an excitation wavelength of 344 nm and an emission wavelength of 405 nm in real time for the first 10 min at 15 s intervals. Calculation of kinase activity in terms of pmole $s^{-1}$of phosphorylated substrate is discussed below under "Calculation of Src activity."

Src Km [ATP] determination. Km [ATP] of purified Src variants were measured using the same assay described above. Briefly, 2-fold (8-data points) serial dilution of ATP starting at 1 mM, was incubated with 5 nM of Src (WT or mutant) in kinase reaction buffer (76 mM HEPES, pH 7.5, 5 mM $MgCl_2$, 150 mM NaCl, 3.8 mM EGTA, 0.2 mg/mL BSA, 150 mM Sodium orthovanadate (Na_3VO_4)) and 20 mM of Src fluorogenic peptide(EEEIYGE-(DAP- Pyrene)-EA). Raw fluorescence units were measured immediately on an Envision fluorometer (Perkin Elmer) with an excitation wavelength of 344 nm and an emission wavelength of 405 nm in real time for 90 min at 15 min interval. Calculation of kinase Km [ATP] is discussed below under "Calculation of Src activity."

Inhibitor IC_{50} and K_I determination. For all IC_{50} determination experiments, a kinase titration was performed as described above prior to inhibitor testing in order to ensure that the kinase concentration assayed was in the linear range. Inhibitors (initial concentration = 30 μM, 3-fold serial dilutions, 10 data points in triplicate) were assayed in triplicate against Src variants ([WT^{FL}] = 4 nM ; [WT^{3D}, $W285T^{FL}$,$W285T^{CD}$ and $V326K^{FL}$] = 5 nM; [$E283M^{FL}$ and $E283M^{CD}$] = 4.5 nM, and [$E283D^{3D}$ and $E283D^{CD}$] = 7.5 nM) in assay buffer (76 mM HEPES, pH 7.5, 5 mM $MgCl_2$, 150 mM NaCl, 3.8 mM EGTA, 0.2 mg/mL BSA, 150 mM Sodium orthovanadate (Na_3VO_4)). Briefly, kinase was incubated with inhibitors and 1 mM ATP for 30 min in a 384-black assay plate (Corning, #3573). 20 μM of Src fluorogenic peptide (Ac-EEEIYGE(Dap-Pyrene)-EA-NH_2) was then added to plate and incubated for 2 h.Raw fluorescence units were measured on Envision (Perkin Elmer) with excitation wavelength of 344 nm and emission wavelength of 405 nm. Data was analyzed using GraphPad Prism software, IC_{50} determination and Ki calculation has been discussed below under "Calculation of IC_{50}, K_I, Km."

SFK yeast growth assay and identification of resistant mutants. Src variant library[12] (previously described in Ahler et. al., 2019) was treated with various conformation selective inhibitor and resistance mutations were identified based on variant activity score. Upon inhibitor treatment, yeast harboring sensitive Src variants have their growth rescued, while those expressing resistant Src variants continue to grow poorly. Time points were sampled throughout growth, plasmids extracted, barcodes amplified and deeply sequenced on an Illumina NextSeq run as described above. We calculated the standard deviation of activity scores of all synonymous variants (for dasatinib STD= 0.327) identified in our dataset. Next we assigned all variants as 'resistant' that scoed 2x of this value (= 0.654).

SH3 domain pulldown assays. Generation of immobilized SH3 domain ligand: 20 mL of a 50% slurry of SNAP-capture pulldown resin was placed in a micro-centrifuge tube. The resin was washed (3x, 10 bead volumes) with pulldown buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM DTT

and 0.2 mg/mL BSA). 8 mM of SNAP tag–polyproline peptide fusion (VSLARRPLPPLP) was loaded onto the resin at a final volume of 50 mL per 10 mL of bead in pulldown buffer (Leonard et al., 2014). The resin was rotated at room temperature for 1 h and then washed (3x, 10 bead volumes) prior to performing pulldown assays.

SH3 pulldown assays performed with Src WT and mutants: Src (100 nM) in 50 mL of pulldown buffer was incubated with 5 mL of the immobilized SH3 domain ligand. The resin was rotated at room temperature for 1 h. Post incubation, the flow-through was collected, and the resin was washed (3x, 10 bead volumes). To elute the retained kinase, 50 mL of 1x SDS loading buffer was added and the beads were boiled at 95C for 10 min. All samples were separated via SDS–PAGE and visualized by western blotting with Src antibody (Cell Signaling, #2109). The loaded amount of kinase was diluted to one third of the original concentration using 1x loading dye before running it on the gel as "I" (Input) along with the kinase eluted ("E") off the resin. The scanned blots were quantified with ImageStudio Lite software and signal corresponding to input protein was scaled to original loaded kinase amount. The percentage of kinase retained on the resin (% retained Src) was then calculated on the basis of the loaded and eluted fraction based on curve fitting of immunoblot signal intensity to a Src titration.

Limited proteolysis of Src WT and mutant with thermolysin. Limited proteolysis of Src was modified from.[17] Briefly, 3-domain Src (WT, W285T, E283D) or autoinhibited Src (WT^{EEI} and $W285T^{EEI}$) was diluted to 1 mM in proteolysis buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.5 mM $CaCl_2$). Proteolysis was initiated by adding 3.8 mM Thermolysin (Promega, #V4001) stock solution to kinase at a final concentration of 60 nM. 20 μL of this mixture was then added to 10 μL of 50 mM EDTA to terminate proteolysis at various time points (0, 2, 4, 8, 16, 32, 64, 128, 256 min). The quenched samples were analyzed by SDS-PAGE (12 % Bis Tris gel in SDS running buffer) and

stained with SYPRO Ruby (ThermoFisher Scientific: catalog number S12000) according to manufacturer's protocol. Band intensities were analyzed by ImageStudioLite imaging software. Percent protein remaining was computed based on band intensity at 0 min and was plotted against time on GraphPad Prism 8.4.2. The curve was fit to an exponential decay equation using GraphPad Prism 8 software to obtain half-life of each Src variant.

HDX-MS of Src in presence of conformation selective inhibitors. 3-domain Src (WT or W285T) or catalytic domain Src (WT and W285T) was diluted to 0.2 mg/mL in protein dilution buffer (50 mM HEPES, pH 7.8, 150 mM NaCl, 1mM DTT, 5% glycerol) at room temperature for 30 min. 10 mL of this was then added to 90 mL of buffered D_2O (prepared 5 mL with 4.5 mL of D_2O and 0.5 mL of 10x protein dilution buffer and 0.2 mg/mL of peptide standard Glu-1-Fibrino peptide (CAS: 103213-49-6, Sigma)) to initiate deuteration at 22°C. Deuterium exchange was quenched after 3 s, 1 min, 30 min, and 20 h by adding the reaction to 100 mL of ice-cold quench buffer (0.2% formic acid, 8M Urea, 0.1% trifluoroacetic acid, final pH dropped to 2.5) in order to lock deuterium in place and unfold the protein. All time points were collected in triplicate. Samples were immediately frozen in a dry ice/ethanol bath and stored at −80°C. Undeuterated samples were prepared the same way except using buffered H_2O instead of D_2O.

Frozen samples were thawed on a 5°C block for 4 minutes prior to injection onto a loading loop. The loaded sample was passed over a custom packed pepsin column (Porcine pepsin immobilized on POROS 20-AL resin; 2.1 x 50 mm column) [20] kept at 12°C with a flow of 0.1% trifluoroacetic acid (TFA) and 2% acetonitrile (ACN) at 200 μL/min. Digested peptic fragments were trapped onto a Waters XSelect CSH C18 XP VanGuard Cartridge (2.1 x 5 mm, 2.5 μm). After 5 minutes of loading, digestion, and trapping, peptides were resolved on an analytical column (Waters BEH 1 x 100 mm, 1.7μm, 130Å) using a gradient of 3 % to 40 % solvent B for 9 minutes (A: 0.1 % FA, 0.025 % TFA, 2

% ACN; B) 0.1 % FA in ACN). The LC system was coupled to a Thermo Orbitrap performing full scans over the m/z range of 300 to 1500 at a resolution of 30,000. The MS source conditions were set to minimize loss. [21] Undeuterated samples were run prior to and at the end of all the LC-MS runs.

During the analytical separation step, a series of 250 μL injections were used to clean the pepsin column: 1) 0.1 % Fos-12 with 0.1 % TFA; 2) 2 M GndHCl in 0.1 % TFA; 3) 10 % acetic acid, 10 % acetonitrile, 5 % isopropanol (PMID 22993047 and 29299838). After each gradient the trapping column was washed with a series of 250 μL injections: 1) 10 % FA; 2) 30 % trifluoroethanol; 3) 80 % methanol; 4) 66 % isopropanol, 34 % ACN; 5) 80 % ACN. During the trap washes the analytical column was cleaned with three rapid gradients. [22]

Peptic peptides were identified from data-dependent acquisition (DDA) experiments on undeuterated samples by exact mass and tandem mass spectrometry (MS/MS) spectra using Protein Prospector [23] filtering with a score cutoff of 15. Mass shifts were determined using HD-Examiner v2 (Sierra Analytics). The Glu-1-Fibrino internal standard peptide (CAS: 103213-49-6, Sigma) peptide was checked in all samples to verify that back-exchange levels were consistent in all experiment.[24] We binned peptides based on their fractional deuteration levels. We assigned the deepest shade of red to the peptide GEVWMGTWNGTTRV which showed the maximal differential fractional deuteration level between for both 3-domain and catalytic domain Src. We then normalized color intensity from red to white for the rest of the peptides and mapped these to the crystal structure of 3-domain of catalytic domain Src.

Maleimide labeling and mass spectrometry. Purified 1 mM catalytic domain Src (WT or W285T) was diluted in mass spectrometry buffer (50 mM HEPES, pH 7.6, 150 mM NaCl, 5% glycerol and 0.02% (wt/vol) n-Dodecyl β-D-maltoside (DDM))[25] and treated with 100 mM N-ethyl maleimide

(NEM) in a LoBind 1.5 mL eppendorf tube at 25°C for 30 min. The NEM labeling reaction was quenched with 20 mM of DTT in NH_4HCO_3 solution. Protein was then precipitated using 0.02% deoxycholate and 10% trichloroacetic acid on ice for 10 min. The precipitated protein was pelleted by spinning at 10,000x rpm for 15 min. The pellet was dried with 10 mL acetone and re-suspended in peptide solubilization buffer (8 M urea, 200 mM Tris-HCl, pH 8.0, 2.4 mM iodoacetamide, 0.001% DDM) by vortexing briefly. The mixture was incubated in the dark for 30 min. Trypsin digestion solution (0.5 mg/mL of trypsin in 1 mM $CaCl_2$, 200 mM Tris-HCl, pH 8.0) was added and the protein was digested overnight at 37°C. Peptide was desalted using C18 ZipTips (Milipore) and each sample was run on the Finnigan LTQ Ion trap. [M+2H] peptide masses for both NEM and iodoacetamide modified Cysteine containing peptides was analyzed using Xcalibur MaxQuant software. Experiment was repeated in triplicate.

Quantification and statistical analysis. For all statistical tests (unless otherwise noted), a two-tailed Student's t test was used to compare means between two samples. A one-way ANOVA with post hoc Tukey's HSD test was used to compare means between more than two samples. Significance is denoted as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Statistical tests were performed in R or GraphPad Prism.

Calculation of K_I, IC_{50} and Km. IC_{50} values were calculated in GraphPad Prism using the "One-Site Fit log IC50." Km[ATP] values were determined using GraphPad Prism using "Plot Michaelis-Menten" option. Ki values for all SFKs were calculated using the Cheng-Prusoff equation at 1 mM ATP and calculated Km[ATP] .

Classification of Src variants from activity scores. "Activity Scores" are the inverse of the Enrich2 output score function. To define "gain of function," "functionally-neutral," and "loss of function"

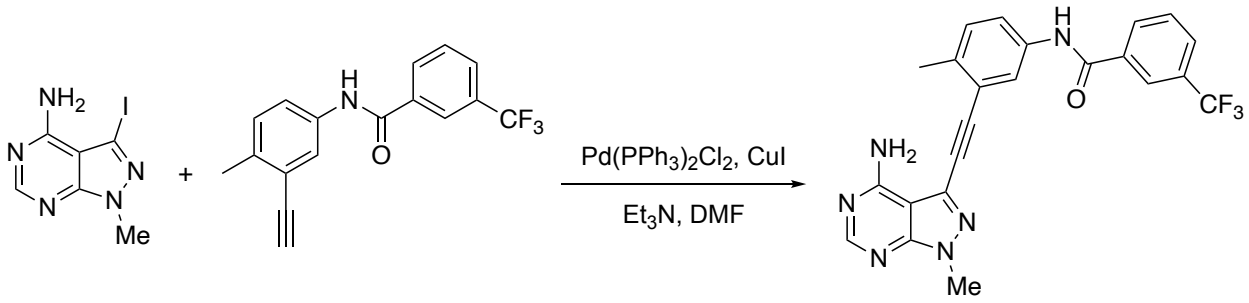
activity classifications of Src variants we calculated ± 2 standard deviations of the mean activity score of synonymous variants. All nonsynonymous variants with scores greater than the upper bound were classified as "gain of function," all those with scores less than the lower bound were classified as "loss of function," and all others were classified as "functionally-neutral."

Calculation of Src phosphotransferase activity. Slopes were calculated from the linear portion of *in vitro* kinase enzyme assay to first obtain raw fluorescence count per sec. This was then divided by fluorescence change per picomoles of phosphorylated substrate (obtained from the slope of a standard curve of raw fluorescence versus phosphorylated substrate) to calculate phosphorylated substrate produced in pmoles.

General Synthetic Procedures

All chemicals purchased from commercial suppliers were used without further purification unless otherwise stated. Reactions were monitored with thin-layer chromatography (TLC) using silica gel 60 F254 coated glass plates (EM Sciences). Compound purification was performed with an IntelliFlash 280 automated flash chromatography system using pre-packed Varian SuperFlash silica gel columns (Hexane/EtOAc or CH_2Cl_2/MeOH gradient solvent). A Varian Dynamax Microsorb 100-5 C18 column (250 mm x 21.4 mm), eluting with H_2O/CH_3CN or H_2O/MeOH gradient solvent (+0.05% TFA), was used for preparatory HPLC purification. The purity of all final compounds was determined by analytical HPLC with an Agilent ZORBAX SB-C18 (2.1 mm x 150 mm) or Varian Microsorb-MV 100-5 C18 column (4.6 mm x 150 mm), eluting with either H_2O/CH_3CN or H_2O/MeOH gradient solvent (+0.05% TFA).

Elution was monitored by a UV detector at l = 220 nm and l = 254 nm, with all final compounds displaying > 95% purity. Nuclear magnetic resonance (NMR) spectra were recorded on Bruker 300 or 500 MHz NMR spectrometers at ambient temperature.



3-Iodo-1-methyl-1H-pyrazolo[3,4-d]pyrimidin-4-amine (73 mg, 0.27 mmol, 1.0 equiv.) was dissolved in anhydrous DMF (3.5 mL) under nitrogen. Triethylamine (110 mg, 150 μL, 1.1 mmol, 4.0 equiv.), N-(3-ethynyl-4-methylphenyl)-3-(trifluoromethyl)-benzamide (120 mg, 0.40 mmol, 1.5 equiv.), bis(triphenylphosphine)palladium(II) dichloride (9.3 mg, 0.013 mmol, 0.05 equiv.), and copper (I) iodide (5.1 mg, 0.027 mmol, 0.10 equiv.) were added to the above solution sequentially. The reaction was heated at 50 °C for overnight and then quenched with a saturated NH_4Cl aqueous solution (5 mL). The resulting mixture was diluted with ethyl acetate (40 mL) and the organic phase was washed with a saturated $NaHCO_3$ aqueous solution (10 mL), brine (10 mL), and then dried over anhydrous Na_2SO_4. Purification by flash chromatography on silica gel using a gradient of EtOAc/Hexane (0:100 to 100:0) afforded N-(3-((4-amino-1-methyl-1H-pyrazolo[3,4-d]pyrimidin-3-yl)ethynyl)-4-methylphenyl)-3-(trifluoromethyl)benzamide as a pale brown solid (70 mg, 58%). HPLC purity>99%.

2.6 References

[1] Balzano, D., Santaguida, S., Musacchio, A., and Villa, F. (2011) A general framework for inhibitor resistance in protein kinases, *Chem Biol 18*, 966-975.

[2] Cocco, E., Scaltriti, M., and Drilon, A. (2018) NTRK fusion-positive cancers and TRK inhibitor therapy, *Nat Rev Clin Oncol 15*, 731-747.

[3] Zhang, J., Yang, P. L., and Gray, N. S. (2009) Targeting cancer with small molecule kinase inhibitors, *Nat Rev Cancer 9*, 28-39.

[4] Simmons, B., Bryant, J., Neiman, K., Spencer, L., and Arheart, K. (1990) The role of handwashing in prevention of endemic intensive care unit infections, *Infect Control Hosp Epidemiol 11*, 589-594.

[5] Apperley, J. F. (2007) Part I: mechanisms of resistance to imatinib in chronic myeloid leukaemia, *Lancet Oncol 8*, 1018-1029.

[6] Persky, N. S., Hernandez, D., Do Carmo, M., Brenan, L., Cohen, O., Kitajima, S., Nayar, U., Walker, A., Pantel, S., Lee, Y., Cordova, J., Sathappa, M., Zhu, C., Hayes, T. K., Ram, P., Pancholi, P., Mikkelsen, T. S., Barbie, D. A., Yang, X., Haq, R., Piccioni, F., Root, D. E., and Johannessen, C. M. (2020) Defining the landscape of ATP-competitive inhibitor resistance residues in protein kinases, *Nat Struct Mol Biol 27*, 92-104.

[7] Rossari, F., Minutolo, F., and Orciuolo, E. (2018) Past, present, and future of Bcr-Abl inhibitors: from chemical development to clinical efficacy, *J Hematol Oncol 11*, 84.

[8] Lovly, C. M., and Shaw, A. T. (2014) Molecular pathways: resistance to kinase inhibitors and implications for therapeutic strategies, *Clin Cancer Res 20*, 2249-2256.

[9] Hari, S. B., Perera, B. G., Ranjitkar, P., Seeliger, M. A., and Maly, D. J. (2013) Conformation-selective inhibitors reveal differences in the activation and phosphate-binding loops of the tyrosine kinases Abl and Src, *ACS Chem Biol 8*, 2734-2743.

[10] Krishnamurty, R., Brigham, J. L., Leonard, S. E., Ranjitkar, P., Larson, E. T., Dale, E. J., Merritt, E. A., and Maly, D. J. (2013) Active site profiling reveals coupling between domains in SRC-family kinases, *Nat Chem Biol 9*, 43-50.

[11] Leonard, S. E., Register, A. C., Krishnamurty, R., Brighty, G. J., and Maly, D. J. (2014) Divergent modulation of Src-family kinase regulatory interactions with ATP-competitive inhibitors, *ACS Chem Biol 9*, 1894-1905.

[12] Ahler, E., Register, A. C., Chakraborty, S., Fang, L., Dieter, E. M., Sitko, K. A., Vidadala, R. S. R., Trevillian, B. M., Golkowski, M., Gelman, H., Stephany, J. J., Rubin, A. F., Merritt, E. A., Fowler, D. M., and Maly, D. J. (2019) A Combined Approach Reveals a Regulatory Mechanism Coupling Src's Kinase Activity, Localization, and Phosphotransferase-Independent Functions, *Mol Cell*.

[13] Peraro, L., Deprey, K. L., Moser, M. K., Zou, Z., Ball, H. L., Levine, B., and Kritzer, J. A. (2018) Cell Penetration Profiling Using the Chloroalkane Penetration Assay, *J Am Chem Soc 140*, 11360-11369.

[14] Fowler, D. M., Araya, C. L., Fleishman, S. J., Kellogg, E. H., Stephany, J. J., Baker, D., and Fields, S. (2010) High-resolution mapping of protein sequence-function relationships, *Nat Methods 7*, 741-746.

[15] Fowler, D. M., and Fields, S. (2014) Deep mutational scanning: a new style of protein science, *Nat Methods 11*, 801-807.

[16] Barouch-Bentov, R., Che, J., Lee, C. C., Yang, Y., Herman, A., Jia, Y., Velentza, A., Watson, J., Sternberg, L., Kim, S., Ziaee, N., Miller, A., Jackson, C., Fujimoto, M., Young, M., Batalov, S., Liu, Y., Warmuth, M., Wiltshire, T., Cooke, M. P., and Sauer, K. (2009) A conserved salt bridge in the G loop of multiple protein kinases is important for catalysis and for in vivo Lyn function, *Mol Cell 33*, 43-52.

[17] Agius, M. P., Ko, K. S., Johnson, T. K., Kwarcinski, F. E., Phadke, S., Lachacz, E. J., and Soellner, M. B. (2019) Selective Proteolysis to Study the Global Conformation and Regulatory Mechanisms of c-Src Kinase, *ACS Chem Biol 14*, 1556-1563.

[18] Foda, Z. H., Shan, Y., Kim, E. T., Shaw, D. E., and Seeliger, M. A. (2015) A dynamically coupled allosteric network underlies binding cooperativity in Src kinase, *Nat Commun 6*, 5939.

[19] Tong, M., Pelton, J. G., Gill, M. L., Zhang, W., Picart, F., and Seeliger, M. A. (2017) Survey of solution dynamics in Src kinase reveals allosteric cross talk between the ligand binding and regulatory sites, *Nat Commun 8*, 2160.

[20] Wang, L., Pan, H., and Smith, D. L. (2002) Hydrogen exchange-mass spectrometry: optimization of digestion conditions, *Mol Cell Proteomics 1*, 132-138.

[21] Walters, B. T., Ricciuti, A., Mayne, L., and Englander, S. W. (2012) Minimizing back exchange in the hydrogen exchange-mass spectrometry experiment, *J Am Soc Mass Spectrom 23*, 2132-2139.

[22] Fang, J., Rand, K. D., Beuning, P. J., and Engen, J. R. (2011) False EX1 signatures caused by sample carryover during HX MS analyses, *Int J Mass Spectrom 302*, 19-25.

[23] Baker, P. R., and Chalkley, R. J. (2014) MS-viewer: a web-based spectral viewer for proteomics results, *Mol Cell Proteomics 13*, 1392-1396.

[24] Zhang, Z., Zhang, A., and Xiao, G. (2012) Improved protein hydrogen/deuterium exchange mass spectrometry platform with fully automated data processing, *Anal Chem 84*, 4942-4949.

[25] Kahsai, A. W., Rajagopal, S., Sun, J., and Xiao, K. (2014) Monitoring protein conformational changes and dynamics using stable-isotope labeling and mass spectrometry, *Nat Protoc 9*, 1301-1319.