Stony Brook University

Molecular mechanisms for specific kinase inhibitors and kinase promiscuity towards inhibitors

A Dissertation Presented

by

George Georghiou

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Molecular and Cellular Pharmacology

Stony Brook University

August 2014

Stony Brook University

The Graduate School

George Georghiou

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Dr. Markus A. Seeliger, PhD
Department of Pharmacological Sciences

Dr. W. Todd Miller, PhD
Department of Physiology and Biophysics

Dr. Miguel Garcia-Diaz, PhD
Department of Pharmacological Sciences

Dr. Yibing Shan, PhD
D.E. Shaw Research

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Molecular mechanisms for specific kinase inhibitors and kinase promiscuity towards inhibitors

by

George Georghiou

Doctor of Philosophy

in

Molecular and Cellular Pharmacology

Stony Brook University

2014

Dysregulated kinase activity is implicated in a wide variety of diseases due to the central role of kinases in many cell signaling pathways. Currently all small molecule kinase inhibitors used clinically target the ATP-binding pocket, which is highly conserved among all kinases due to their common use of ATP. Because of this conservation, targeting individual kinases is difficult, but has been achieved in the past with several drugs, the most famous of these being the Abl kinase inhibitor imatinib. While a number of broad-spectrum kinase inhibitors are in clinical use, specific kinase inhibition is favorable to avoid potential undesirable side effects. Typically, the focus is on the specificity of the kinase inhibitor and not on the ability of a kinase to bind multiple inhibitors promiscuously. It is important to understand the specificity of a kinase for inhibitors since some promiscuous kinases bind a large number of kinase inhibitors, in some cases with higher affinity than their intended targets. It is unknown how these promiscuous kinases are able to accommodate such a wide variety of inhibitors. To address how inhibitors can

achieve greater specificity, I characterized and determined the molecular mechanism of unusually specific macrocyclic peptide inhibitors of Src. These compounds are the first to distinguish Src kinase from other Src family kinases. To address kinase promiscuity, I examined how the receptor tyrosine kinase DDR1 binds to different classes of inhibitors. I determined that the promiscuity of DDR1 is due to a large hydrophobic pocket in the active site of the kinase that is formed when the kinase is in an inactive conformation. This inactive conformation is stabilized by unique interactions within the kinase domain. Consequently, disruption of these interactions by mutation increases the activity of the DDR1. These studies demonstrate that specific kinase inhibition can be achieved by targeting unique binding pockets in the active site of a kinase, decreasing the chance that new small molecule inhibitors will bind to promiscuous kinases like DDR1.

Dedication Page

I would like to dedicate this work to my family and friends. Without their support, none of this would have been possible.

Table of Contents

List of Figures

Chapter 1:

Chapter 3:

Chapter 4:

List of Tables

Chapter 3:

Chapter 4:

List of Abbreviations

ATP – Adenosine triphosphate
PKA – Protein kinase A
PI3K – Phosphatidylinositol-3 Kinase
S/TK – Serine/Threonine kinase
TK – Tyrosine kinase
DSK – Dual specificity kinase
RTK – Receptor tyrosine kinase
CTK – Cytoplasmic tyrosine kinase
DDR – Discoidin domain receptor
SFK – Src family kinase
KD – Kinase domain
N-lobe – Amino terminal lobe of the kinase domain
C-lobe – Carboxy terminal lobe of the kinase domain
P-loop – Phosphate binding loop
A-loop – Activation loop
CDK- Cyclin dependent kinase
CAMK – Calcium/Calmodulin kinase
SH2 – Src homology domain 2
SH3 – Src homology domain 3
EGFR – Epidermal growth factor receptor
IRK – Insulin receptor kinase
PDGFR – Platelet derived growth factor receptor
VEGFR – Vascular endothelial growth factor receptor
FGFR – Fibroblast growth factor receptor
MuSK – Muscle specific kinase
Eph – Ephrin receptor
ALK – Anaplastic lymphoma kinase
PTB – Phosphotyrosine binding domain
JAK – Janus kinases
FAK – Focal Adhesion Kinase
RSV – Rous Sarcoma Virus
HER2 – Human epidermal growth factor receptor 2
IGF-1R – Insulin like growth factor 1 receptor
HGFR – Hepatocyte growth factor receptor
ERK – Extracellular signal regulated kinase
ECM – Extracellular matrix
NSCLC – Non-small cell lung cancer
DS – Discoidin domain
DS-like – Discoidin-like domain
JM – Juxtamembrane region
MMP – Matrix metalloproteinase
MC - Macrocycle

Acknowledgments

To my family:

To my parents Andreas and Christina, thank you providing me with the love and support that without I would have not been able to accomplish so much. Katherine and Chris, thank you for your support and keeping me sane over the years when times were hectic. I look forward to seeing Daphne grow up and seeing how she will also in turn make you two grow as parents. Demetri and Susan, who initially got me interested in science as kid, and always were there for the momentous occasions of my life. Andrea, thank you for always checking in and letting me know how proud you are of me. You always provided the confidence boost I needed. And to my grandmother, whose wishes were for me to pursue my dreams and to see me accomplish great things in life. Thank you for everything you've done for me, and your continued love and support.

To my friends:

Morgan: we've been friends since we were in grade school, and have filled it with great memories. We've have had a lot of random adventures together over the years, exploring every inch of Long Island and taking epic road trips all over. You've been there when I needed someone to listen and have always given me sound advice. I'm proud of what we've accomplished together over the years and without you this PhD would have been impossible.

Brian: you've been the best drinking partner, band mate, teacher, and most importantly a good friend. Without you, I would have never picked up the bass guitar and found my greatest passion outside of science. Thanks for keeping me sane and getting me into the craziest scenarios every time we went out to keep life interesting.

Maggie, Leanne, Jess, Anand, Kazi, and Larry: we've been friends since our freshmen year at Stony Brook, and we've been making each year more special by filling it with great memories from crazy dinner parties and legendary BBQs. Our hangouts have made my time during graduate school all the better.

To my lab mates, former and current:

Ben, Derek, Grace, Ivan, Mike T, Mike P-G., Weibing and Zack: you were the ones that truly made earning this PhD a positive experience. I still believe we go down as the most sarcastic lab in the department, but through all the sarcasm, we always were supportive of one another. We've all lent a hand to one another when we need it and have learned from each other in the process.

And finally to my advisors, former and current:

Stella: You took me into your lab as a high school student and gave me a real perspective on how science is done. You and your lab always made sure to make me feel like I was family. The years I spent in your lab as a high school student and an undergraduate student inspired to me stay on and pursue my PhD. Thank you for all the opportunities and support you've given me.

Markus: As a new faculty member, you took a chance on someone who had no experience with structural biology. Luckily taking that chance worked out great for the both of us. You managed to teach me more than I ever thought I could know and provided the guidance I needed to grow as a scientist. You've been a great friend, and an amazing mentor. Thank you for everything and may your lab see continued success.

Publications

Georghiou, G.*, Kleiner, R,E.*, Pulkoski-Gross, M., Liu, D., Seeliger, M.A. Highly Specific bi-substrate competitive Src inhibitors from DNA-templated macrocycles. **Nature Chemical Biology**, 2012. 8, p366–374

*These authors contributed equally to this work

Chapter 1: Introduction

1.1 Kinases

Kinases are a class of signaling proteins that belong to the phosphotransferase family of enzymes due to their ability to catalyze a molecular reaction known as phosphorylation. Phosphorylation is the transfer of a high-energy phosphate (PO_4) from a phosphodonor, such as adenosine triphosphate (ATP), onto a substrate. Kinases can be defined into four classes based on the substrate they phosphorylate: 1) protein kinases, such as Protein Kinase A (PKA) and Src[1]; 2) lipid kinases such as phosphatidylinositol-3-kinase (PI3K) and sphingosine kinase; 3) carbohydrate kinases, such as hexokinase and phosphofructokinase; and 4) nucleotide kinases, such as thymidine kinase and nucleoside-monophosphate kinase. Each class of kinase is involved in a multitude of cellular processes that are tightly regulated to maintain homeostasis in a cell. Should kinases become dysregulated the well-orchestrated cellular events they help coordinate and maintain are no longer controlled, potentially leading to cell death or disease. It is not uncommon to find kinase signaling as part of the underlying pathology of a disease, particularly protein and lipid kinases[2]. Therefore, kinases are very attractive pharmaceutical targets[3,4]. For the last two decades, 50-70% of research and development efforts in the pharmaceutical industry have been dedicated towards the discovery of kinase inhibitors[4].

1.2 Eukaryotic Protein Kinases

Protein kinases mediate almost every cellular process, such as cell migration, metabolism, differentiation, and proliferation. Based on their critical role in mediating cellular processes, it is not surprising to find dysregulated kinase activity playing a role in diseases such as cancer,

hypertension, Parkinson's disease, and immunological diseases[2-4]. It is because of their central role in cell signaling and diseases that there is a need to understand how protein kinases function.

Eukaryotic protein kinases phosphorylate amino acids that have the potential to form reversible phosphate ester bonds. These amino acids include serine, threonine, tyrosine, histidine, cysteine, aspartate, lysine, and glutamate[1,5-7]. In prokaryotes, the phosphorylation of histidine, aspartate, and glutamate are the most prevalent forms of phospho-acceptors, however in eukaryotes, serine, threonine and tyrosine became the more prevalent targets for phosphorylation[6,8,9]. Interestingly, higher eukaryotes still phosphorylate histidine, aspartate, lysine and glutamate, however little is known about the kinases involved.

The human genome encodes for 538 protein kinases, which are referred to as the human kinome (Figure 1.1)[10,11]. The human kinome is divided into three primary groups: serine/threonine kinases (S/TK), tyrosine kinases (TK), and dual specificity kinases (DSK)[10]. Dual specificity kinases bear much resemblance to serine/threonine kinases however they phosphorylate serine, threonine, and tyrosine residues. Out of the 538 human kinases, 384 are S/TKs and only 90 are TKs.[12-14]. S/TKs are classified into 7 families, such as the AGC and STE, and further classified into 87 subfamilies[10,11]. The 90 TKs are classified into 2 families; receptor tyrosine kinases (RTKs) and cytoplasmic tyrosine kinases (CTKs), and further classified into 20 RTK subfamilies and 9 CTK subfamilies. An example of an RTK subfamily is the discoidin domain receptor (DDR) family and an example of CTK subfamily is the Src family kinases (SFKs)[12,15]. Members of these two TK subfamilies will be the primary focus of the work in this thesis.

1.2.1 The protein kinase domain

All eukaryotic protein kinases utilize ATP as a substrate; the catalytic domain of the kinase facilitates the transfer of the γ phosphate of ATP. The catalytic domain, which will be referred to as the kinase domain (KD), is a structurally conserved domain amongst all the kinases[16]. It has a bi-lobal structure, containing an amino (N)-terminal lobe that consists of β sheets (and helix αC), and a carboxy (C)-terminal lobe primarily composed of α helices. Connecting these two lobes is the hinge region, which outlines the active site located between the two lobes[17,18] (Figure 1.2).

Within the KD there are conserved motifs and secondary structures that play a role in the regulation and catalytic activity of kinases. These include the phosphate-binding loop (P-loop), αC-helix, and the activation loop (A-loop). There are also conserved triple amino acid motifs that are found throughout most of the human kinome: the C-lobe H/Y-RD motif and the A-loop DFG motif[17,19-22].

The P-loop (also referred to as the glycine rich loop) is an N-terminal anti-parallel β sheet motif that forms the "roof" of the active site by sitting at the bottom of the N-lobe of the kinase[23]. This loop hydrogen bonds to the α and β phosphates of ATP to properly align the γ phosphate for transfer onto a protein substrate[24]. Within the P-loop is a conserved salt bridge between a lysine and glutamic acid, with both residues opposite each other in an anti-parallel β sheet motif[25] (Figure 1.2).

The αC-helix is a highly conserved α helix in the N-lobe of the KD and plays a critical role in the regulation and catalytic activity of the kinase (Figure 1.2). Originally described in the first structure of PKA, the αC-helix was the third helix found in the N-lobe of PKA. It is C-terminal to the αA- and αB-helices of PKA, which are not highly conserved secondary structures

of the N-lobe amongst all kinases[23]. The αC-helix contains a conserved glutamate that forms a salt bridge with the catalytic lysine in the active site of the kinase. It also acts as a means of regulating kinase activity by rotating outwards, breaking the salt-bridge between the αC-helix glutamate and catalytic lysine, thus inactivating the kinase[26,27]. This rotation is particular to one of two known inactive kinase conformations.

The KD contains two triple amino acid motifs necessary for catalytic activity. The first of these motifs is the histidine/tyrosine-asparagine-aspartate (H/Y-RD) motif. This motif is in the C-lobe of the kinase and is N-terminal to the A-loop. The aspartate of this motif coordinates the hydroxyl acceptor on the protein substrate. The histidine/tyrosine of this motif forms hydrophobic contacts with the phenylalanine of the aspartate-phenylalanine-glycine (DFG) motif in the A-loop. The histidine/tyrosine of this motif will also form hydrogen bonds to the carbonyl of the Asp of DFG motif[28].

The second triple amino acid motif is the aspartate-phenylalanine-glycine (DFG) motif at the N-terminus of the A-loop (Figure 1.2). The aspartate of this motif interacts and coordinates with the γ-phosphate of ATP for transfer via a magnesium ion[24]. The phenylalanine of the DFG motif forms hydrophobic contacts with the αC-helix by the phenyl group of the amino acid occupying a hydrophobic pocket underneath the helix[28]. Occupying this pocket contributes to the stability of the salt bridge between a conserved glutamate on the αC-helix and the catalytic lysine. The phenylalanine of this motif interacts with the histidine/tyrosine of the H/Y-RD motif to properly align the aspartate of the DFG motif in the active site. The aspartate and phenylalanine of the DFG motif also serve as part of a regulation mechanism for kinases. The amino acids are able to undergo a 180° flip, where the aspartate faces out of the active site, and phenylalanine

faces into the active site[29]. This inactivates the kinases and is a hallmark for one of the two known inactive conformations that kinases can adopt[27].

In addition to the DFG motif, the A-loop also contains the site(s) for kinase autophosphorylation. There are either single or multiple residues that undergo autophosphorylation in the A-loop. Autophosphorylation occurs typically in trans, meaning another kinase will phosphorylate the autophosphorylation residue(s). For S/TKs these are usually serine/threonine and for TKs a tyrosine. Autophosphorylation is necessary for kinases to become fully active. The autophosphorylated residue locks the A-loop in an extended conformation by forming salt-bridges or hydrogen bonds via the phosphate group with surrounding amino acids, keeping the kinase in an activated state[27,28]. This residue becomes dephosphorylated via phosphatases, which causes the A-loop to collapse back onto the kinase, allowing the kinase to adopt an inactive conformation[19].

1.2.3 Protein Kinase Regulation

Since protein kinases are important cellular signaling enzymes, their own activity needs to be tightly regulated. Therefore, regulatory elements have evolved to help orchestrate the many cellular processes protein kinases participate in. Regulatory elements can be diverse, ranging from co-factors that bind directly to the kinase domain, to different protein domains attached to the kinase domain. Besides controlling kinase activity, regulatory elements determine cellular localization and substrate protein specificity[5]. For receptor kinases, their regulatory domain is typically an extracellular ligand-binding domain, such as epidermal growth factor receptor

(EGFR), which requires the bindings of epidermal growth factor (EGF) to stimulate kinase activity[15].

There are several different forms of kinase regulation for non-receptor kinases. For example, cyclin-dependent kinases (CDKs) require the co-factor cyclin to bind to stimulate kinase activity so that the cell cycle can progress. The binding of cyclin is reversible, and upon its breakdown or dissociation from its binding site on the CDKs, the CDK becomes inactive preventing it from promoting cell cycle progression[30,31]. Another example of kinases regulated by co-factor binding is the calcium/calmodulin kinases (CAMKs). The binding of calmodulin and calcium to the regulatory domains activates the kinase by dissociating the kinase domains from the inactive holoenzyme complex[32].

Cytoplasmic protein kinases that do not require co-factor binding for activity are typically regulated by intrinsic domains. These regulatory domains affect enzymatic activity, control cellular localization, and determine substrate specificity[5,22,33-36]. An example of this is the protein tyrosine kinase Src. Src contains two regulatory domains: the Src homology 2 (SH2) and Src homology 3 (SH3) domains. These domains inhibit Src activity by binding to specific motifs on the kinase, keeping it in a 'closed' inactive state[37]. The SH3 domain binds to a PxxP motif on the linker between the SH2 and kinase domain[38]. By binding this motif, the SH3 presses onto the back of the kinase, forcing the N- and C-lobes closer and reduces the size of the active site[39]. The SH2 domain binds the phosphorylated C-terminal tail of Src kinase domain, which then allows the SH2 domain to press onto the back of the C-lobe of the kinase domain[40]. The combined actions of SH2 and SH3 binding to the kinase domain forces the kinase into an inactive conformation, preventing it from phosphorylating protein substrates. Another example is Abl kinase. Like Src, it has an SH2 and SH3 domain; however Abl lacks the C-terminal

phosphorylation site that binds to the SH2 domain. As a substitute for this, Abl binds its unique myristoylated N-terminal domain in the myristic acid binding pocket within the C-lobe of the kinase domain. The binding of the N-terminal domain locks the SH2 and SH3 against the back of the kinase and Abl is inactivated.

1.2.4 Protein Kinase Domain Conformations

The kinase domain is a dynamic structure that can access several conformations. Currently, there is only one known active conformation and two known inactive conformations that kinases can adopt. It is unknown whether all kinases can access both inactive conformations[19,29,41].

The active conformation is defined by the aspartate of the DFG motif facing into the active site, the phenylalanine of the DFG motif occupying a hydrophobic pocket underneath the αC-helix, and the αC-helix rotated inwards, so that the salt bridge between the conserved glutamate on the αC-helix and the catalytic lysine is formed (Figure 1.3)[18,26]. The A-loop will also extend outwards, away from the kinase domain, allowing the autophosphorylation residue(s) to be positioned for trans-autophosphorylation[19].

There are two inactive conformations that kinases can potentially adopt. One conformation is the Abl/c-Kit-like inactive conformation. It is defined by the aspartate and phenylalanine of the DFG motif undergoing a 180° flip, which causes the aspartate to face out of the active site, preventing it from coordinating the phosphate groups of ATP and the magnesium ion required for phosphorylation (Figure 1.3)[42]. The side chain of phenylalanine faces into the active site and occludes the ATP binding pocket. Phosphatases can de-phosphorylate the A-loop,

causing the A-loop to collapse back onto the kinase domain, blocking protein substrate from binding to the kinase domain.

The other inactive conformation that kinases can access is the Src/Cdk-like inactive conformation[43]. Starting from the active conformation, the αC-helix rotates outwards, breaking the salt bridge between the αC-helix glutamate and catalytic lysine (Figure 1.3)[19,26]. Similar to the Abl/c-Kit-like inactive conformation, the A-loop also collapses back onto the kinase domain, preventing protein substrate from binding to the kinase domain.

1.3 Protein Tyrosine Kinases

Ninety of the 538 human kinases are protein tyrosine kinases. These are further categorized into cytoplasmic tyrosine kinases (CTKs) and receptor tyrosine kinases (RTKs)[10,12,13,44,45]. These two families are divided into 20 RTK subfamilies and 10 CTK subfamilies, based on sequence and domain homology[10,15].

1.3.1 Receptor Tyrosine Kinases

Fifty eight of the 90 human tyrosine kinases are RTKs and are classified into 20 subfamilies[15]. These kinases have an overall similar architecture, containing an extracellular ligand-binding domain connected through a single α helical transmembrane domain to an intracellular tyrosine kinase domain, plus additional C-terminal or juxtamembrane regulatory domains (Figure 1.4)[38]. Several RTKs also contain pseudokinase domains. Pseudokinases lack kinase activity due to amino acid substitutions within catalytically important regions and

typically have an allosteric regulatory effect on RTK activity[46]. The 20 RTK subfamilies are: the epidermal growth factor receptor (EGFR) family, the insulin receptor kinase (IRK) family, the platelet-derived growth factor receptor (PDGFR) family, the vascular endothelial growth factor receptor (VEGFR) family, the fibroblast growth factor receptor (FGFR), the PTK7 family, the Trk family, the Ror family, the muscle-specific kinase (MuSK) family, the Met family, the Axl family, the Tie family, the ephrin receptor (Eph) family, the Ret family, the Ryk family, the discoidin domain receptor (DDR) family, the Ros family, the LMR family, the anaplastic lymphoma kinase (ALK) family, and the STYK1 family (Figure 1.4)[15]. Each of these families is categorized based on their extracellular domains and sequence homology[15].

Most RTKs follow a similar mechanism of activation, beginning with an extracellular ligand binding the receptor, causing homo- or heterodimerization with another member of its RTK family, leading to a trans-autophosphorylation of the dimerized kinases and the subsequent phosphorylation of downstream proteins[2,38,45,47]. There are several cases where the RTK dimers already form in the absence of ligand, such as the DDRs; however, the kinases involved in the dimer are not active until a ligand binds to the extracellular domains. These RTKs are kept inactive by a form of cis-autoinhibition by intramolecular interactions within the kinase's intracellular region. RTKs will remain inactive until ligands bind, which leads to the disruption of the inhibitory complex by trans-autophosphorylation.

Several RTKs are cis-autoinhibited through their A-loops on the kinase domain[48,49]. In the IRK family, the A-loop contains a YxxxYY motif, where the 2nd tyrosine of this motif faces into the active site, forming a hydrogen bond with the αD-helix of the C-lobe. This interaction stabilizes a conformation of the A-loop that occludes the active site[50-52]. Trans-autophosphorylation of all three tyrosine residues of this motif upon insulin binding disrupts this

complex, allowing the A-loop to uncouple from the C-lobe and activates the kinase[53]. In FGFR RTKs a similar mechanism is used; however tyrosines in the A-loop do not directly block the substrate protein binding site. In this case, the tyrosines form interactions that stabilize the inactive conformation of the kinase, indirectly blocking the protein binding site, and leaving the ATP binding site open[2,22,38]. Trans-autophosphorylation disrupts these interactions and allows the kinase to adopt the active conformation[47].

In contrasts, there are RTKs that are cis-autoinhibited by regions outside the kinase domain. The best example of this cis-autoinhibition is seen in MuSK, Kit, Flt3 and Eph family RTKs, where the juxtamembrane region makes extensive inhibitory contacts with the kinase domain[54-57]. These contacts are along the A-loop and other regions of the KD. The specific intramolecular interactions between the juxtamembrane and KD differ for each kinase, however they are all mediated through key tyrosine residues on the juxtamembrane region. Mutations in Kit- and PDGFR-family RTKs can disrupt interactions with the juxtamembrane and kinase domain, constitutively activating the kinase. These mutations are frequently found in a number of cancers related to constitutive Kit and PDGFR activity[15]. C-terminal regions of some RTKs can also participate in cis-autoinhibition of an RTK. This is best exemplified with Tie2, where the kinase appears to be in an active conformation but the C-terminal tail containing a tyrosine autophosphorylation sites blocks the active site, preventing protein substrates from binding[58]. It is hypothesized that autophosphorylation of this tail, triggered by ligand binding, will disrupt this autoinhibitory complex.

Trans-autophosphorylation of RTKs not only disrupts cis-autoinhibition, but also allows RTKs to serve as signaling hubs through their kinase domain, juxtamembrane, and C-terminal regions[5,59]. Upon autophosphorylation of the A-loop, tyrosines along the kinase domain,

juxtamembrane, and C-terminal regions also become phosphorylated, acting as docking sites for other proteins containing phospho-tyrosine binding domains, such as SH2 or PTB domains[60-63]. Binding of proteins with SH2/PTB domains to RTK phosphotyrosine motifs recruits other potential downstream targets of RTKs for phosphorylation, or attracts scaffold proteins to assemble large signaling hubs through the recruitment of other signaling proteins. These hubs can also attract ubiquitin ligases which will in turn ubiquinate RTKs, and target them for degradation[15,64]. The ubiquitin-mediated degradation of RTKs serves as a negative feedback mechanism that directs the termination of an RTK's influence on a signaling network.

When RTKs become dysregulated or dysfunctional they can cause diseases such as cancer. Dysregulated RTK activation in cancer is controlled by four potential mechanisms: autocrine activation, chromosomal translocation, RTK overexpression, or gain-of-function mutations[15]. These mechanisms can lead to cells becoming tumorigenic, and has made RTKs a pharmacologically attractive target for treating cancers.

1.3.2 Cytoplasmic Tyrosine Kinases

Thirty two tyrosine kinases are classified as cytoplasmic tyrosine kinases and are broken down into 10 subfamilies[12,13]. The subfamilies are: the Src family kinases (SFKs), Abl family kinases, Tec family kinases, Csk family kinases, Syk family kinases, Ack family kinases, Janus kinase (JAK) family, focal adhesion kinase (FAK) family, FRK family kinases, and FER family kinases (Figure 1.5). These kinases are grouped based on sequence and domain homology. Unlike the RTKs, CTKs are more modular, containing a variety of domains attached to the KD[5]. The attached domains can serve as activity modulators, determine subcellular localization of a

kinase, act as docking sites, and determine substrate protein specificity. These additional domains can be thought of as 'beads on a string' and are found either N- or C-terminally from the KD.

The domains attached to the kinase will dictate where kinases will localize and which proteins are substrates for a particular kinase[36,65,66]. This was best demonstrated by the replacement of the regulatory SH3 and SH2 domains of the SFK Hck with PDZ domains, which redirected Hck to another set of proteins instead of its normal targets[33]. In Abl family kinases the actin and DNA binding domains localize the kinases to protein substrates in the cytoplasm and in the nucleus. The SH2 and SH3 domains of Abl dictate which proteins it will phosphorylate in these environments[2,12]. The Tec family kinases contain an N-terminal pleckstrin homology (PH) domain, which allows them to localize to phosphatidylinositol lipids on the plasma membrane[12]. The FAK family kinases also have unique domains that contribute to their cellular localization. The integrin-binding domain allows FAK kinases to bind to collagen fibers, while the focal adhesion-binding domain allow FAK to bind to cytoskeleton proteins[12].

Many of the CTKs contain the phosphotyrosine-binding SH2 domain (Figure 1.5). In most instances it serves the same inhibitory function that is seen with the binding of the phosphorylated C-terminal tail of the Src kinase domain (described above); however, this is not always the case[40]. The SH2 domain of Fes kinase has a positive effect on kinase activity by interacting with the tip of the N-lobe of the KD, stabilizing the active conformation. The SH2 domain in this instance, through electrostatic and packing interactions, stabilizes the αC-helix so that it faces inwards and keeps the salt bridge between the αC-helix glutamate and catalytic lysine intact. The SH2 domain is able to press against the N-lobe of the kinase only when phosphotyrosine ligand binds to the SH2 domain[67,68]. In Abl kinase, the SH2 domain has both a

stimulatory and inhibitory function. In the inactive conformation, the SH2 domain presses down on the KD to bring the myristoylated N-terminus of Abl into proximity to bind to the myristic acid binding pocket on the C-lobe of the kinase[42,69]. In the active conformation the SH2 domain stimulates kinase activity in a fashion similar to that seen in Fes kinase. The interactions that the SH2 domain has with Abl in the active conformation differ from those in Fes[69]. In Abl, the interaction is centered on an isoleucine and mutation of this isoleucine in the cancer-causing BCR-Abl fusion protein results in the loss of pro-oncogenic function[70]. The SH2 domain can also help localize CTKs by binding to phosphotyrosines on other proteins, such as RTKs and scaffold proteins[65].

A majority of CTKs contain an SH3 domain as well (Figure 1.5). Unlike the SH2 domain, which binds phosphotyrosine motifs, SH3 domains bind poly-proline motifs (PxxP)[36,37,66,71]. SH3 domains can function as kinase regulators, such as with Src kinase (described earlier), and localize kinases to proteins that contain a poly-proline motif. This also contributes to the protein substrate selectivity of many CTKs by binding specific poly-proline motifs on substrate proteins[36,66].

Like RTKs, CTKs have also been implicated in a number of diseases, the most prominent being cancer. CTK dysregulation in cancer is mediated by four mechanisms: dysregulation of upstream protein signaling, chromosomal translocation, CTK overexpression, or gain-of-function mutations. One of the best studied examples is BCR-Abl, where a break at breakpoint cluster region (BCR) on chromosome 22 is translocated to the Abl gene on chromosome 9, resulting in the fusion kinase BCR-Abl[3,72]. This fusion kinase lacks the ability to be regulated due to the loss of its myristoylated N-terminal tail and causes chronic myeloid leukemia (CML)[73]. Brk is another example of CTKs involved in cancer. Brk overexpression in

mammary epithelial cells sensitizes them to epidermal growth factor signaling (EGF), that when combined with other mutations can lead to the breast cancer[74]. Similar to RTKs, CTKs have also become a very attractive pharmacological target[3].

1.4 Src and Src family kinases (SFKs)

1.4.1 Discovery of Src and SFKs

Src was the first viral proto-oncogene to be discovered[75,76]. The virus carried a gene referred to as the Rous Sarcoma Virus (RSV) gene, which coded for the protein v-Src. Expression of v-Src by infected cells resulted in tumorigenesis. Using DNA sequencing a cellular homologue of Src (c-Src) was discovered[77]. It was later determined that both v-Src and c-Src were CTKs. Interestingly, c-Src was not able to transform normal cells. Comparing the DNA sequence of the v-Src and c-Src revealed that v-Src was missing a portion of the C-terminus of the protein. Through cellular and biochemical experiments the C-terminal tail was discovered to be critical for regulating Src kinase activity, which explained why v-Src was able to transform normal cells and c-Src was not.

Using sequence comparison other CTKs similar to Src were discovered. These kinases were Fyn, Yes, Blk, Yrk, Fgr, Hck, Lck, and Lyn[2,78]. Each had an overall conserved DNA sequence encoding for a unique N-terminal domain (refereed to as an SH4 domain, (SH standing for Src Homology)), followed by an SH3 domain, an SH2 domain, and a kinase domain (referred to originally as an SH1 domain) (Figure 1.6)[79,80]. There is high degree of sequence conservation amongst the SFKs. Yes kinase is the most homologous to Src (76% homologous to Src),

followed by Fyn (70%), Fgr (66%), Hck (58%), Blk (56%), Lyn (55%), and Lck (54%). The homologous structure of the SFKs translates to a similar mechanism of regulation where the SH2 domain binds the phosphorylated C-terminal tail of the kinase domain and locks the kinase in an inactive complex[37].

Of the SFKs, only Src, Fyn, Yes, and Yrk are expressed ubiquitously in human cells[81]. The other SFKs have different expression patterns, primarily in different types of immune and neuronal cells[82-84]. Hck, for example is exclusively expressed in myeloid cells, while Blk is expressed solely in B cells. Fgr and Lyn kinases are also expressed in both myeloid and B cells. Lyn is also expressed in brain tissue, along with Lck. Lck is also found in T cells and natural killer (NK) cells.

SFKs are highly conserved amongst organisms. The sequence homology of mouse and chicken Src compared to human Src is 99% and 95% respectively. Homologous variants of the SFKs have been found in unicellular organisms, such as the choanoflagellate *Monosiga brevicollis*[85]. MbSrc4 is homologous to mammalian Src, having similar catalytic activity and domain architecture. Unlike mammalian Src, MbSrc4 contains a lipid binding C2 domain that allows it to bind to the plasma membrane[86]. The C2 domain replaces the myristoylated N-terminal region found in mammalian Src. The high sequence homology of Src in unicellular and multicellular organisms suggests that Src signaling is ancient and had a common ancestor that existed in earlier eukaryotes[85].

1.4.2 SFKs are involved in multiple signaling pathways

The SFKs are involved in many different signaling pathways in the cell. They mediate signal transduction from RTKs, G-protein coupled receptors (GPCRs), and steroid receptors[81]. Signaling from these receptors involve SFKs in pathways related to cell growth and proliferation[87,88]. SFKs also play a critical function in cell migration and adhesion by phosphorylating FAK and p120-catentin[89]. SFKs also mediate the cross talk between GPCRs, RTKs, and steroid receptor signaling pathways. The involvement of SFKs in these pathways has shown that SFKs play a critical role in oncogenic signaling, where they promote tumor cell growth and metastasis by being overexpressed or constitutively active due to upstream factors, such as RTKs, becoming dysregulated[90].

SFKs also mediate immune cell development and activation. In T-cells, Lck and Fyn kinases are activated by the antigen receptor, leading to the phosphorylation of downstream proteins involved in the T-cell signaling cascade[82]. Lyn, Fyn and Blk have a similar involvement in B-cell signaling[83]. Hck, Fgr, and Lyn also mediate myeloid cell activation. These cellular responses are not entirely mediated by the specific SFKs mentioned, but can be mediated by Src and Yes as well. In all instances, the activation of SFKs ultimately leads to the activation of transcription factors necessary for immune cell proliferation and differentiation.

1.4.3 Src kinase signaling in disease

Of the SFKs, Src is the one that is most often implicated in cancer[2]. In both normal and oncogenic cells c-Src regulates cell adhesion and migration[75]. Its regulation of FAK, PI3K, EGFR, human epidermal growth factor receptor 2 (HER2/ErbB2), PDGFR, insulin-like growth factor 1 receptor (IGF-1R), c-Met/hepatocyte growth factor receptor (HGFR), VEGFR, Akt,

STAT3, integrin receptor and erythropoietin receptor (EpoR) signaling pathways has made it a very attractive target for inhibition in the treatment of multiple cancers[90]. In particular, the discovery of the involvement of c-Src in tumor growth and metastasis in breast, colorectal, pancreatic, lung, bone and ovarian cancers has made it a potential therapeutic target in halting the progression of these cancers[90-92].

In many of these cancers, c-Src kinase activity and protein expression levels are elevated when compared to normal cells. The elevated c-Src activity (whether it be due to overexpression or upstream activators of c-Src being dysregulated) has been implicated less in promoting tumor cell proliferation but more in tumor cell metastasis and invasion[75,90]. Tumor cell metastasis and migration require cytoskeletal rearrangement and focal adhesion turnover, which is mediated by FAK signaling[93]. Once tumor cells have metastasized to other organs, Src signaling is necessary for tumor cell survival by inducing apoptotic resistance in response to chemokine signaling by immune cells. c-Src also mediates metastatic tumor cell survival through extracellular signal-regulated kinase (ERK)-dependent activation of cell survival signals[90]. Furthermore, c-Src signaling confers cell death resistance to metastatic tumor cells that have separated from the extracellular matrix (ECM). In colorectal cancer Src signaling can also contribute to angiogenesis through FAK, VEGFR, and HGFR signaling[90].

c-Src activity can also confer therapeutic resistances in cancer. In CML, Src can directly interact through kinase-dependent and –independent mechanisms with BCR-Abl, causing inhibitor resistance[90,94]. Src, in combination with EGFR signaling, can also confer inhibitor resistance in steroid receptor driven cancers. For example, in estrogen receptor driven breast cancer, Src activation contributes to tamoxifen resistance and is associated with poor patient response to tamoxifen[75,90]. Finally, c-Src activation contributes to the kinase inhibitor resistance

of RTKs in cancer. This is best exemplified in HER2 driven metastatic breast cancer. HER2 activates Src directly and Src signaling promotes c-Src protein synthesis and stability. This upregulated activation and expression of Src promotes the upregulation of EGFR activity and the dephosphorylation of the tumor suppressor protein PTEN, thus upregulating tumor cell proliferation. As a result of this, Src activity confers resistance to the effect of the HER2 inhibitor trastuzumab[95]. A similar effect of upregulated Src activation is seen in EGFR driven non-small cell lung cancer (NSCLC), where Src activation confers EGFR resistance to cetuximab by promoting activation of downstream PI3K/Akt signaling cascades and facilitating EGFR nuclear translocation[90,96].

In cellular and mouse cancer models, inhibition of Src has proven to be highly effective in reducing tumor growth and metastasis in cancer models[90,92,97]. It would therefore be highly beneficial to target Src and SFKs in combination with other cancer drugs as a new means of improving cancer treatment. Currently, there are no drugs that inhibit Src specifically.

1.5 Discoidin Domain Receptors (DDRs)

1.5.1 DDRs are unique RTKs that bind to extracellular collagen

Discoidin domain receptors (DDRs) are a unique RTK family with only two members, DDR1 and DDR2 (Figure 1.4)[98-100]. Originally identified through homology cloning based on their catalytic domain, they were considered orphan receptors until it was discovered that they are activated by collagen[101]. The DDRs act as extracellular matrix sensors for the cell, bridging two receptor families; the extracellular matrix receptors and RTKs. The DDRs are activated by

extracellular collagen in a slow and sustained manner[98,99]. Their activation can take up to 18 hours with continuous collagen stimulation to reach full tyrosine kinase activity[98,99]. The activation time is unusually slow when compared to the other RTKs, which are usually activated on the second to minute timescale[15,102]. Similar to some RTKs, such as the IRK family, the DDRs are expressed as monomers and form constitutive dimers as they are transported to the cell membrane. These dimers are inactive until collagen binds to their extracellular domains[103,104].

DDR1 and 2 bind different types of extracellular collagen. While fibrillar collagens, like collagens I and III, bind and activate both DDRs, the ability to bind non-fibrillar collagens differs between the two[98]. For example, DDR1 can bind to the basement membrane collagen IV and collagen VIII, while DDR2 can bind collagen II and collagen X.

1.5.2 DDR expression and isoforms

DDR1 and DDR2 are expressed in a wide variety of tissues, and play important roles in both development and in adult organisms[102]. DDR1 mRNA is found in high levels in the brain, lungs, kidneys, spleen and placenta, while DDR2 mRNA is found at high levels in skeletal and heart muscles, the kidneys, and lung tissues[105]. DDR1 is primarily expressed in epithelial cells while DDR2 is found in mesenchymal cells[106]. Both DDRs are expressed in the developing nervous system, as well as in immune cells.

There are five known isoforms of DDR1 and only one for DDR2 (Figure 1.7). DDR1 isoforms (referred to as DDR1a, b, c, d, and e) are a result of alternative splicing of the 17 exons that encode the protein. The extracellular region is encoded by exons 1-8, the transmembrane by exon 9, the cytosolic juxtamembrane region by exons 10-12, and exons 13-17 encode for the

kinase domain[101]. The two most common isoforms of DDR1 are DDR1a and DDR1b[100]. These are a result of alternative splicing of exons 10-12, where DDR1a lacks exon 11 and DDR1b contains exons 11-12. DDR1c differs from DDR1a and b by containing an additional cryptic splice site at the 5' end of exon 14, resulting in an additional 6 amino acids in the kinase domain[101]. DDR1d and e are truncated or non-active forms of DDR1. DDR1d is formed by the exclusion of exons 11 and 12, resulting in a frame shift mutation and the truncation of the protein. This isoform lacks the kinase domain as a result. DDR1e is formed also as the result of exon 11 and 12 deletion however no frame shift mutations occur since a portion of exon 10 is also deleted. The resulting protein is not truncated, but codes for an inactive kinase domain by a portion of the ATP binding pocket not being expressed due to the deletion of exon 12. DDR2 mRNA is comprised of 19 exons and has no splice variants[101,107].

DDR1a and b are expressed at different stages in development and in the adult organism and have functional differences. For example, DDR1b is the predominant isoform expressed during embryogenesis, while DDR1a is commonly found in human mammary cells in the adult organism[108]. As an example of functional differences, the loss of 37 amino acids in DDR1a by the exclusion of exon 11 results in different functions between DDR1a and DDR1b. DDR1b can interact with the adaptor protein ShcA while DDR1a can interact with the FGFR substrate-2[60,98,102]. As a result, DDR1a is unique compared to DDR1b since this interaction grants it the ability to trigger cell migration in leukocytes. This suggests that different DDR1 isoforms are 'tuned' for different signaling pathways.

1.5.3 DDR structure

The DDRs are composed of an extracellular discoidin homology (DS) domain, which binds collagen; a discoidin-like (DS-like) domain that contributes to collagen-induced receptor activation; and an extracellular juxtamembrane region, which is heavily glycosylated and contains matrix metalloproteinase (MMP) sites (Figure 1.7)[109-111]. Following the extracellular domains is a single α helical transmembrane domain that mediates collagen-independent receptor dimerization; a large juxtamembrane region, which contains phosphotyrosines that serve as docking sites for proteins containing SH2 or PTB domains; and an intracellular kinase domain that is homologous to the IRKs (Figure 1.7)[105].

The DS domain (also known as the factor 5/8 type C domain due to homology with blood coagulation factors) is homologous to the lectin discoidin I, which was first described as a galactose binding lectin of the slime mold *Dictyostelium discoideum* (Figure 1.7)[112]. This globular domain is approximately 160 amino acids long and interacts with the globular 182 amino acid long DS-like domain. Between DDR1 and DDR2, the DS domain is 59% sequentially conserved, while the DS-like domain is only 51% sequentially conserved. The difference between DDR1 and DDR2 in the DS domain is responsible for the specific types of collagen that the two kinases bind[113].

The structures of the extracellular domains of DDR1 and DDR2 have revealed that the DS domain is composed of an eight stranded β-barrel arranged in two antiparallel β sheets[110,111]. The top of the barrel is characterized by five protruding loops, while the bottom of the barrel is formed by connections between the antiparallel β sheets. Collagen binding is mediated by the formation of a trench from the loops that protrude from the top of the β barrel, which allows for triple helical collagen to fit and make appropriate contacts[109]. Highly conserved di-sulfide bonds hold the N- and C-termini of the DS-domain together. The relative affinity for collagen for a

single DS-domain is modest, however the affinity increases greatly when the DS-domain is homodimerized to another DDR kinase[113,114].

The DS-like domain is also a β-barrel arranged in two antiparallel β sheets; however there is a long insertion between the β1 and β2 strands, as well as five additional strands; two N-linked glycosylation sites and a calcium binding site (Figure 1.7). Unlike the DS domain, the DS-like domain lacks the disulfide between the N and C termini. The loss of this bond is compensated by other cysteines that form an intramolecular disulfide bridge. The DS and DS-like domain are connected via a short linker region, and form a compact structure where the two domains are offset from one another by a 90° rotation. The long insertion of the DS-like domain forms extensive contacts with the flat bottom of the β barrel of the DS domain (Figure 1.7)[110,111]. This region may be subject to conformational changes as a result to collagen binding to the DS domain; however this has yet to be proven.

The DS-like domain connects to the transmembrane domain through an extracellular juxtamembrane (JM) region, which is proline rich and contains many N-linked glycosylation sites and several predicted O-link glycosylation sites (Figure 1.7)[115]. In DDR1, the extracellular JM is 50 amino acids long and only 32 amino acids long in DDR2. Unlike the DS and DS-like domain, the extracellular JM is not well conserved between DDR1 and DDR2[116-118]. Interestingly, the extracellular JM of DDR1, but not DDR2, has been shown to be cleaved by MMP-14, -15, and -16, which releases the entire extracellular region. This may act as means of regulating collagen-induced receptor activation and also suggests that DDR1 and DDR2 have different means of regulating receptor activation[119,120].

Connecting the extracellular domains to the intracellular domains is a single α helical transmembrane domain. The transmembrane region plays an important role in the formation of

DDR homodimers by forming contacts with an adjacent DDR as they are transported to the plasma membrane. The dimerization of DDRs is not entirely based on interactions in the transmembrane domain, but additional interactions in the homodimer between extracellular domains and intracellular domains[105,117]. The transmembrane region of DDR1 contains several leucines that are critical for DDR1 activation. These leucines are thought to form a leucine zipper motif, which may cause conformational changes in DDR1 upon collagen binding activating the kinase[104]. This conformational change may involve a rotation of the transmembrane domains within the dimer. The intracellular domains rotate to face one another, allowing trans-autophosphorylation to take place. A similar function for transmembrane leucine zippers needed for the activation of receptors has been described in the erythropoietin receptor[104].

Following the transmembrane region is the intracellular JM region. This region contains several key phosphotyrosines that are necessary for interactions between the DDRs and other signaling/adaptor proteins[60]. Many of these adaptor protein interactions are mediated by phosphotyrosine binding domains, such as SH2 and PTB domains (Figure 1.7)[60,113,121,122]. These adaptor proteins may use the DDRs as scaffolds for a signaling hub, allowing the collagen-stimulated activation of the DDRs to communicate with a number of different signaling pathways. In DDR1, isoforms b and c contain an additional 37 amino acids in the JM region and contains the NPxY motif that allow the b and c isoform to interact with the adaptor protein ShcA (Figure 1.7)[60,117]. As mentioned earlier, DDR1a is missing these residues; however, as a result, has the appropriate motif for FGFR substrate-2 to interact its intracellular JM region, which DDR1b and c lack. DDR2 differs from DDR1 in that only one phosphotyrosine site has been found in its intracellular juxtamembrane region, which allows it to interact with the ShcA adaptor protein[60].

The kinase domains of the DDRs are similar to many of the other RTKs, bearing the greatest sequence homology to TrkA (approximately 68% homologous)[123]. The structure of the DDR1 kinase domain was solved recently by Canning et al. and will be discussed in a later chapter, alongside the structures of the kinase domain of DDR1a that were solved in this thesis project[124].

1.5.4 DDR-mediated signaling

DDRs are involved in multiple signaling pathways relating to ECM remodeling, cell migration, adhesion, proliferation and differentiation[102,105,107,113,117]. These signaling pathways are initiated by DDRs in a context and cell dependent manner. Mediated by phosphotyrosine interactions (DDR1a having 13 tyrosines, DDR1b and c 15 tyrosines, and DDR2 14 tyrosines) in the juxtamembrane and kinase domains, DDR1 and 2 can bind adaptor and other signaling proteins that contain either an SH2 or PTB domain (Figure 1.7)[60,122]. There is insufficient data showing whether all tyrosines get phosphorylated in either DDR1 or 2, however independent studies have found a multitude of proteins that interact with phosphotyrosines on DDR1 and 2.

For DDR1, these proteins include scaffold proteins (such as Crk, and Nck1/2), docking proteins (ShcA), guanine exchange factors (PLC-γ1 and Vav1/2), SFKs (Src, Yes, Lyn, and Fyn), Csk, STATs (STAT1a/b, STAT3, and STAT5b), PI3K (p85a/b and p110a/b/d), and phosphatases (SHIP-1/2 and Shp1/2) (Figure 1.7)[60,125]. This results in collagen-stimulated DDR1 activation contributing to pathways like the NF-κB pathway by ShcA and subsequent activation of TRAF6 promoting anti-apoptotic/pro-survival signals, or the PI3K/Akt pathway by direct phosphotyrosine interaction on the kinase domain[126]. DDR1 can also directly activate or regulate

the Ras/Erk MAPK signaling cascade in a cell type dependent manner[108]. Through interactions with Shp-2 and DDR1 activity stimulating MMP-1, -2, and -7 expression and activation, DDR1 promotes ECM remodeling and cell migration (Figure 1.7)[105,127,128]. There is also significant crosstalk between DDR1 and the integrin receptors through the protein Bim-1. DDR1 signaling can either counteract or cooperate with integrin signaling, resulting in different outcomes. In cooperative signaling, DDR1 can maintain precursor cells in an undifferentiated state, which is useful during early embryogenesis. When counteracting integrin signaling, DDR1 promotes epithelial cell differentiation (Figure 1.7)[129]. Further crosstalk between integrins and DDR1 promote cell adhesion to basement membranes or to surrounding cells. DDR1 also plays a critical role TGF-β signaling through activation of DDR1 by Wnt5a signaling and collagen stimulation (Figure 1.7)[130,131]. This is necessary for proper mammary gland development in adult tissue[132].

For DDR2, less is known about its intracellular signaling partners. Recent studies have revealed that only one adaptor protein, ShcA, interacts with activated DDR2 and several downstream effectors, such as Shp-2, Nck1, Lyn, PLCL2 (phospholipase C-like 2), and phosphatidylinositol-4-phosphate 3-kinase[113,117,122]. There is insufficient data to suggest whether any of these interactions are directly with DDR2 or through other proteins. Similar to DDR1, it is hypothesized that DDR2 can promote ECM remodeling and cell migration by stimulating increased expression and activation of specific MMPs and activating Shp-2. A recent study has shown that DDR2 can be activated by insulin receptor activation; however the mechanism by which this happens has not been fully explored[122,133].

Knockout mouse models of the DDRs have also given insight into the role of DDR signaling in development. DDR1 -/- mice were runted compared to normal mice and experienced

aberrant morphology and physiology[105,113]. Almost 80% of the DDR1 -/- female mice were unable to give birth due to blastocysts not being able to implant[134]. Those that were able to give birth were unable to nourish their pups due to mammary gland epithelium defects. Further problems were exhibited in the kidneys of older DDR1 -/- mice where the glomerular basement membrane was swollen, leading to increased serum proteins in their urine (referred to as proteinuria)[113,135]. Defects to calcification of bones were also found in DDR1 -/- mice. DDR2 -/- mice developed a rare form of dwarfism, relating to issues with bone growth. This dwarfism is identical to spondylo-meta-epiphyseal dysplasia, a rare genetic disorder found in humans that is the caused by mutations to the DDR2 gene[136]. DDR2 -/- mice also had significantly impaired wound healing[137].

1.5.5 DDRs in disease

It is not very surprising that like other PTKs, the DDRs have been implicated in a number of diseases. Both DDR1 and DDR2 are overexpressed, activated, or inactivated in a number of cancers, fibrobrosis, atherosclerosis, and osteoarthritis[108].

DDR1 has been implicated in a number of metastatic cancers, including breast cancer, non-small cell lung cancer, gliomas, lymphoma, hepatocellular carcinomas, pancreatic cancer, prostate cancer, and ovarian cancer[108]. Interestingly, DDR1 mutation or overexpression is not sufficient to cause cancer and most likely requires other oncogenic mutations to drive tumorigenesis and metastasis[138]. There have also been disease models demonstrating that DDR1 contributes to the pathology of atherosclerosis, as well as lung and kidney fibrosis[139-143]. In all

instances, DDR1 expression and activity is pro-inflammatory. Intriguingly, in atherosclerosis, DDR1 activation is anti-fibrotic; however in kidney fibrosis is pro-fibrotic.

DDR2 has also been implicated in squamous cell lung cancers and breast cancer metastasis[138,144]. Like DDR1, mutation or overexpression of DDR2 is not sufficient to drive cancer and requires additional oncogenic mutations for tumorigenesis and metastasis. Furthermore, DDR2 has been implicated in liver fibrosis. Unlike DDR1, which promotes fibrosis, DDR2 activation is anti-fibrotic and promotes hepatic stellate cell invasion and proliferation[108]. There is also evidence that DDR2 contributes to osteoarthritis by being overexpressed in damaged cartilage[108]. This overexpression induces the expression of certain MMPs, which contribute to the pathology of the disease.

1.5.6 DDR1 promiscuously binds small molecule kinase inhibitors

Recently, an exhaustive screen of small molecule kinase inhibitors against a large panel of human kinases has shown that some kinases bind inhibitors more promiscuously than others (kinases that were classified as promiscuous bound more than 8 inhibitors with a $K_D \leq 100$ nM)[145]. Of these promiscuous kinases, the majority were disease relevant kinases and intended targets of the inhibitors screened, such as PDGFR, p38α, Abl, and VEGFR. Interestingly, DDR1 was also amongst these promiscuous kinases. Of the 38 inhibitors screened, DDR1 bound 11 of them with affinity greater than 100 nM. This was an unexpected finding since none of the small molecule inhibitors were designed for targeting DDR1 and DDR1 has not been established as a medically interesting target for disease treatment. Unexpectedly, DDR1 binds several of these kinase inhibitors with greater affinity than their intended targets. This is best exemplified by

DDR1 being the tightest binding kinase for imatinib, dasatinib, and nilotinib; three clinically used kinase inhibitors designed for inhibiting BCR-Abl in CML treatment[123]. As a result, when treating diseases like cancer with small molecule kinase inhibitors, normal cells may be affected due to DDR1 readily binding these inhibitors, leading to undesirable side effects[146]. Therefore, it is important to understand how DDR1 promiscuously binds these inhibitors.

1.6 Small molecule inhibition of protein kinases

Protein kinases have become the most targeted group for drug development[3,4]. Their role in diseases like cancer have made them an attractive target and specific inhibition of them offers the benefits of lesser side effects compared to using DNA-alkylating agents as chemotherapeutics[146-149]. However, developing specific kinase inhibitors is a challenge. Most small molecule inhibitors target the ATP binding site, a highly conserved, deep hydrophobic pocket in the active site of all kinases[4]. Due to the common use of ATP by all PTKs and the millimolar amounts of ATP in the cell, small molecule inhibitors have to bind more tightly than ATP and bind specifically to the kinase it's targeting. Developing specific inhibitors that target other sites on the kinase have potential for specific inhibition, such as targeting the protein substrate binding site and allosteric sites. The best example of protein substrate competitive inhibitors was demonstrated by the tyrphostins, which were specific but lacked potency in cell[150,151]. Allosteric inhibitors offer the opportunity for being highly specific since, such as the compounds GNF-2 and -5 for Abl kinase or CI-1040 for MEK kinases, however there are few known allosteric sites[152-154].

ATP competitive inhibitors are classified based on the kinase conformation they bind to[148]. Type I inhibitors, such as dasatinib, bind the active kinase conformation (Figure 1.3)[148,155]. Inhibitors that bind to the Abl/c-Kit like-inactive conformation, such as imatinib, are classified as Type II inhibitors (Figure 1.3)[148,156,157]. In the Abl/c-Kit-like inactive conformation, a hydrophobic pocket underneath the αC-helix that is typically occupied by the phenylalanine of the DFG motif is exposed[156]. Type II inhibitors bind to both the ATP binding site and the exposed hydrophobic pocket. Due to low conservation of this hydrophobic pocket, Type II inhibitors can be tailored to target a select few kinases. Compared to Type II inhibitors, Type I inhibitors are typically less selective since they only target the highly conserved ATP binding pocket[148].

ATP competitive small molecule inhibitors can be highly effective in the clinic. The best example of this is the BCR-Abl inhibitor imatinib[158,159]. Imatinib increased the 5 year survival rate of BCR-Abl driven CML patients from 20% to almost 90%[160]. Furthermore, imatinib only inhibits a select amount of kinases (DDR1, PDGFR, c-Kit, and VEGFR), making it still one of the more specific inhibitors clinically used[123,161-163]. Since the success of imatinib, pharmaceutical companies have attempted to develop specific small molecule kinase inhibitors[3]. Success in developing specific small molecule kinase inhibitors has been achieved for other kinases, such as EGFR, VEGFR, Her2, and JAK. As of 2013, there are currently 23 clinically approved small molecule inhibitors for treating not only cancer, but other diseases such as myelofibrosis, rheumatoid arthritis, and cerebral vasospasm in humans[4].

Unfortunately, kinase inhibitors also have to overcome drug resistance mutations[164,165]. Using the example of BCR-Abl, patients can potentially develop mutations that render the kinase resistant to imatinib[166-168]. The most common form of imatinib resistance is the mutation of the

gatekeeper threonine[165,169]. This conserved threonine regulates access to a hydrophobic pocket adjacent to the ATP-binding pocket. In cancer, the gatekeeper threonine is mutated to an amino acid with a bulkier hydrophobic sidechain, such as isoleucine, which blocks imatinib binding[165]. Besides the gatekeeper threonine, mutations in the P-loop often confer resistance in BCR-Abl to imatinib[164,165]. Inhibitors that could inhibit imatinib resistance mutations were needed and led to the creation of 2nd generation BCR-Abl small molecule inhibitors dasatinib, and later nilotinib[169,170]. The compounds offered increased potency but had reduced specificity compared to imatinib[161,171]. Similar scenarios are seen in all clinically targeted kinases, such as EGFR and VEGFR, and have driven the need for the development of inhibitors that can overcome inhibitor resistance mutations and affect as few kinases as possible[165].

While these small molecule inhibitors have had much therapeutic benefit in the clinic, and target a select few kinases, there are undesirable effects by the inhibition of off-target kinases seen in many patients[146]. Though most kinase inhibitors are well tolerated clinically compared to cytotoxic chemotherapies, off target effects affect a patient's quality of life following treatment. Imatinib, for instance, is associated with severe heart failure. Nilotinib can cause pancreatitis, and cardiovascular defects. In some patients, dasatinib may cause gastrointestinal bleeding and pericardial effusions[146]. Many of these side effects are caused by the inhibition of off target kinases in normal cells, such as promiscuous kinases like DDR1[146]. There is a need for the development of more specific inhibitors, and also a need to understand how promiscuous kinases bind a variety of inhibitors.

1.7 Using Src and DDR1 as model systems for understanding the kinase-inhibitor specificity relationship

For this thesis project, I will use Src and DDR1 kinases as models for studying how specific kinase inhibition of Src can be achieved and why DDR1 binds promiscuously to small molecule inhibitors.

Highly specific macrocyclic peptide inhibitors of Src kinase were discovered by Dr. David Liu and did not inhibit the other SFKs[172]. Although inhibitors like dasatinib inhibit Src and the SFKs, there are currently very few inhibitors that can distinguish between Src, SFKs and Abl[173]. Furthermore there is substantial evidence that Src inhibition may be beneficial in the treatment of several cancers[90]. By examining how these compounds achieve their specificity and assess whether the compounds are efficacious in cells, new kinase inhibitors may be developed following the model of how these compounds target Src specifically.

Kinase promiscuity for small molecule inhibitors can lead to undesirable side effects when treating diseases like cancer[146]. To better understand how kinases are able to bind a wide variety of inhibitors, I will examine how DDR1 promiscuously binds small molecule kinase inhibitors. Although DDR1 has been implicated in a number of diseases, and its coincidental inhibition by clinically used inhibitors may be well tolerated, there are still side effects to kinase inhibitor therapies that do affect the patient's quality of life during treatment. It would be beneficial to understand why DDR1 is promiscuous and use this information for the development of more specific kinase inhibitors.

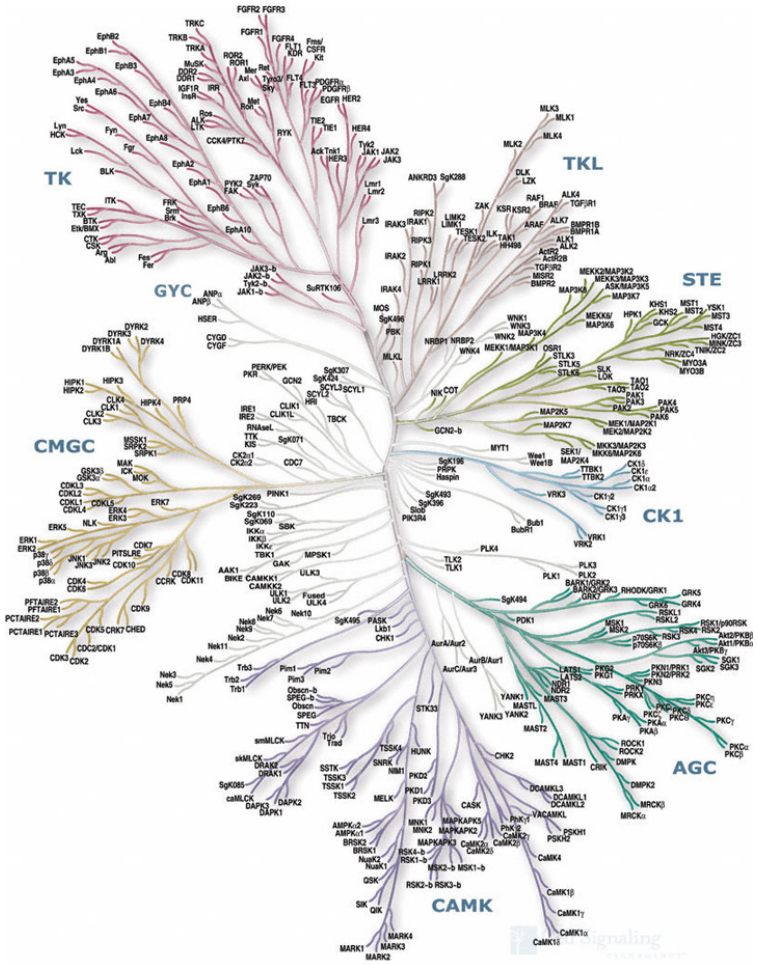


Figure 1.1 – The human genome encodes for 538 protein kinases. The kinases are classified based on whether amino acid they phosphorylate and sequence homology. There are 384 S/TKs and 80 TKs. Adapted from Manning G et al., 2002[10].

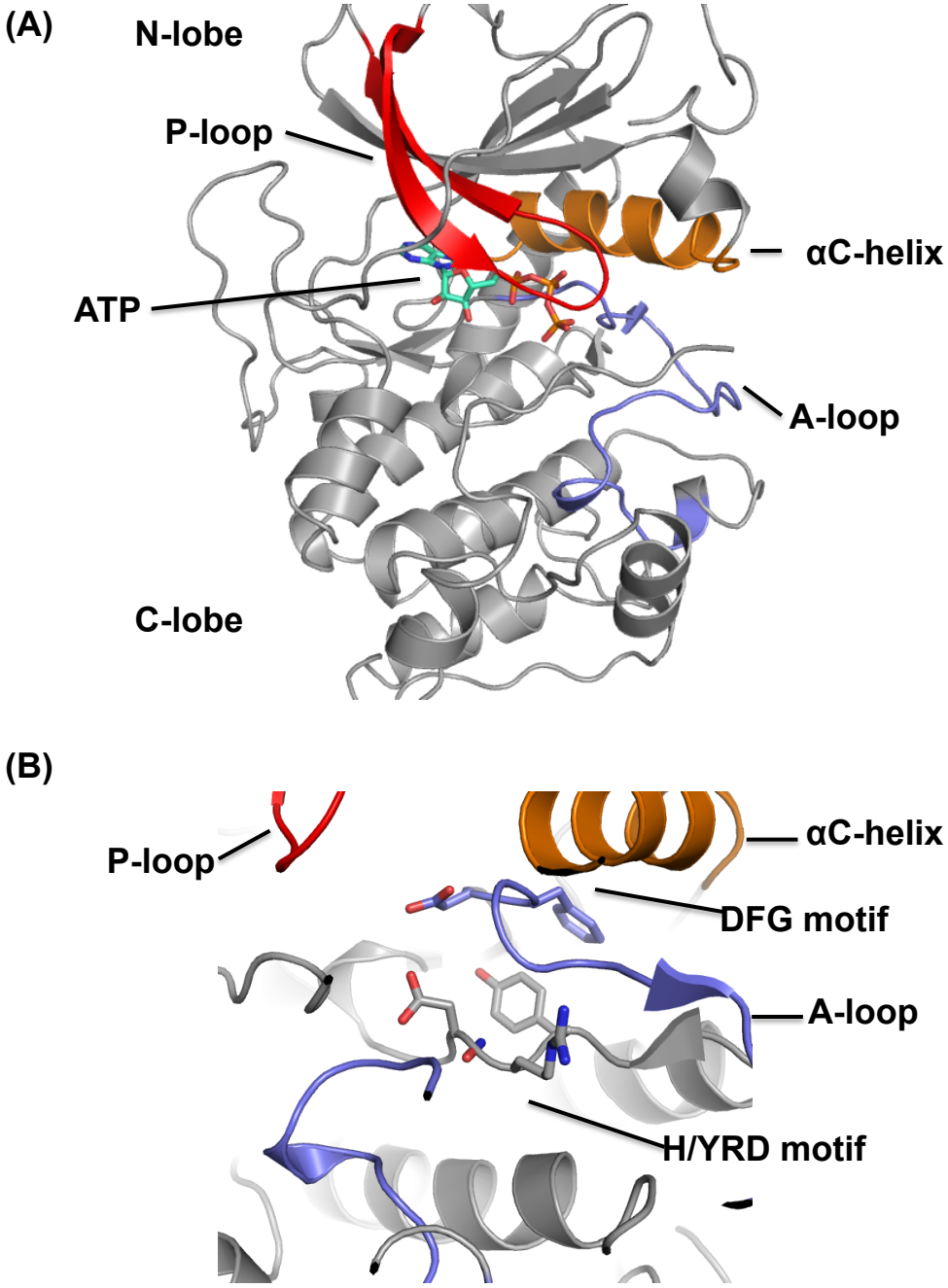


Figure 1.2 – The structurally conserved kinase domain. (A) Bi-lobal structure of the kinase domain. The P-loop is highlighted in red, the αC-helix in orange, and the A-loop in blue. ATP is located in the active site and is hydrogen bonds with the P-loop. (B) The H/YRD and DFG motifs are highly conserved amongst all protein kinases. Figure was generated using PDB entry 1ATP[24].

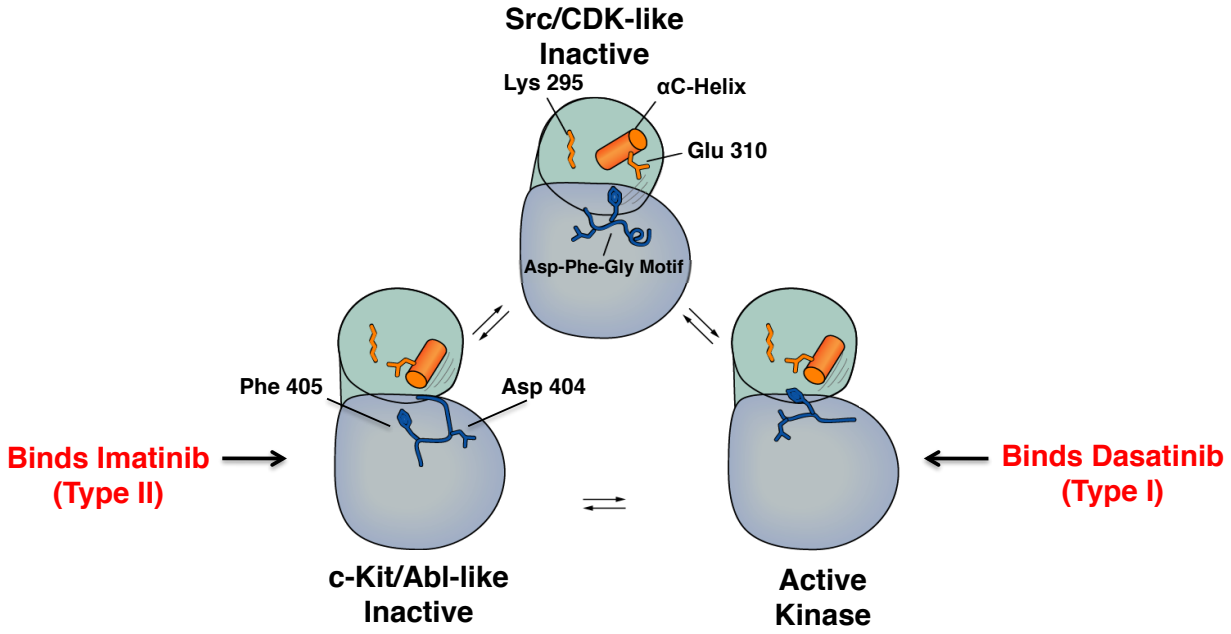


Figure 1.3 – Kinases can adopt multiple conformations. There is one active conformation and two inactive conformations that kinases can adopt. ATP competitive small molecule kinase inhibitors that can bind the active conformation are referred to as Type I inhibitors. Type II inhibitors bind the Abl/c-Kit-like inactive conformation. Adapted from Seeliger MA, et al. (2007)[174].

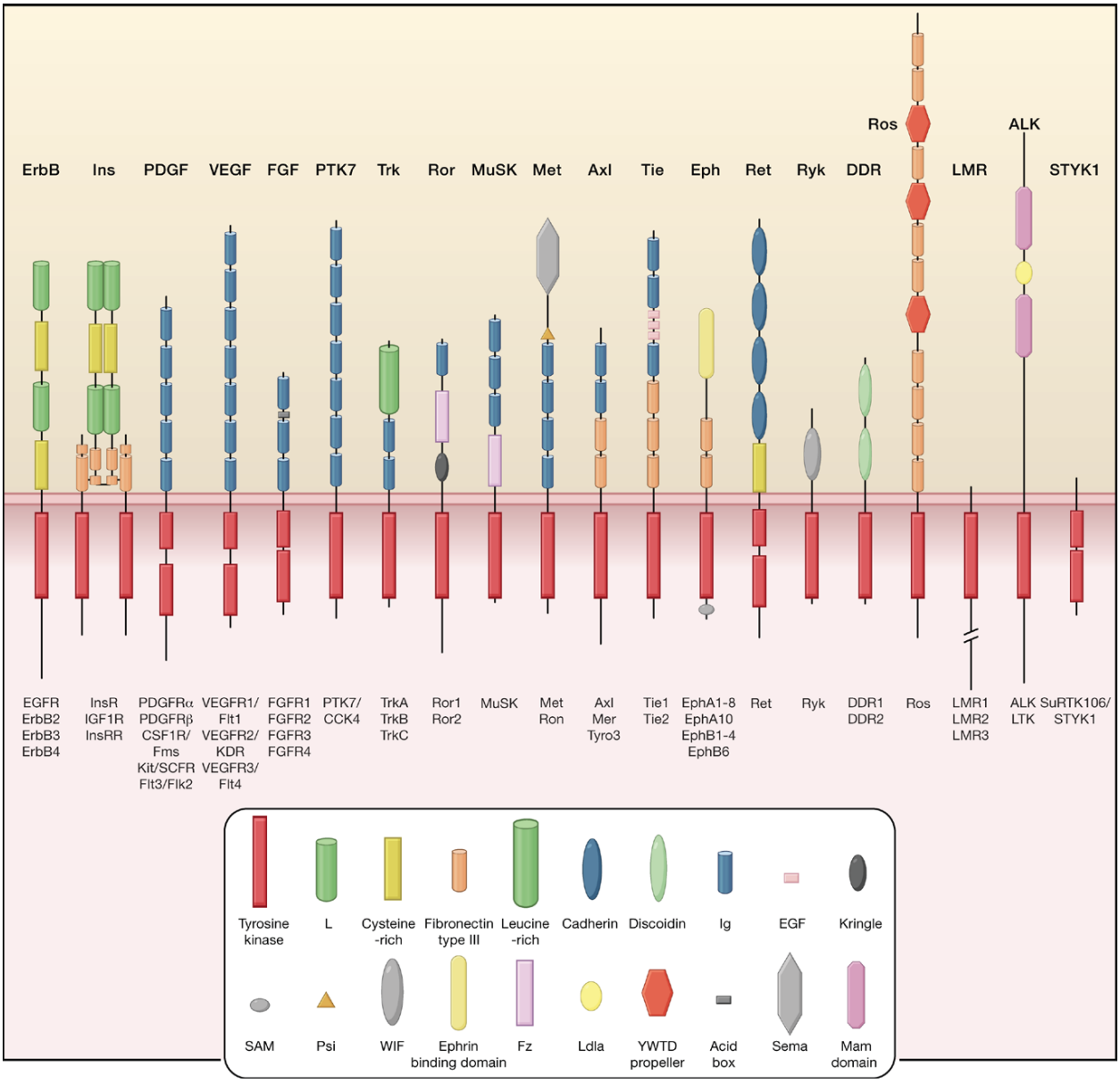


Figure 1.4 – Receptor tyrosine kinases families and domain architecture. Adapted from Lemmon MA and Schlessinger J (2010)[15].

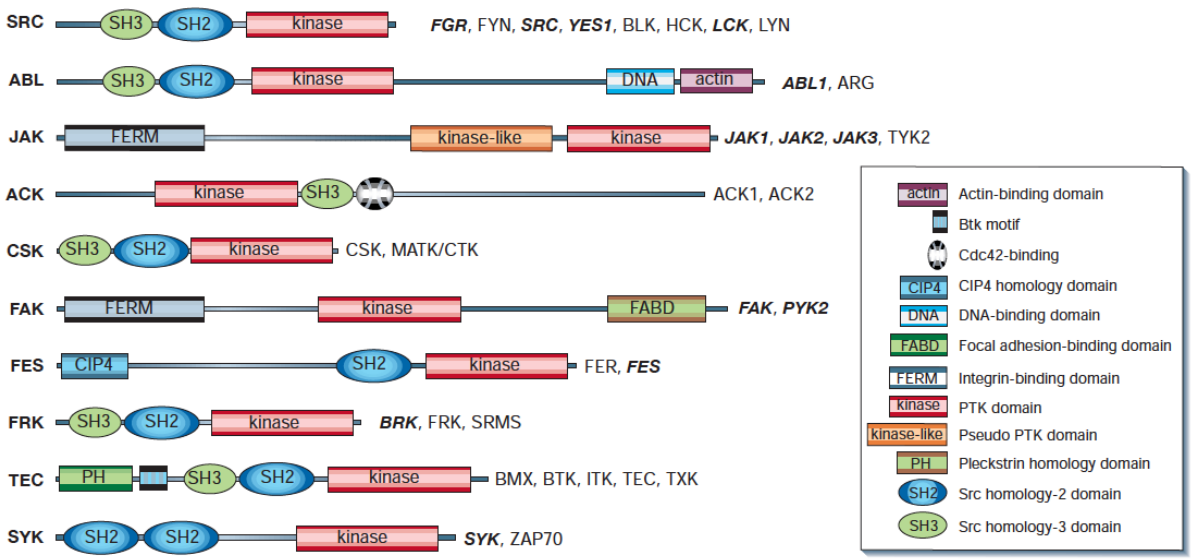


Figure 1.5 – Cytoplasmic tyrosine kinase families and domain architecture. Adapted from Blume-Jensen P and Hunter T. (2001)[2].

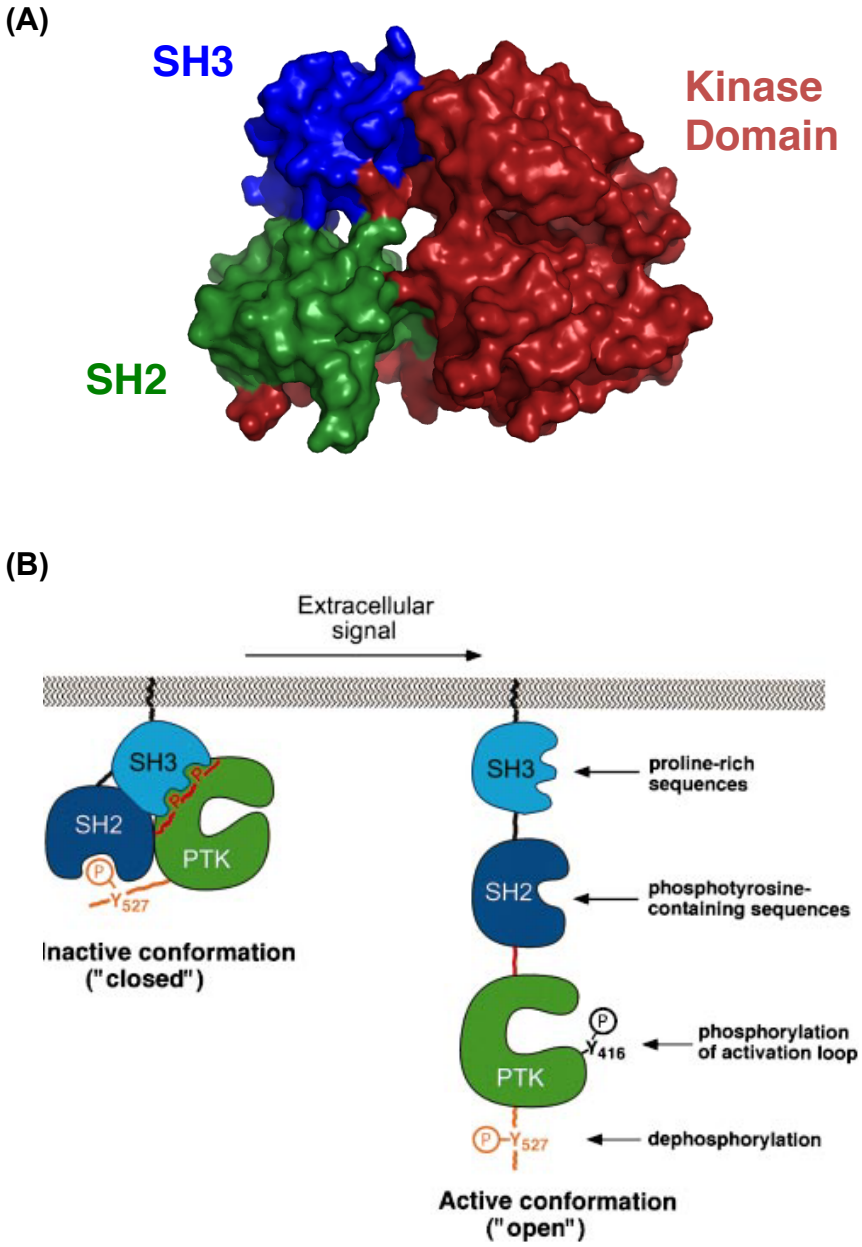


Figure 1.6 – Src family kinases have a similar domain architecture that results in conserved mechanisms of autoregulation and activation. (A) X-ray protein crystal structure of inactive Src_{83-533} (PDB entry 2SRC)[43]. (B) The SH2 and SH3 domains lock Src in an inactive conformation through two interactions: binding of the phosphorylated C-terminal tail of the kinase to the SH2 domain and binding of the poly-proline motif in the SH2-kinase domain linker to the SH3 domain. When Src is activated, the SH2 and SH3 domain release the kinase domain. Adapted from Hubbard SR et al. (1998)[38].

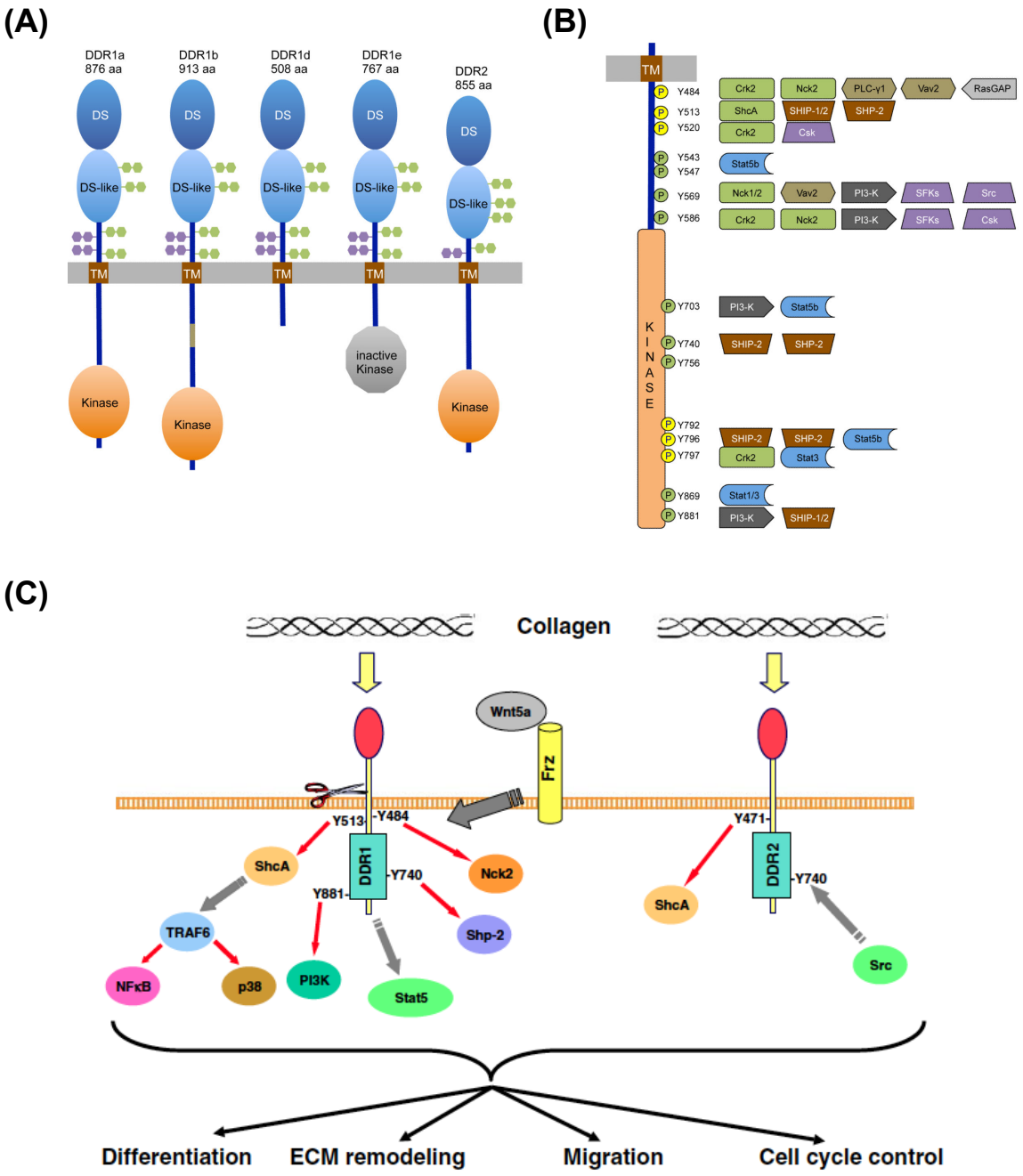


Figure 1.7 – DDRs are a unique family of RTKs that bind extracellular collagen. (A) Isoforms of DDR1 and DDR2. There are 5 isoforms of DDR1 (a, b, c, d, and e) and only 1 of DDR2. DDR1a and b are the most commonly expressed isoforms of DDR1. DDR1d and e are inactive isoforms. Adapted from Leitinger B (2014)[105]. (B) Phosphotyrosine along the intracellular juxtamembrane region and kinase domain mediates DDR1 interactions with adaptor and signaling proteins that contain SH2 or PTB domains. Adapted from Leitinger B (2014)[105]. (C) DDR1 and DDR2 signaling pathways. Adapted from Vogel, W et al. (2006)[102].

Chapter 2: Methodology

2.1 Molecular genetics

2.1.1 Src, Hck, Lck, and Abl kinases

Kinase-domain constructs of human c-Abl (residues 229–512), chicken c-Src (residues 251–533 and residues 83–533), mouse Lck (residues 227–509), and human Hck (residues 166–445) were cloned as previously described[175,176]. The vectors for Src, Hck, Lck, and Abl kinase domain encoded for an N-terminal His6-tag, followed by a TEV cleavage site, and kanamycin resistance. Mutations were introduced into chicken c-Src (residues 251–533; Q275G, C277Q, E280V, L297M, T338I, and R419P) by site-directed mutagenesis and were verified by DNA sequencing.

2.1.2 DDR1

2.1.2A DDR1a constructs

The constructs of human DDR1a kinase domain (residues 526-876) and DDR1a kinase domain with intracellular juxtamembrane (residues 485-526) were amplified from a pDNR-Dual mammalian expression vector containing full length human DDR1a (residues 1-876; purchased from the ASU Gene Repository) and subcloned into the pFastBac HTb vector (Invitrogen) using BamHI and XhoI restriction sites located after the His6-tag and TEV protease cleavage site. Individual mutations were introduced into pFastBac HTb DDR1a$_{526\text{-}876}$ (D671N, Y755A, and Y759A) by site directed mutagenesis and verified by DNA sequencing.

Full length DDR1a (1-876) was subcloned from pDNR-Dual $DDR1a_{1-876}$ and inserted into the mammalian expression vector pEF1-V5-His A (Invitrogen) using BamHI and NotI restriction sites. The resulting construct contained C-terminal V5 and His6 tags, and was ampicillin resistant.

2.1.2B General protocol used for generation of $DDR1a_{485-526}$ and $DDR1a_{526-876}$ baculovirus

Cloning and baculovirus generation for the kinase domain of DDR1a utilized the Bac-to-Bac baculovirus expression system (Invitrogen). This system uses the pFastBac HTb vector that is kanamycin resistant and encodes for a N-terminal His6-tag and a TEV protease cleavage site. In order to generate the baculovirus for protein expression, pFastBac HTb containing the DNA sequence of the protein of interest must be transformed into the *Escherichia coli* DH10Bac strain. DH10Bac cells transpose the insert DNA from pFastBac HTb into a bacmid. These are grown on agar plates coated in 50 μg/mL kanamycin, 7 μg/mL gentamicin, 10 μg/mL tetracycline, 100 μg/mL X-gal, and 40 μg/mL IPTG for 48 hours at 37°C. Kanamycin, gentamicin, and tetracycline are needed to select for the bacmid, pFastBac HTb plasmid, and a helper plasmid needed for transpositioning the insert from pFastBac HTb to the bacmid, respectively. X-gal and IPTG are needed to conduct blue-white screening of DH10Bac cells.

Blue-white screening is used since the bacmid contains the *lac* operon with the *lacZ* gene in the middle of its transposition site. The gene encodes for β-galactosidase. β-galactosidase expression is induced in the presence of IPTG and the expressed enzyme breaks down X-gal, turning DH10Bac cells blue. If the insert from pFastBac HTb is transposed into the bacmid successfully, the *lacZ* operon is disrupted and the cells remain white in the presence of IPTG and

X-Gal. White colonies were picked and grown in 10 mL of LB with 50 μg/mL kanamycin, 7 μg/mL gentamicin, and 10 μg/mL tetracycline for 24 hours at 37°C. DH10Bac liquid cultures were spun down at 4000 G at 4°C in a swing bucket centrifuge. Bacmid DNA was purified using ethanol precipitation as described in the Bac-to-Bac expression system protocol. Bacmid DNA was reconstituted in water and quantified on the Nano-Drop by measuring O.D at 260 nm. Typical yields of bacmid DNA were between 5-8 μg/mL.

To generate baculovirus, *Spodoptera frugiperda* (Sf9) insect cells were plated out in a 6-well tissue culture dish at a confluence of 0.8 x 10^6 cells/mL per well and transfected with purified bacmid using Cellfectin. Transfection solutions were prepared by mixing 5 μL of bacmid with 100 μL serum-free SF900-II medium (Life Technologies) and 6 μL of Cellfectin reagent with 100 μL Sf900-II medium separately. The two mixtures were combined and incubated together for 15 to 45 minutes. 0.8 mL of serum-free Sf900-II medium was added to the transfection solution to raise the final volume to 1 mL. Sf9 cells were washed twice with serum-free media prior to the addition of the 1 mL transfection solution. The cells were incubated for 5 hours at 27°C. Following the incubation, the transfection solution was removed and 2 mL of serum free Sf-900 II media was added to the cells. Sf9 cells were then incubated for 72 hours at 27°C to generate P1 baculovirus.

After 72 hours, the medium was harvested from the Sf9 cells and spun down at 4000 G at 4°C in a tabletop centrifuge. The baculovirus was filtered sterilized using a 0.22 μm syringe filter. To generate the P2 virus, a 300 mL suspension culture of Sf9 cell at 0.8 x 10^6 cells/mL was infected with 10 mL of P1 virus and incubated for 5-7 days at 27°C. Virus was harvested by spinning cells down at 4000 rpm in an SLC-6000 rotor and filter sterilization through 0.45μm and 0.22μm vacuum filters. P3 virus (the virus used for protein production) was generated by

infecting a 500 mL Sf9 suspension culture at 0.8 x 10 cells/mL with 20 mL of P2 virus and incubated for 7-10 days at 27°C. P3 baculovirus was harvested and purified by the same method P2 baculovirus was harvested. Baculoviruses were stored at 4°C, and protected from light.

2.2 Protein expression and purification

Protein expression and purification described in this section is based off protocols established previously by Seeliger et al. and Day et al. for Src and DDR1 respectively[123,175]. For DDR1a$_{485-526}$ and DDR1a$_{526-876}$, expression of the kinase domain was attempted using the *E. coli* expression system described in Section 2.2.1. Unfortunately the expressed protein was either insoluble or inactive and as a result we used the Bac-to-Bac expression system discussed in sections 2.1.2A and 2.2.2.

2.2.1 Src, Hck, Lck, and Abl

All Src, Hck, Lck and Abl constructs were transformed into the *E. coli* BL21-DE3 strain that had been transformed previously with vectors for the tyrosine phosphatase YopH, the chaperone protein GroEL, and the adaptor protein Trigger Factor (this *E. coli* strain is referred to as BL21-DE3 YGT and was previously described by Seeliger et al.)[175]. Transformed cell cultures were grown in 1 L 2XYT medium with 50 μg/μL kanamycin at 37°C to an O.D. of 0.4 then grown at 16°C for 1 hour in a shaking incubator set to 200 rpm. After one hour, cells grew to an approximate O.D. of 0.8, and protein expression was induced using 1 mM IPTG. Cells were left shaking to express protein for 14-16 hours overnight at 16°C before harvesting.

Following overnight protein expression, cells were spun down at 4500 g for 5 minutes at 4°C in the SLC-6000 rotor. Cells were resuspended in an ice-cold buffer made up of 20 mM Tris pH 8, 20 mM Imidazole pH 8, 500 mM NaCl, and 5% glycerol. Resuspended cells were lysed by sonication for 30-second pulses over 5 minutes on ice. Lysates were subjected to centrifugation for 40-60 minutes at 20,000 G at 4°C using the SS-34 rotor. The supernatant of the spun down lysates was then loaded onto a GE HisTrap Fast Flow 5mL NiNTA column for purification by FPLC. The column was washed for 2 column volumes with a buffer containing 20 mM Tris pH 8, 20 mM Imidazole pH 8, 500 mM NaCl, and 5% glycerol. Protein was eluted in a gradient from 0-40% of elution over 5 column volumes with elution buffer (20 mM Tris pH 8, 500 mM Imidazole, 500 mM NaCl, and 5% glycerol). Eluted protein was dialyzed in 4 L of 20 mM Tris pH 8, 100 mM NaCl, 5% glycerol, and 1 mM DTT overnight at 4°C in the presence of TEV protease to cleave the His6-tag off the kinase in dialysis tubing with a molecular weight cut off of 3,000 Da. His-tag cleavage was verified by running the dialyzed protein on a 15% SDS-PAGE gel at 200 V for 1 hour and staining with Coomassie. The dialyzed protein was diluted 1:1 with cold, sterile filtered H_2O and loaded onto a GE Q Fast Flow 5 mL anion exchange column. The loaded column was washed for 2 column volumes using a wash buffer (20 mM Tris pH 8, 1 mM DTT, and 5% glycerol). The protein was eluted in a gradient from 0-40% over 5 column volumes with elution buffer (20 mM Tris pH 8, 1 M NaCl, 5% Glycerol, and 1 mM DTT). Final buffer conditions for the eluted proteins were assumed to be 20 mM Tris pH 8, 250 mM NaCl, 5% Glycerol, and 1 mM DTT. Protein purity was assessed by SDS-PAGE and Coomassie staining. Protein concentrations were determined by using the protein's extinction coefficient and the measured O.D. at 280 nm by Nano-Drop. Typical yields for Src kinase were between 20-60 mg

of protein per liter of *E. coli*. For Hck, Lck, and Abl, typical yields were between 5-15 mg per liter of *E. coli*. Proteins were snap frozen in liquid N_2 and stored at -80°C.

2.2.2 DDR1a$_{485-526}$ and DDR1a$_{526-876}$

His6-tagged DDR1a$_{485-526}$ and DDR1a$_{526-876}$ were expressed in *Spodoptera frugiperda* (Sf9) insect cells and purified following the Bac-to-Bac system protocol from Invitrogen. Sf9 cells were grown in Sf-900-II medium supplemented with 5% fetal bovine serum and 1x antibiotic/antimitotic solution. A 1 L culture of Sf9 cells at 0.8 x 10^6 cells/mL was infected with 30 mL of P3 virus and was incubated for 72 hours. Infected cells were centrifuged for 5 minutes at 3000 G at 4°C using the SLC-6000 rotor. Cells were resuspended in lysis buffer (50 mM Tris pH 8.5, 5 mM β-mercaptoehanol, 100 mM NaCl, 1 mM PMSF, and 1% NP-40) in a ratio of 5 mL lysis buffer per 1 g of cells. Lysates were centrifuged at 20,000 G for 60 minutes at 4°C using the SS-34 rotor. The resulting supernatant was loaded onto NiNTA resin on a gravity column in the cold room (4°C). The loaded resin was washed for 10 column volumes (1 column volume = volume of resin used; typically 1 mL of resin was used to purify 2 L of infected Sf9 cells) with a wash buffer (20 mM Tris pH 8.5, 20 mM imidazole, 500 mM NaCl, 5 mM β-mercaptoehanol, and 10% glycerol). Resin was subjected to a high salt wash to remove any additional protein contaminants in a buffer containing 20 mM Tris pH 8.5, 1 M NaCl, 5 mM β-mercaptoehanol and 10% glycerol followed by 2 column volumes of wash buffer. The protein was eluted off the resin by buffer comprised of 20 mM Tris pH 8.5, 125 mM NaCl, 200 mM Imidazole, 10% glycerol and 5 mM β-mercaptoehanol. The His6-tag of the protein was then cleaved by TEV protease digest overnight at 4°C. His6-tag cleavage was verified by running the

digested protein on a 15% SDS-PAGE gel at 200 V for 1 hour and stained with Coomassie. The cleaved protein was loaded on a GE 60/600 Superdex Gel Filtration Column and eluted in a buffer of 20 mM Tris pH 8.5, 125 mM NaCl, 20 mM Imidazole, 10% glycerol and 5 mM β-mercaptoehanol by FPLC. Concentration of protein was calculated using the extinction coefficient and the measured O.D. at 280 nm. The typical yields for $DDR1a_{485-526}$ and $DDR1a_{526-876}$ were approximately 2 mg of protein per liter of Sf9 cells at 0.8 x 10^6 cells/mL. Protein was then snap frozen in liquid N_2 and stored at -80°C.

2.3 X-ray protein crystallography

X-ray protein crystallography was used to determine the structures of Src and DDR1 described in chapters 3 and 4, respectively. This technique allows the study of protein-drug interactions on a molecular and atomic scale. These structures can used to determine molecular mechanisms for kinases, ranging from autoregulation to how they bind substrate peptide. The general strategy used for identifying crystallization conditions, data collection, structure determination, and model refinement is outlined below.

2.3.1 Crystal Screens

Crystallization conditions were determined using commercially available sparse matrix screens from Hampton Research. These 96-well screens (PEG/Ion, PEG/Rx, Salt/Rx, and Crystal Screen) contain varying crystallization conditions that differ in the amount of precipitant, salt, buffer, and additive solutions used to find ideal conditions for a protein crystal to form. As a

default, protein•drug complexes were formed at 10 mg/mL protein with 2-fold excess inhibitor, unless otherwise noted. The co-complex is mixed in a 1:1 ratio with mother liquor from the crystallization screens listed above in a 96-well crystallization plate using the sitting drop vapor diffusion method. Typical volumes used were 0.5 μL protein•drug to 0.5 μL mother liquor, and 75 μL of mother liquor in the well reservoirs of the crystallization plate. The plate was sealed with clear tape and incubated overnight at room temperature. Screens were checked daily for protein crystals. In the conditions where protein crystals formed, 24-well optimization screens were designed around the initial condition.

In the optimization screens, ranges of precipitant, salt or buffer were screened to further improve crystal quality. In most instances, crystals from the 96-well screens were clustered. Ideal crystals are individual, and well defined three dimensionally. The 24-well screens also used higher volumes of protein•drug complex (1 μL instead of 0.5 μL), and mother liquor (1 μL to be mixed with the protein•drug complex, 500 μL in the reservoir). As a result, the time needed for the well to reach equilibrium is greater and can affect the rate of crystal growth. The optimization screens were done in successive rounds. In each round the range of precipitant, salt buffer, and additive screened was smaller than the last. This was done until individual crystals are grown.

If individual crystals could not form, or were too small, micro seeding was used to improve crystal quality. Seeds for micro seeding were generated by crushing pre-existing crystals of the same protein (the ligand bound to the protein does not to be the same) until they were no longer visible. The seeds form nucleation points for crystals to grow from. This requires the use of less precipitant in the optimization screens. Seeds were added to the mixture of

protein•drug complex with mother liquor by touching a cat whisker into the seed solution and touching it to the drop.

When ideal crystals were grown, they were harvested in mother liquor containing 20% cryoprotectant for storage in liquid N_2. For our studies, glycerol and ethylene glycol were screened as cryoprotectants to see which would result in better diffraction during data collection.

2.3.2 X-ray diffraction and data collection

X-ray diffraction data were collected at the National Synchrotron Light Source at Brookhaven National Laboratories beamline x29. Data for all protein-drug complexes were collected at 100K and 1.075 Å wavelength.

2.3.3 Data processing and model building

Data were processed using indexing and scaling software suites such as HKL2000, iMosflm, and autoPROC[177-179]. The indexing software is designed to determine the space group (symmetry pattern) of the crystal, as well as the size of the unit cell. Scaling software integrates and scales the data by determining the signal to noise ratio (I/σ) of the data, and scales the entire data set according to the I/σ and space group. Data quality was assessed using Xtriage in the Phenix crystallography suite[180].

All protein structures solved in this thesis were done by molecular replacement using Phaser[181]. Multiple search models were used and the best scoring search model was used for refinement. The model of the structure was built in Coot[182] and refined in Phenix[180]. The

geometric quality of the refined model was assessed with MolProbity[183] and the structure validation tools in the Phenix suite[180].

2.4 Biochemical Assays

2.4.1 Src activity and IC_{50} assays

In vitro kinase inhibition assays were performed using a continuous spectrophotometric assay. This assay is a coupled kinase assay that measures kinase activity coupled to the oxidation of NADH to NAD^+ by measuring NADH absorbance at 340 nm[184]. ATP is consumed as the kinase phosphorylates its substrate peptide, resulting in the formation of ADP. ADP is regenerated to ATP by pyruvate kinase transferring the phosphate of phosphoenol pyruvate onto ATP. The resulting pyruvate is a substrate for lactate dehydrogenase, which oxidizes NADH to NAD^+ to convert the pyruvate to lactate. As a result, the consumption of ATP is in a 1:1 ratio with the oxidation of NADH. This assay can be adapted for multiple kinases and a variety of substrate peptides[184]. The assays were performed in 96-well ELISA plates and read using a plate reader.

For the Src inhibition assays,[184] 100 μM of a Src-optimal substrate peptide (AEEEIYGEFAKKK)[185,186] were combined with 5 μM ATP for MC1, MC2, and MC9 (see chapter 3). Concentrations of kinase used for these assays were as follows: 0.125 μM for Src kinase domain, 0.33 μM for $Src_{83\text{-}533}$, 0.33 μM for Src Q275G, 0.4 μM for Src C277Q, 0.8 μM for Src L297M, 0.4 μM for Src E280V, 0.042 μM for Hck, 0.25 μM for Lck, and 0.5 μM for Abl. For MC4b, and MC25b, the assay was performed using 250 μM ATP and 300 μM Src-optimal

peptide. The concentration of kinase used in these assays were as follows: 0.0125 μM for Src kinase domain, 0.1 μM for Src_{83-533}, 0.033 μM for Src Q275G, 0.02 μM for Src C277Q, 0.04 μM for Src L297M, 0.02 μM for Src E280V, 0.0042 μM for Hck, 0.025 μM for Lck, and 0.05 μM for Abl. Titrations of MC1, MC2, MC4b, MC9, and MC25b (ranging from 0 μM to 83.3 μM) were performed at 30 °C as described before for imatinib[176] to determine the concentration at which 50% of the initial kinase activity is inhibited (IC_{50}).

2.4.2 DDR1a activity assays

The continuous assay described above was not sensitive enough to detect DDR1 kinase activity. Therefore, *in vitro* kinase assays for $DDR1a_{526-876}$ were done by measuring substrate peptide phosphorylation using [γ-^{32}P] ATP in a phosphocellulose paper binding assay[187]. This assay has higher sensitivity that can measure the low kinase activity of DDR1. Reactions were done in 20 mM Tris pH 7.4, 10 mM $MgCl_2$, 400 μM ATP, 200 μM Axltide (KKSRGDYMTMQIG), 0.5 μM $DDR1a_{526-876}$ or $DDR1a_{485-876}$, and 50-100 cpm/pmol [γ-^{32}P] ATP. Reaction mixtures were incubated for 60 minutes at 30°C. The reactions were stopped by the addition of TCA and centrifuged to separate the precipitated kinase from the soluble substrate peptide. Supernatant was blotted onto phosphocellulose paper and washed three times with phosphoric acid. The phosphocellulose paper was dried and read on a scintillation counter.

2.5 Binding Assays

Dissociation constants (K_D) can be used to assess how tightly a ligand can bind to a protein. In chapter 3, binding assays were used to determine how well the macrocycle inhibitors compete with ATP and substrate peptide for the active site of Src to demonstrate the bi-substrate competitive nature of the compounds. To do this, the change in fluorescence anisotropy of fluorescein labeled macrocycles was measured as they bind to Src kinase. In principle, this works by measuring the change of the rotational rate of the fluorescein-labeled macrocycles in solution. As they bind to the kinase, the rate of rotation decreases since the macrocycle is bound to a larger molecule. The rate of rotation is slowed since the kinase is significantly larger than a macrocycle and will therefore spin slower than the unbound macrocycles in solution.

2.5.1 Affinity of macrocyclic peptide inhibitors for Src kinase

The change in fluorescence anisotropy of fluorescein labeled MC1, MC2 and MC9 at 518 nm upon excitation at 492 nm was monitored with a HORIBA Jobin Yvon FluoroMax-4 (Edison, NJ) spectrofluorimeter. Src kinase domain (residues 251-533) was titrated to 0.5 μM of the fluorescein-labeled macrocycle, in 100 mM Tris pH 8.0, 10 mM $MgCl_2$ at 25°C. Src_{83-533} was titrated to 0.05 μM fluorescein-labeled macrocycle in 100 mM Tris pH 8.0, 10 mM $MgCl_2$ at 25°C. After equilibratrion, the increase in the fluorescence anisotropy of the fluorescently labeled ligand was recorded and fitted against a quadratic binding equation in Kaleidagraph (Synergy Software, Reading PA) to yield the dissociation constant (K_D).

2.6 Cellular Assays

Cellular experiments are an excellent way of determining how meaningful *in vitro* data is. Since *in vitro* experiments are focused around purified proteins, the impacts of kinase inhibition and kinase mutations are limited to the biochemistry observed. By examining them in the context of the cell, the effects of how particular inhibitors or mutations affect signaling pathways can be studied. This data can be translational to the clinic, showing how effective an inhibitor may potentially be in treating a particular disease or how a mutation drives a disease.

2.6.1 Src specific macrocyclic peptide inhibitors efficacy in cell

The in cell potency of the macrocycles has yet to be accurately determined. To do so, HEK 293T cells were used to determine the efficacy of the macrocycle inhibitors on inhibiting autophosphorylation of endogenous Src in non-transformed cells. By determining the efficacy on inhibiting Src in normal cells, a baseline could be developed to compare how efficacious the macrocycles may be in treating cancer lines where Src signaling is dysregulated.

HEK 293T cells were grown in Dulbecco's modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum at 37°C in 5% CO_2. To test inhibitor efficacy, we treated the cells with 5 μM dasatinib, 20 μM MC4b, 20 μM MC25a or a DMSO vehicle. Following treatment, cells were collected and lysed, and run on SDS-PAGE. Following SDS-PAGE, lysates were transferred onto nitrocellulose paper and blocked with 5% BSA for 1 hour at room temperature. Lysates were probed via Western blot for autophosphorylated Src using Src phospo-Y418 antibody (Biosource) and total Src using GD11 Src specific antibody (Upstate).

2.8.2 DDR1a Cellular Expression and Activation

DDR1 is an unusually slow RTK involved in many cell signaling pathways[102]. In an effort to better understand the role of DDR1 signaling and its slow activation in the cell, an assay was designed that would examine how mutations to the kinase domain of DDR1 affect its activity and the phosphorylation of downstream targets. A recent study showed how DDR1 signaling regulates megakaryocyte migration[188]. Using this signaling pathway and cell migration as a medium for measuring DDR1 activity (signaling pathway is described in chapter 4), a cellular system was established that could be used to study the impacts of DDR1 mutations on both cell migration and DDR1 activation. The experiment described below is the preliminary experiment that will be used to establish this assay.

pEF1-V5-His A DDR1a$_{1\text{-}876}$ was transfected into the UT7 human megakaryocyte cell line grown in a medium of Dulbecco's modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum and thrombopoietin at 37°C in 5% CO_2. To test DDR1a activation, transfected UT7 cells were treated with 20 μg/mL of collagens I and IV. Stimulated cells were collected at 0, 2, 16, and 24 hours. Cells were lysed and DDR1a was purified from lysates by pull down with a V5-tag antibody. The purified DDR1a was run on an SDS-PAGE gel then transferred to a nitrocellulose membrane. The membrane was probed for autophosphorylated DDR1 using an anti-phosphotyrosine antibody via Western blotting.

Chapter 3: Determining the molecular mechanism of Src specific macrocyclic peptide inhibitors

This chapter has been adapted from Georghiou G*, Kleiner RE*, et al. Nature Chemical Biology, (2012)

*These authors contributed equally to this work

Work in this chapter is credited to several people:

George Georghiou was responsible for protein purification, generation of Src kinase mutants, biochemical/biophysical characterization of the MC compounds, and protein crystallography.

Dr. Ralph E. Kleiner was responsible for generation of the Src specific MC compounds and the Z-LYTE assay data reported in Table 3.1.

Michael Pulkoski-Gross was responsible for MC binding studies.

Jevon Fragoso was responsible for determining the IC_{50} of the MC compounds for Src^{R419P}.

Saadat Aleem was responsible for the cellular potency of the MC compounds in HEK 293T cells.

Kip E. Guja was responsible for crystallographic data processing for the Src•MC25b structure.

Dr. W. Todd Miller (mentor of S. Aleem) advised on the cellular potency of the MC compounds in HEK 293T cell experiments.

Dr. David R. Liu (mentor of R.E. Kleiner) advised on and reviewed all data collected for the publication in Nature Chemical Biology.

Dr. Markus A. Seeliger (mentor of G. Georghiou, M. Pulkoski-Gross and J. Fragoso) advised on and reviewed all data collected for the publication in Nature Chemical Biology.

3.1 Introduction

Because protein kinases have a central role in cell signaling, the discovery and development of protein kinase inhibitors have been the focus of intensive research since the late 1980s[4,189]. Currently, all clinically–approved small-molecule inhibitors of protein kinases target the binding site for ATP, the common substrate in the enzymatic reaction of all 538 human kinases[10,11]. The high sequence conservation within the kinase ATP-binding pocket, however, makes the development of specific inhibitors of protein kinases challenging. When such specificity is achieved, kinase inhibitors can become effective drugs, such as the c-Abl kinase inhibitor imatinib, which treats chronic myelogenic leukemia[158,190,191].

Non–ATP competitive inhibitors have the potential to show excellent specificity because of the more varied nature of their binding sites among kinases. Allosteric inhibitors, such as GNF-2 (which targets Abl kinase) and CI-1040 (which targets MEK kinases), have been shown to be potent and highly specific[152,154,192]. Likewise, substrate peptide–competitive inhibitors such as tyrphostins (which target EGFR kinase) have the potential to be highly specific because of the sequence variation in the substrate peptide–binding sites of protein kinases[193,194].

In addition to inadequate specificity, vulnerability to drug-resistance mutations is another major problem facing the use of kinase inhibitors as therapeutics[195]. One of the most common imatinib resistance mutations in Abl kinase, Thr315Ile, confers resistance to other available kinase inhibitor therapeutics[196]. This threonine (Thr338 in chicken c-Src) regulates access to a hydrophobic pocket adjacent to the ATP-binding pocket and is therefore often referred to as the gatekeeper. Replacement of the gatekeeper residue in Src, Abl, PDGFR and EGFR kinases with hydrophobic residues increases kinase activity and leads to the transformation of Ba/F3 cells[197].

Most small-molecule kinase inhibitor discovery efforts rely on combinatorial or diversity-oriented synthesis and high-throughput screening. Dr. David Liu of Harvard University developed DNA-templated synthesis as a method of translating DNA sequences into small-molecule libraries that can be directly subjected to *in vitro* selections for desired properties, including target affinity[198-201]. Selection-based approaches are typically much more efficient than screening because they enable the simultaneous evaluation of all library members in one experiment regardless of library size, obviating the time and infrastructure demands of screening[202].

Recently, the Liu lab reported the synthesis and selection of a 13,824-membered DNA-templated macrocyclic peptide library[172,203]. The macrocyclic peptides generated were chemically diverse, containing different macrocyclic backbones, and three variable building blocks that are referred to as the A, B, and C groups. The building blocks were composed of both natural and unnatural amino acids, such as phenylalanine and cyclopropylalanine. They identified from this library a series of macrocycles that inhibit Src with half-maximum inhibitory concentration (IC_{50}) values as potent as 680 nM. Two of these macrocyclic compounds (MC2 and MC9) showed a remarkable level of specificity, inhibiting Src kinase, but not Abl kinase or closely related Src-family kinases, including Hck (Figure 3.1 and Table 3.1)[172]. Through structure activity relationships (SAR), they determined that the inhibition of Src by the macrocycles was dependent on the macrocycle building blocks, as well as the conformation of the macrocyclic backbone.

In this study, we developed second-generation macrocycles based on MC2 and MC9 with potencies as high as $IC_{50} \leq 4$ nM. The most potent macrocycles inhibit Src kinase activity in mammalian cells. We determined the inhibition mechanism of these inhibitors and the structural

basis of their unusual specificity by solving the X-ray co-crystal structures of three macrocycles bound to the Src kinase domain. Our studies reveal that the macrocycles force the kinase to adopt the Src/CDK-like inactive conformation. In addition, the inhibitors occupy the ATP-binding pocket of the kinase and simultaneously disrupt the substrate peptide–binding patch. We identified three amino acid substitutions between Src and Hck that determine the specificity of macrocycles for Src over Hck. Finally, we discovered that macrocycles derived from MC2 and MC9 potently inhibit two inhibitor resistance mutations we modeled in Src. Taken together these findings establish that macrocycles can serve as Src kinase inhibitors with exceptional kinase selectivity, potency against common drug-resistant forms of tyrosine kinases and activity in cultured mammalian cells. Additionally, our results reveal that the macrocycles studied here achieve these properties through a bisubstrate mode of kinase inhibition that locks the kinase in an inactive conformation.

3.2 Results

3.2.1 Improvement of Src-inhibiting macrocycle potency

As ATP-competitive kinase inhibitors must compete with low millimolar ATP concentrations in the cell, nanomolar *in vitro* potency (typically measured in the presence of ATP concentrations near the K_M for ATP) is often required for a kinase inhibitor to demonstrate cellular activity at micromolar concentrations[204]. We therefore sought to improve the potency of pyrazine-containing MC2 and nitrophenylalanine–containing MC9 by the systematic optimization of macrocycle building blocks (Figure 3.1 and Table 3.1). Derivatives of MC9 were

synthesized as previously described[172] and were assayed against Src_{83-533}. Because we previously observed the importance of the nitrophenyl group at the A position in MC9 for Src kinase inhibition[172], we first replaced building blocks at the B and C positions (Figure 3.1 and Table 3.1). Holding the amino acids at the A and C positions constant, we synthesized and assayed variants of MC9 containing phenylalanine (MC10), cyclohexylalanine (MC11) and pentafluorophenylalanine (MC12) at the B position in place of furylalanine (Figure 3.1 and Table 3.1). Though substituting fluorophenylalanine at this position abolished activity (IC_{50} >10 mM) and substituting cyclohexylalanine resulted in a ~10-fold increase in IC_{50}, the phenylalanine derivative (MC10) inhibited Src_{83-533} with 3-fold enhanced potency (Figure 3.1 and Table 3.1).

Next we varied building blocks at the C position. Holding the A and B positions constant as nitrophenylalanine and phenylalanine, respectively, we substituted phenylalanine (MC13), diphenylalanine (MC14), 1-naphthylalanine (MC15) and cyclohexylalanine (MC16) at the C position in place of cyclopropylalanine (Figure 3.1 and Table 3.1). Phenylalanine and naphthylalanine decreased Src_{83-533} inhibition potency by increasing the IC_{50} from three- to five-fold, whereas diphenylalanine resulted in a 16-fold increase in IC_{50}. Cyclohexylalanine, however, improved inhibition potency by more than ten-fold, resulting in MC16 with Src_{83-533} IC_{50} = 6 nM (Figure 3.1 and Table 3.1). These results collectively demonstrate that modest modifications in the size and shape of the macrocycle side chains can result in substantial gains in potency. Notably, none of the amino acids that increased Src inhibition potency were present at the appropriate position in the DNA-templated library from which the initial macrocycles were discovered. This is consistent with the accuracy of the structure-activity relationships resulting from the *in vitro* selection for Src binding[205].

We next installed more subtly altered building blocks into the partially optimized MC16. We probed the importance of nitrophenylalanine at the A position by replacing the nitro group at the para position with methyl (MC17), chloro (MC18), bromo (MC19), trifluoromethyl (MC20), cyano (MC21), carbamoyl (MC22) or *tert*-butyl (MC23) substituents. Remarkably, all analogs, except the electronically similar cyanophenylalanine, showed substantial reductions in potency (Figure 3.1 and Table 3.1). We implemented similar changes to the optimized B position building block, phenylalanine, by replacing it with tyrosine (MC24), fluorophenylalanine (MC25a) or methylphenylalanine (MC26) (Figure 3.1 and Table 3.1). Introduction of a methyl or hydroxyl group onto the phenyl ring resulted in a substantial decrease in Src inhibition. In contrast, the fluorophenylalanine derivative MC25a retained the activity of MC16 (Figure 3.1 and Table 3.1).

The analogous building block substitutions in the pyrazine-containing macrocycle structures (starting from MC2), changing furylalanine to phenylalanine and cyclopropylalanine to cyclohexylalanine, resulted in macrocycle MC4a, which had >150-fold greater potency against Src kinase than the parent compound MC2 (Figure 3.1 and Table 3.1). We could not accurately measure the potency of MC4a because its IC_{50} of ≤ 4 nM was equivalent to $\geq 50\%$ of the enzyme concentration required in the *in vitro* kinase assay. These findings support a similar mode of binding for the B and C building blocks in MC2- and MC9-derived macrocycles.

We also studied the effect of modifying the macrocycle peptide backbone on Src kinase inhibition. We systematically replaced each amide in the backbone of the improved nitrophenylalanine–containing macrocycle MC25a with an *N*-methylamide, generating MC27–MC31, and we performed the same *N*-methylamide backbone scan in the pyrazine-based macrocycle MC4a (generating macrocycles MC5–MC8; Figure 3.1 and Table 3.1)[206]. Nine of the

ten possible *N*-methylamide macrocycles were synthesized; methylation of the diaminobutyric acid α-nitrogen of MC4a impaired macrocyclization. All *N*-methylamide–containing macrocycles had substantially lower activity against Src kinase, with the exception of compounds MC5, MC27 and MC28, in which the methyl group resides on the scaffold diamino acid rather than on one of the three side chain–containing building blocks (Figure 3.1 and Table 3.1). The sensitivity of the macrocycle backbone to N-methylation suggests the importance of backbone hydrogen bonds or backbone conformation for kinase inhibition.

3.2.2 Specificity of macrocyclic Src kinase inhibitors

The development of small-molecule kinase inhibitors with specificity within SFKs has proven challenging. Notable successes in this area are a family of Lck-selective thienopyridine compounds[207,208] and a catechol-based Src-selective inhibitor[209]. Macrocycles MC1, MC2, and MC9 inhibit Src kinase with exceptional specificity over the closely related Src-family kinases Hck and Lck or c-Abl kinase[172]. To understand the origin of this unusual specificity, we characterized in greater depth the activity and specificity of three previously reported macrocycles, MC1, MC2, and MC9, as well as two second-generation compounds with greatly improved potency, MC4b and MC25b (the C-terminal carboxylate analogs of MC4a and MC25a) (Figure 3.1 and Table 3.1). Because MC4b and MC25b are equally as potent against Src kinase as carboxamides MC4a and MC25a, but offer higher aqueous solubility, we used these compounds for the biochemical and structural studies described below. Macrocycles MC1, MC2, and MC4b share a diaminobutyric acid scaffold, *cis*-olefin stereochemistry, and a pyrazine group

in position A, whereas macrocycles MC9 and MC25b contain an ornithine scaffold, *trans*-olefin stereochemistry, and a nitrophenylalanine in position A (Figure 3.1 and Table 3.1).

The original characterization of kinase inhibition was performed with a commercial kinase assay (Z'-LYTE, Invitrogen) that relies on a fluorescently labeled peptide kinase substrate. For further characterization, we used a continuous spectrophotometric kinase activity assay that is easily adapted to different substrate peptides as well as a wider range of peptide concentrations[184]. Using this assay, we found that the original compounds MC1, MC2, and MC9 inhibited 50% of Src kinase domain activity (IC_{50}) at 60 μM, 15 μM, and 6.8 μM, respectively, in the presence of 5 μM ATP and 100 μM Src optimal substrate peptide (Figure 3.2). The IC_{50} values of these compounds against Hck, Lck, and Abl kinase were over 100 μM (Figure 3.2). Consistent with their improved potency in the Z'-LYTE assay, the second-generation compounds MC4b and MC25b showed more than 100-fold higher potency than the first-generation compounds. Because the potency of the second-generation compounds approached the concentration of enzyme used in the assay, we increased the ATP concentration to 250 μM and the peptide concentration to 300 μM to determine their IC_{50} values. Under these more stringent conditions, MC4b and MC25b inhibited the Src kinase domain with IC_{50} values of 0.13 μM and 0.099 μM, respectively, and showed only modest potency against the Src-family kinases Hck (IC_{50} = 0.86 μM and 8.4 μM, respectively) and Lck (IC_{50} = 2.4 μM and 6.1 μM, respectively) (Figure 3.2). We noted that all of the compounds tested (MC1, MC2, MC4b, MC9, and MC25b) inhibit Src kinase constructs that include the regulatory SH3 and SH2 domains (chicken c-$Src_{83–533}$) approximately two- to ten-fold more potently than the isolated kinase domain (chicken c-$Src_{251–533}$) (Figure 3.3). The selectivity of the most potent second-generation compound, MC25b,

for Src over Lck, Hck, and Abl kinase was maintained when the corresponding three-domain constructs were used.

The high specificity of the macrocycles for Src kinase over Lck and Hck is unusual because Src, Lck, and Hck typically behave very similarly towards known kinase inhibitors[210]. In a recent study addressing the selectivity of kinase inhibitors, only 2 out of 38 compounds could discriminate Src from Hck with the 5-fold selectivity shown by MC4b, and none could do so with the 100-fold selectivity shown by MC25b. In the same study, only 2 out of 38 compounds showed the 10-fold selectivity of MC4b for Src over Lck, and none showed the 60-fold selectivity of MC25b[145]. The molecular basis of the remarkable specificity of these macrocyclic kinase inhibitors is revealed in the experiments described below.

3.2.3 ATP- and substrate peptide–competitive inhibition

The large size and unusual specificity of the macrocyclic Src inhibitors led us to speculate that they may interfere with the binding of not only ATP but also the substrate peptide. To probe this possibility, we tripled the concentration of the Src optimal substrate peptide in the kinase assay and observed a 25–50% increase in the apparent IC_{50} values for MC1, MC2, and MC9 (Figure 3.4). These results suggest that the macrocycle compounds could be substrate peptide competitive. For comparison, we also raised the ATP concentration 50-fold, which resulted in a two-fold increase in IC_{50} for MC1, MC2, and MC9, consistent with the known ATP-competitive behavior of the compounds[172]. To clarify the inhibitory mechanism of the compounds further, we synthesized fluorescein-labeled derivatives of MC1, MC2 and MC9 (carboxyfluorescein-MC1, carboxyfluorescein-MC2, and carboxyfluorescein-MC9, respectively)

and measured their binding affinity to the Src kinase domain in a fluorescence anisotropy assay (Figure 3.4). We found that the Src kinase domain binds the fluorescently labeled macrocycles MC1, MC2, and MC9 with dissociation constants (K_D) of 3.1 μM, 2.1 μM, and 0.52 μM, respectively. The binding affinity of the compounds decreased in the presence of 250 μM of the nonhydrolyzable ATP analog adenylyl-imidodiphosphate (AMP-PNP) to 11 μM, 3.1 μM, and 1.5 μM, respectively, as would be expected for an ATP-competitive inhibitor (Figure 3.4). Consistent with the increase in IC_{50} observed in the kinase assay, the presence of 300 μM Src optimal substrate peptide increased K_D values of the fluorescein-macrocycle conjugates by 7-fold for MC1, 1.5-fold for MC2, and 6-fold for MC9 (Figure 3.4). Because of their increased potency, we tested the substrate-competitive behavior of the improved compounds MC4b and MC25b by varying the ATP concentration from 250 μM to 2.5 mM and the substrate peptide concentration from 300 μM to 900 μM. The ATP- and substrate peptide–competitive nature of both compounds is maintained under these conditions (Figure 3.4). These results collectively indicate that the macrocycles are both ATP- and substrate peptide–competitive inhibitors.

3.2.4 Protein crystallography and structure determination for Src kinase domain bound to pyrazine- and nitrophenyl-containing macrocycles

To determine the structural basis of Src inhibition, we solved three X-ray crystal structures of MC1, MC4b, and MC25b in complex with the c-Src kinase domain. Described below are the crystallization conditions and structure determination for each Src-macrocycle complex. See Table 3.2 for crystallographic statistics for each of the structures.

3.2.4A Src•MC1

The complex between MC1 and c-Src kinase domain was formed in a solution of 190 μM kinase domain, 476 μM MC1, 50 mM Tris (pH 8.0), 125 mM NaCl, 5% DMSO, and 2.5% glycerol. Using the PEG/Ion sparse matrix screen (Hampton Research), we determined ideal crystallization conditions that were used in successive rounds of crystal optimizations screens utilizing the hanging drop vapor diffusion method. Src•MC1 was mixed in a 1:1 ratio (1 μL Src•MC1 complex + 1 μL mother liquor) with a mother liquor containing 0.1 M Bis-Tris (pH 6.5), 12% PEG 3350, and 1% Tacsimate (pH 6.0). Crystals were grown overnight at 24°C. Crystals were cryoprotected in mother liquor plus 20% ethylene glycol, frozen, and stored in liquid nitrogen until screened for diffraction.

Data collected from Src•MC1 crystals were processed in space group $P2_1$ with DENZO and Scalepack via the HKL2000 suite[177]. While the unit cell parameters are almost compatible with space group p222, processing of the data in p222 yielded poor statistics and molecular replacement with phaser failed[181]. Analysis of the data for possible twinning in space group p21 with phenix.xtriage showed significant pseudo-merohedral twinning with the h, -k, -l operator[211]. We therefore included the twin law in further refinement with Phenix, which improved refinement statistics and electron density maps. The structure was solved by molecular replacement using the kinase domain of human c-Src (PDB: 1Y57) (residues 260–520) without the αC-helix (residues 298–310) and the A-loop (residues 400–425) as a search model in Phaser[181,212].

3.2.4B Src•MC4b

Src kinase domain and MC4b were co-complexed by concentrating a 10-fold mixture of 10 μM of the kinase domain and 12.5 μM MC4b in 50 mM NaCl, 20 mM Tris (pH 8.0), and 5% glycerol. This method was used to form the co-complex as an alternative to mixing kinase directly to the drug at the final concentration. If Src and MC4b were mixed together directly without concentrating, thin, rod-like crystals grew that diffracted poorly. After several rounds of spare matrix screening and hanging drop optimization screens, crystals were successfully grown in a mother liquor containing 200 mM ammonium sulfate, and 3% glycerol overnight at 24°C. Crystals were then cryoprotected in mother liquor with 20% glycerol, frozen, and stored in liquid nitrogen.

After data collection Src•MC4b data were processed in space group P321 using Mosflm and Scala in iMosflm[178]. The structure was phased by molecular replacement using the kinase domain of inactive c-Src (PDB: 2SRC)[213] (residues 250-533) without the αC-helix (residues 298-310) and A-loop (residues 400-425) as a search model in Phaser[181].

3.2.4C Src•MC25b

The complex between MC25b and c-Src kinase domain was formed in a solution of 200 μM kinase domain, 500 μM MC25b, 20 mM Tris (pH 8.0), 125 mM NaCl, 2.5% DMSO, and 2.5% glycerol. Curiously, crystals did not grow in any of the spare matrix screens. To facilitate crystal growth, we created micro seeds from Src•dasatinib crystals using previously described crystallization conditions[214]. Micro seeds were added directly to the Src•MC25b mixture prior to

screening. Crystallization conditions were determined using sparse matrix screens and were refined using the hanging drop optimization screens. Src•MC25b was mixed in a 1:1 ratio (1 μL Src•MC25b complex + 1 μL mother liquor) with mother liquor consisting of 14% PEG 5000 MME, and 0.3 M NaH_2PO_4. Crystals were cryoprotected in mother liquor plus 20% glycerol, frozen, and stored in liquid nitrogen.

Data collected from Src•MC25b crystals were processed using XDS[215] and Aimless[216] as implemented in the autoPROC pipeline[179]. Phases were obtained by molecular replacement using the kinase domain of c-Src in complex with the macrocyclic inhibitor MC4b (PDB code 3U4W) with the αC-helix (residues 298–310), A-loop (residues 400–425), and ligand removed as a search model in Phaser[217,218]. The diffraction data were strongly anisotropic, with resolution limits (defined by $CC_{1/2}$ and $I/\sigma I$) of 2.5 Å in a* and c* reciprocal space directions, but only 3.0 Å along the b* direction[216,219]. For this reason, an anisotropic correction was carried out using the Anisotropy Diffraction Server[220]. Briefly, data were truncated that fell outside an ellipse centered at the reciprocal lattice origin and having vertices at ½.2, 1/3.0, and ½.5 Å along a*, b*, and c* axes, respectively. Isotropy was approximated by applying a negative scale factor along b* (−12 $Å^2$) with no correction along a* or c*. These anisotropically scaled data then were used for refinement in Phenix[211].

3.2.5 Structural basis for Src kinase inhibition by macrocycles

We solved three structures of Src kinase domain bound to MC1, MC4b, and MC25b at a resolution of 2.2 Å, 1.9 Å, and 2.5 Å, respectively. The structures show conserved bi-lobal structure of the kinase domain, with the P-loop, and αC-helix fully resolved in all three

structures. The A-loop was only fully resolved in our structures of Src kinase domain bound to MC4b and MC25b. In the Src•MC1 structure, A-loop residues 407-423 are unresolved. To our surprise, the kinase domain adopts the Src/CDK-like inactive conformation; to our knowledge, these structures are the first in which the isolated kinase domain of Src has been observed in this conformation. We describe the structures of the 2nd generation macrocycles below.

MC4b binds the active site of the Src kinase domain (Figure 3.5). The macrocycle forms five direct hydrogen bonds and four water-mediated hydrogen bonds with the kinase. Notably, the backbone of the macrocycle forms seven of these nine hydrogen bonds, explaining the loss of macrocycle potency upon backbone amide methylation (Table 3.1; MC5-MC8). A total of 1,392 Å2 of molecular surface is buried from solvent upon binding of MC4b to the kinase.

Consistent with the macrocycle structure-activity relationships described above, MC4b occupies three distinct binding sites (Figure 3.5). The pyrazine group from building block A binds the ATP-binding pocket and forms a hydrogen bond with the backbone of the kinase, similar to the binding mode of adenine (Figure 3.5)[197]. The phenylalanine side chain of building block B occupies a hydrophobic pocket between the β3-αC loop in the amino terminal N-lobe of the kinase and the DFG motif at the beginning of the A-loop (Figure 3.5). The outward rotation of the αC-helix and the disruption of the salt bridge between Lys295 and Glu310 open up this hydrophobic pocket, which is lined by Val281, Lys295, Leu297, Ile336, and Leu407. In the active conformation of the kinase, cyclohexylalanyl or phenylalanyl side chains at position B would clash with the side chain of Lys295 and Phe307, explaining the incompatibility of the bound macrocycles with the active conformation of the enzyme. The cyclohexylalanyl side chain of building block C faces into an amphipathic binding pocket around residues Phe278, Leu407, Ile411, Tyr416, Asp386, Arg388, and Asn391 (Figure 3.5). The C-terminal carboxyl group of

MC4b, which represents the site of attachment of DNA in the library, faces the solvent and does not interact with the kinase, showing how the DNA-linked macrocycle could bind Src during *in vitro* selection (Figure 3.5)[172].

The structure of Src bound to MC25b shows how the macrocycle binds to the active site of Src in a similar fashion to MC4b (Figure 3.5). MC25b belongs to the MC9 series of macrocycles, containing an ornithine scaffold, a trans-olefin, and a nitrophenylalanine in the A position (Figure 3.1 and Table 3.1). When compared to MC4b, the backbone of MC25b is slightly larger due to the additional carbon atom in the ornithine scaffold. MC25b forms six direct hydrogen bonds with the peptide backbone of the kinase and a single water mediated hydrogen bond. The backbone of MC25b, unlike MC4b, forms all of the hydrogen bonds with the kinase. Interestingly, only one of the hydrogen bonds is made between an amide on the backbone of MC25b and a carbonyl group on the kinase. Five of the hydrogen bonds are formed by carbonyl groups on the MC25b backbone and the C-terminal carboxyl group mediates the seventh hydrogen bond. The loss of potency seen in amide-methylated variants of MC25b (MC26-31) is explained most likely by altering the conformation of the backbone of MC25b (Table 3.1). We hypothesize that *N*-methylation results in the formation of cis-peptide bonds instead of trans-peptide bonds in the methylated amides of the building blocks, preventing the carbonyls from forming hydrogen bonds, as well as causing side chain repositioning, and preventing proper binding to their respective pockets[221]. When compared to MC4b, which forms nine hydrogen bonds with Src, MC25b is equally as potent as MC4b even with two less hydrogen bonds. MC25b makes up for the loss of these hydrogen bonds through electrostatic interactions and stronger hydrophobic interactions compared to MC4b. These interactions are

described below. A total of 1,273 $Å^2$ of molecular surface is buried from solvent upon binding of MC25b to the active site of the kinase.

MC25b, like MC4b, also occupies the same three binding sites in the active site of the kinase (Figure 3.5). The nitrophenylalanine of building block A occupies the adenine-binding pocket, however it does not form the same hydrogen bonds as adenine or MC4b. When compared to the positioning of the pyrazine of MC4b, the nitrophenylalanine does not bind as deeply into the pocket. The nitrophenylalanine forms electrostatic interactions with the catalytic lysine (Lys295) and explains the loss of potency seen with the substitutions of nitrophenylalanine in MC17-MC23. Only the cyanophenylalanine containing MC21 retains its potency for Src, since it is capable of forming electrostatic interactions with Lys295 (Table 3.1). The phenyl ring of the nitrophenylalanine occupies a rather hydrophobic space formed by the side chains of Leu273, Val281, Ser345, and Leu393. The fluorophenylalanine in the position of building block B also occupies the same hydrophobic pocket underneath the β3-αC loop as the phenylalanine in MC4b (Figure 3.5). The addition of the fluorine to the phenylalanine at this position allows MC25b to bind slightly deeper into this hydrophobic pocket (approximately 1.1 Å deeper). Replacement of this fluorophenylalanine with tyrosine, phenylalanine, or methylphenylalanine results in a loss of potency, suggesting that fluorophenylalanine is better suited for binding this hydrophobic environment. The cyclohexylalanine of MC25b occupies the same amphipathic pocket as the cyclohexylalanine group of MC4b near the DFG motif (Figure 3.5). The C-terminal carboxy group of MC25b also is solvent exposed like the carboxy group of MC4b; however unlike MC4b, it has a water mediated hydrogen bond with the sulfur of Cys277 on the P-loop of the kinase (Figure 3.5)[222].

3.2.6 Biochemical validation of macrocycle binding

In order to verify that both MC2 and MC9 based macrocycles bind to the Src/Cdk-like inactive conformation, we screened the compounds against the autophosphorylated Src kinase domain. Autophosphorylation locks the kinase in an active conformation, preventing it from switching into the Src/Cdk-like inactive conformation. As expected, MC2, MC4b, MC9 and MC25b could not inhibit autophosphorylated Src (IC_{50}s > 100 μM). Additionally, we tested the potency of MC4b and MC25b against a destabilizing mutant of the Src/Cdk-like inactive conformation, Leu407Gly[174]. Similar to autophosphorylation, the IC_{50}s of MC4b and MC25b were increased by over 200–fold compared to wild-type Src (IC_{50}s for MC4b and MC25b were ≥ 100 μM and 23 μM, respectively).

3.2.7 Structural basis of substrate peptide–competitive behavior

We were interested in the binding mode of MC1 because it has the most pronounced substrate peptide–competitive behavior of the compounds tested (Figure 3.4). MC1 and MC4b belong to the same family of macrocycles, sharing a diaminobutyric acid backbone as well as a pyrazine group in position A. They differ in position B, where MC1 contains cyclohexylalanine compared to phenylalanine in MC4b, and in position C, where MC1 contains styrylalanine instead of cyclohexylalanine in MC4b (Figure 3.1 and Table 3.1). The overall binding mode of MC4b and MC1 is similar, but MC1 binds deeper into the active site of the kinase, and the Cα atoms of groups A, B, and C in MC1 are shifted 1.6–2.6 Å toward helix αC compared to MC4b. Though the three side chains of MC1 and MC4b occupy similar binding pockets of the kinase,

the conformation of the macrocycle peptide backbone differs substantially between the two structures (Figure 3.6). The intramolecular distances within the macrocycle backbone change on average by 0.74 Å and maximally by 4.19 Å between the structures of Src in complex with MC1 and MC4b. These observations suggest that macrocycles of the MC1 family are sufficiently rigid to bind with high affinity but are flexible enough to accommodate subtle changes in binding conformation that optimize interactions with kinase residues.

Similar to the Src•MC4b structure, the side chain groups of MC1 occupy three distinct pockets in the kinase active site. The pyrazine group at the A position binds the ATP-binding pocket in both structures. At the B position, the cyclohexylalanine of MC1 and the phenylalanine of MC4b bind the hydrophobic pocket toward the outwardly rotated helix αC. The styryl group at the C position of MC1 faces the C-terminal end of the kinase's A-loop (residues 404–425), which is disordered between residues 405 and 423. The symmetry of the crystal would allow the A-loop in the Src•MC1 complex to access the same conformation as that observed in the Src•MC4b complex. Upon the binding of MC1 to the kinase, 1,459 $Å^2$ of surface area are buried, comparable to the amount of buried surface in the Src•MC4b complex. However, MC1 forms only three hydrogen bonds with the backbone of the kinase and one hydrogen bond between its carboxamide terminus and the side chain carboxylate of Asp348. The loss of five hydrogen bonds in the Src•MC1 complex compared to the Src•MC4b complex could explain the large difference in potency between these two inhibitors. This is largely the result of a 3.4-Å shift of the backbone Cα atoms of MC1 toward the αC helix.

Although a structure showing the interaction between Src and substrate peptide has not been reported, the closely related kinases Abl (Protein Data Bank (PDB) entry 2G2I[223]) and IRK (PDB entry 1IR3[224]) have been crystallized in complex with substrate peptides. An analysis of

these structures together with ours provides a plausible molecular basis for the observed substrate-competitive nature of our macrocycles. Substrate peptide binding is not compatible with the Src/CDK-like inactive conformation of the A-loop in the Src•MC1 complex, most likely because the peptide-binding patch on the kinase is disrupted. When the substrate peptide from the complex with IRK is aligned onto the Src•MC1 complex, it clashes with the A-loop and the styryl moiety of MC1 (Figure 3.6). The cyclohexyl groups of MC4b and MC25b do not clash with the substrate tyrosine in a model of a docked substrate peptide, but the inactive conformation of the A-loop in the Src•MC4b and Src•MC25b complexes disrupts the substrate-binding patch. Taken together, these observations reveal the molecular basis of the observed substrate peptide–competitive behavior of the macrocycles; the bulky groups in the B position of the macrocycles (for example, phenylalanyl, fluorophenylalanyl, or cyclohexylalanyl) bind a pocket that is only present in the Src/CDK-like inactive conformation in which the salt bridge between Lys295 and Glu310 is disrupted. In this conformation, the outward rotation of helix αC is coupled to a rearrangement of the A-loop, which subsequently disrupts the binding patch for the substrate peptide.

3.2.8 Molecular basis of Src versus Hck inhibition specificity

On the basis of the crystal structures of Src•MC4b and Src•MC25b, we identified four Src amino acids that are within 5 Å of MC4b and MC25b that differ between Src and Hck (Figure 3.7). Two of these Src residues, Gln275 and Cys277, are located in the P-loop of the kinase. The third Src residue, Leu297, is in the β3 strand. The final difference, Src Tyr340 versus Hck Phe334, represents a modest change, and either phenylalanine or tyrosine is found among

kinases that are poorly inhibited by MC4b and MC25b (Lck Tyr318, Abl Phe317). Therefore, this fourth residue was not studied further.

Replacement of Cys277 in Src with the corresponding glutamine from Hck increased the IC_{50} of the mutant protein for the two ornithine-derived compounds MC9 and MC25b by 13-fold and 5-fold, respectively. In contrast, the inhibitory potency of the diaminobutyric acid–derived compounds MC2 and MC4b was hardly affected by the C277Q mutation (Figure 3.7). The structure of MC25b shows that the replacement of the small cysteine side chain with the larger glutamine side chain induces a steric clash between the kinase and the larger ornithine-containing macrocycle backbone and also results in the loss of a water mediated hydrogen bond. Substitutions at the C building block, which is closest to the Cys277 side chain, do not correlate clearly with activity against this mutant. Because the C277Q mutation also removes a nucleophile that might react with the maleamide or fumaramide group found in all of the macrocycles, we performed competition binding experiments and time-dependent inhibition experiments. The results suggest that macrocycle inhibitors do not achieve Src selectivity by reacting with Cys277 (Figure 3.8).

Gln275 in Src is replaced by alanine or glycine in Hck, Lck, and Abl. The Q275G mutation in Src increased IC_{50} values of MC2, MC4b, MC9 and MC25b between 5- and 7-fold (Figure 3.7). Notably, the side chain of Gln275 faces away from both MC4b and MC25b; and hydrogen bonds to a salt bridge between Lys272 and Glu280 (Figure 3.7). This salt bridge is thought to rigidify the otherwise flexible P-loop and its disruption decreases kinase activity[25]. We speculate that the loss of this hydrogen bond in the Q275G mutant destabilizes the P-loop and results in a reduction in potencies for MC4b and MC25b. Likewise, disrupting the Lys272–Glu280 salt bridge by mutating Glu280 to valine in Src destabilizes the P-loop and increases the

IC_{50} of MC4b and MC25b by 4-fold and 12-fold respectively (Figure 3.7). Overall, the results of mutations to the P-loop of Src indicate that the binding of MC4b and MC25b to Src is most likely dependent on the stability of the P-loop. In Hck, Lck, and Abl, which lack the glutamine in the salt bridge, the P-loop may be more flexible and may adopt different conformations that make macrocycle binding more entropically unfavorable. Consistent with this model, some of the most common imatinib resistance mutations are in the P-loop of Abl kinase and include the residues corresponding to Src residues Gln275, Lys272, and Glu280[196].

Leu297 is part of a cluster of hydrophobic residues that form the binding pocket for the B building blocks of MC1, MC4b, and MC25b. Replacement of Leu297 in Src with the corresponding Hck residue (Met292) would probably decrease the size of the binding pocket and would lead to a steric clash, explaining the lowered potencies of MC4b and MC25b for Hck. All of the macrocycles tested contain a bulky five- or six-membered ring in the B position, and the L297M mutation in Src increased the IC_{50} values of MC2, MC4b, MC9, and MC25b by 2- to 5-fold, respectively (Figure 3.7). Because the two regions of sequence divergence between Src and Hck are distinct, their effect on macrocycle potency could be at least partially additive.

Taken together, these mutational studies interpreted in light of the crystal structures suggest that specific interactions involving Src residues Gln275, Cys277, and Leu297 are the molecular basis for the unusual selectivity of the macrocycles for Src versus other Src-family kinases such as Hck.

3.2.9 Macrocycles are active against inhibitor resistance mutants of Src

Inhibitor resistance mutations can arise during treatment with small molecule inhibitors, making kinase inhibition all the more difficult. Resistance mutations in the kinase domain are mainly focused around the hinge region, the P-loop and A-loop of the kinase domain[165,225]. This generates the need for inhibitors that can bind to these mutations. We had already screened the macrocycles against the E280V P-loop mutant we modeled in Src (Figure 3.7). Therefore, we tested the macrocycles against two inhibitor resistance mutants we modeled in Src kinase: the hinge region gatekeeper mutant Thr338Ile (Src^{T338I}) and the A-loop Arg419Pro (Src^{R419P}).

Remarkably, the nitrophenylalanine–containing macrocycles MC9 and MC25b inhibited the three-domain construct of the gatekeeper mutant Src^{T338I} (residues 83–533) with IC_{50} value increases of only 1.8- to 3.1-fold compared to wild-type Src (Figure 3.9). A similar fold change was seen with Src^{T338I} kinase domain, where the IC_{50} increased 4- to 6-fold for MC25b and MC9, respectively. In contrast, the gatekeeper mutation had a much stronger impact on the potency of macrocycles containing the pyrazine building block at position A; for example, macrocycles MC2 and MC4b inhibited the gatekeeper mutants of Src_{83-533} and Src kinase domain with IC_{50} values 15- to 100-fold higher than those against wild-type Src. Similarly, the gatekeeper mutation increased the dissociation constant almost 100-fold for fluorescein-labeled MC2 but only 1.3-fold for fluorescein-labeled MC9. This decrease in potency and affinity is consistent with both the Src•MC4b and Src•MC25b structures, which suggests that substitution of Thr338 with isoleucine would create a steric clash with the pyrazine group found in the A position of MC2 and MC4b. In contrast, macrocycles MC9 and MC25b contain a smaller nitrophenylalanine building block at this position. The structure of MC25b bound to Src shows that the smaller nitrophenyl group avoids steric repulsion with an isoleucine at Src residue 338, enabling compounds MC9 and MC25b to retain their potency against the Src^{T338I} gatekeeper

mutant (Figure 3.9). We hypothesize that a larger hydrophobic amino acid substitution at the gatekeeper threonine in Src (such as the EGFR gatekeeper mutation Thr to Met) will result in a loss of potency for MC9 and MC25b.

The BCR-Abl His396Pro mutation is an A-loop mutation that affects the folding of the A-loop back on to the inactive kinase. In BCR-Abl, this renders the fusion kinase resistant to imatinib[167,226-228]. We modeled this mutation in Src kinase domain by mutating the Arg419 to proline to assess macrocycle potency against A-loop mutations[174]. Interestingly, the pyrazine-containing MC4b was able to retain potency against this mutation, while the nitrophenylalanine containing MC25b was not. MC4b only had a 2.6 fold increase in IC_{50} while the IC_{50} of MC25b increased by over 26-fold when compared to wild-type Src kinase domain (Figure 3.9). In the crystal structures of MC4b and MC25b, Arg419 faces away from the active site. We hypothesize that substituting a proline at this position alters the conformation of the A-loop in the Src/Cdk-like inactive conformation, and would potentially result in the loss of a hydrogen bond for MC25b, but not for MC4b. A carbonyl group of the backbone of MC25b forms a hydrogen bond with Arg388, and the perturbation of the A-loop by R419P would most likely shift Arg388 out of proper hydrogen bonding distance. Furthermore, comparing the structures of MC1 and MC4b to MC25b reveal that the pyrazine-based macrocycle backbone has a degree of flexibility that can accommodate subtle changes in binding conformation. The larger ornithine scaffold of the MC9 family may not be able to tolerate such changes, explaining the loss of potency against Src^{R419P}.

3.2.10 The macrocycles have in cell efficacy

The development of macrocyclic Src kinase inhibitors with low nanomolar *in vitro* potency raised the possibility that these compounds may inhibit Src kinase activity in living cells. We assayed our most potent macrocycles, MC25a and MC4b, against endogenous Src in HEK 293T cells. MC25b was not used since we believe it is less cell permeable based on previous experiments done by the Liu lab[217]. Upon treatment with 20 μM of nitrophenylalanine–containing MC25a, the total amount of autophosphorylated Src was substantially reduced to below 50%. Treatment with pyrazine-containing MC4b at 20 μM only affected the total amount of autophosphorylated Src slightly (Figure 3.10).

We speculate that in addition to the abovementioned competition with high concentrations of intracellular ATP, factors including modest cell membrane permeability may explain the substantial differences between the *in vitro* potency and cell culture activity of these compounds. The data suggests that the EC_{50} for MC25a may be lower than previously reported EC_{50} of 60 μM[217]. Further experiments will need to be done to assess the efficacy of these compounds in cell.

3.3 Discussion

We systematically modified the building blocks of macrocyclic kinase inhibitors identified from the *in vitro* selection of a DNA-templated library and improved their potency by up to 240-fold while maintaining their unusually high specificity for Src kinase. Characterization of the inhibitory mechanism revealed that the compounds are both ATP competitive and substrate peptide competitive. Because the peptide-binding patch on kinases is less conserved

than the ATP-binding pocket, compounds interacting with the former offer rich opportunities to develop kinase-specific inhibitors.

The three-dimensional structures of the Src kinase domain complexed with a first-generation inhibitor and with two improved second-generation compounds reveal that Src adopts the Src/CDK-like inactive conformation. These structures represent, to our knowledge the first time that Src kinase domain alone has been observed in this conformation, which has previously only been seen in structures of larger Src constructs containing N-terminal SH3/SH2 domains. This inactive kinase conformation is incompatible with substrate peptide binding and provides the molecular basis for the substrate peptide–competitive behavior of the inhibitors. Macrocycle binding requires the Src- or CDK-like inactive conformation for at least two reasons: first, the outward rotation of helix αC and the disruption of the salt bridge between Lys295 and Glu310 is required to form a binding pocket for building blocks in the B position; second, the conformation of the A-loop of the kinase in its active form would clash with the backbone of the inhibitors as well as with building blocks in the C position.

The macrocycles studied here inhibit the larger Src constructs about 10-fold more potently than the Src kinase domain. Our structural data supports a model in which the ligands form no extra interactions with the larger constructs, and therefore the observed differences in inhibitory potency are most likely due to the relative stabilities of the conformational states between the isolated kinase domain and the SH3 and SH2 kinase domain constructs. This reasoning suggests that the SH3 and SH2 domains stabilize the Src/CDK-like inactive conformation by approximately 1.3 kcal mol^{-1} [71,229].

On the basis of the structures and subsequent assays of mutant Src kinases, we identified two Src residues in the P-loop (Gln275 and Cys277) and one residue lining a hydrophobic pocket

(Leu297) that largely explain the selectivity of the macrocycles for Src compared with Hck, Lck and, by inference, other Src-family kinases.

The two different families of Src-inhibiting macrocycles developed and characterized here differ mainly in their A-position building blocks and in the length and olefin stereochemistry of their backbones. The nitrophenylalanine–based macrocycles inhibit the gatekeeper mutant of Src kinase with potency comparable to that of wild-type Src. In contrast, on the basis of our structural data, the pyrazine-containing macrocycles are predicted to clash with the T338I mutant and indeed have a much lower ability to inhibit this mutant than the *p*-nitrophenylalanine–based macrocycles. Interestingly, the pyrazine-based macrocycles do maintain potency against Src^{R419P} while the nitrophenylalanine-based macrocycles experience a large loss of potency. This suggests that both series of macrocycles may be useful for targeting different inhibitor resistance mutations depending on where they arise in the kinase domain.

Furthermore, we speculate that the MC compounds may have a modest EC_{50} in cells. The preliminary experiments described above have demonstrated that MC25a reduced the amount of endogenous autophosphorylated Src by greater than 50% at 20 μM when compared to DMSO treated cells. While MC25a is not as potent as our positive control of dasatinib, it demonstrates that the MCs have potential to be used as specific Src inhibitors for studying Src-mediated signaling pathways in cells.

The dysregulation and activity of many kinases is associated with human disease. Highly specific kinase inhibitors, such as the macrocyclic compounds presented here, that compete not only with ATP binding but also with substrate peptide binding could inspire the development of new inhibitors with two qualities unavailable to ATP-competitive inhibitors. First, we have shown that the level of peptide competition is tunable in these macrocycles. Second, a single

peptide-competitive kinase inhibitor could, in principle, reshape signaling pathways downstream of the kinase by favoring the phosphorylation of strongly competing substrate peptides over weakly competing substrate peptides rather than by simply inhibiting the ability of the kinase to phosphorylate all downstream targets. Finally, the potency of some of the macrocycles characterized here, including those that potently inhibit drug resistant mutations, together with their observed activity in cell culture is encouraging for the future development of macrocyclic kinase inhibitors with potential therapeutic relevance.

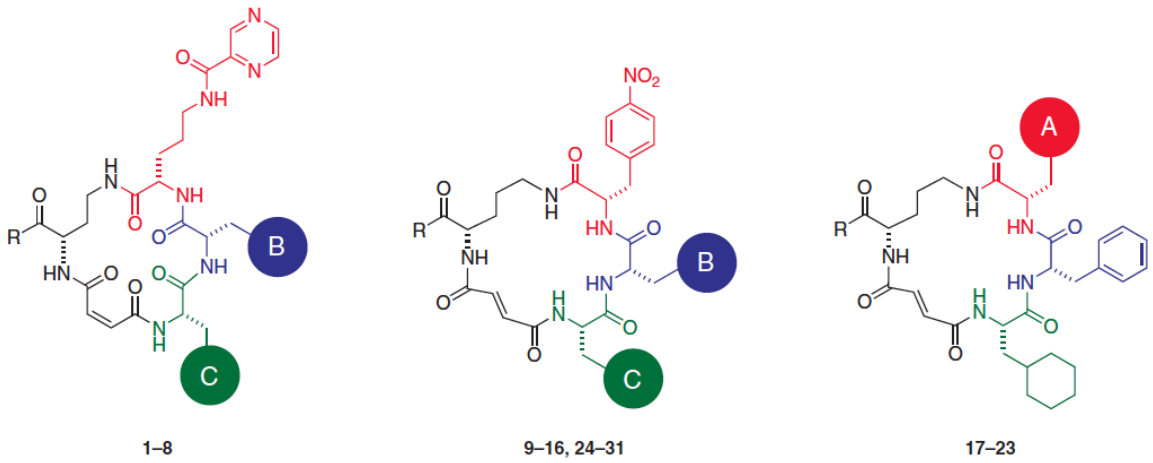


Figure 3.1 - Chemical structures of macrocycles described in this work. The compounds fall into two families: **1-8**, which contain a diaminobutyric acid scaffold, *cis*-olefin, and *N*-pyrazinylcarbonyl-ornithine building block at the A-position, and **9-31**, which contain an ornithine scaffold, *trans*-olefin, and *para*-substituted phenylalanine building block at the A-position.

(A)

IC_{50} [μM]	Src	Hck	Lck	Abl
MC1	60 μM	> 100 μM	> 100 μM	> 100 μM
MC2	15 μM	> 100 μM	> 100 μM	> 100 μM
MC9	6.8 μM	> 100 μM	> 100 μM	> 100 μM

(B)

IC_{50} [μM]	Src	Hck	Lck	Abl
MC4b	0.13 μM	0.86 μM	2.4 μM	> 100 μM
MC25b	0.099 μM	8.4 μM	6.1 μM	> 100 μM

Figure 3.2 – Specificity of the macrocyclic kinase inhibitors. (A) IC_{50} values of MC1, MC2, and MC9 for the kinase domains of Src, Hck, Lck and Abl were determined in the presence of 5 μM ATP and 100 μM Src-optimal peptide. (B) IC_{50} values of MC4b and MC25b for the kinase domains of Src, Hck, Lck and Abl were determined in the presence of 250 μM ATP and 300 μM Src-optimal peptide.

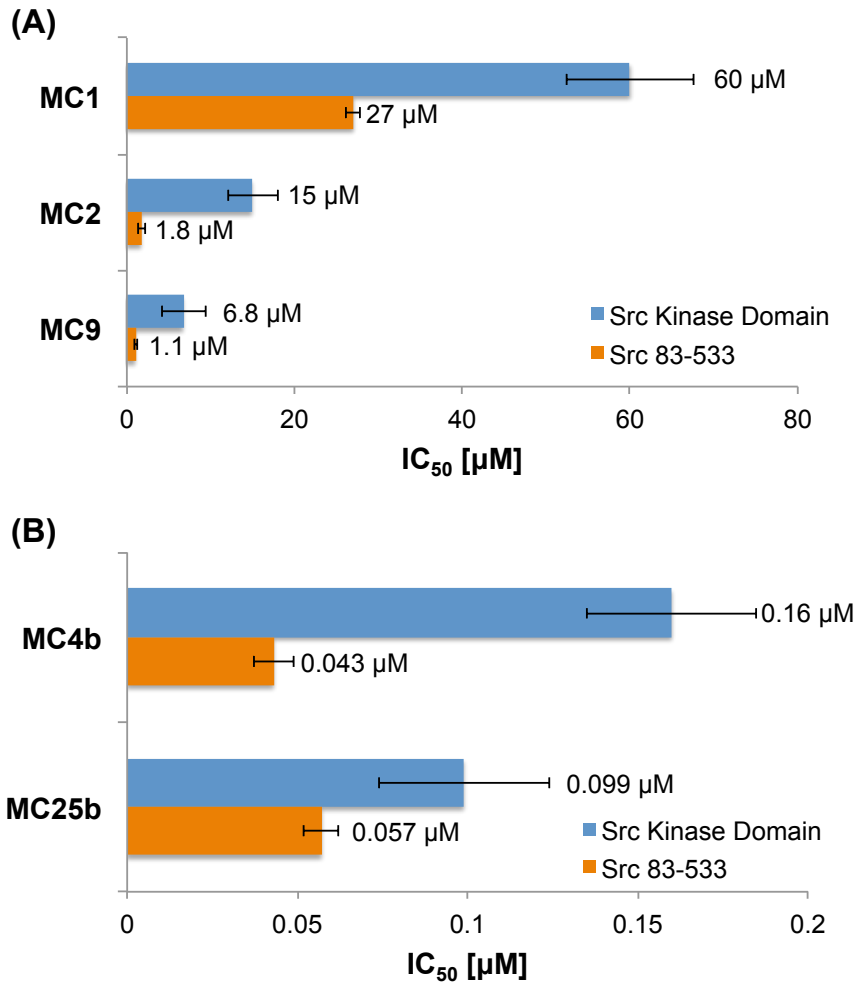


Figure 3.3 - Potency of macrocycle compounds against Src kinase domain constructs. (A) Inhibition of Src kinase domain (chicken c-$Src_{251-533}$) and Src containing the N-terminal SH3 and SH2 domains in addition to the kinase domain (chicken c-Src_{83-533}). IC_{50} values of MC1, MC2, and MC9 were determined in the presence of 5 μM ATP and 100 μM Src-optimal substrate peptide. (B) Inhibition of Src kinase domain and Src_{83-533} by MC4b, and MC25b was determined in the presence of 250 μM ATP and 300 μM Src-optimal substrate peptide. All experiments were performed in triplicate, and data represent mean values ± standard deviation.

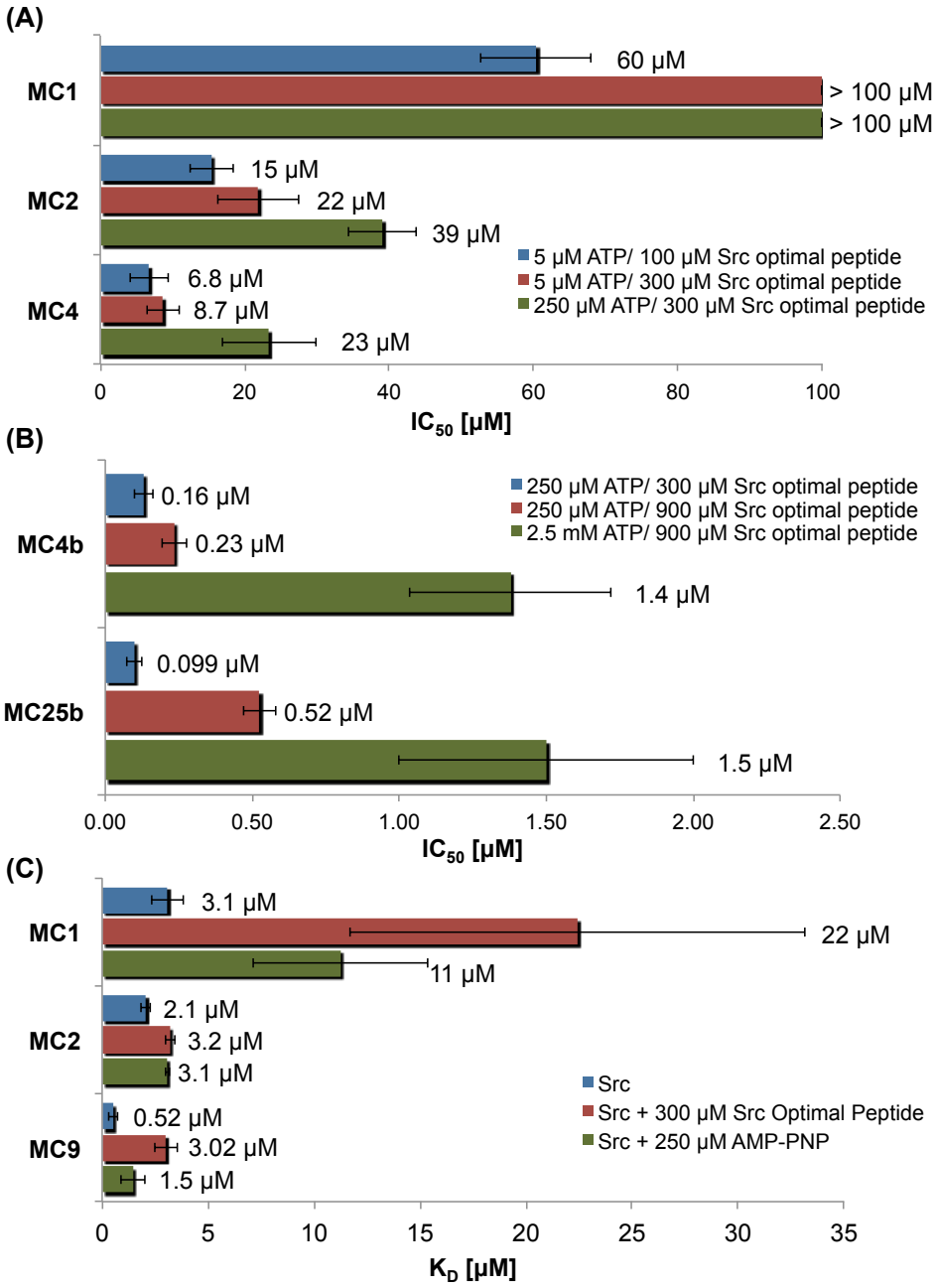


Figure 3.4 - Macrocycle compounds are ATP- and peptide-competitive inhibitors. (A) IC_{50} values of MC1, MC2, and MC9 for Src kinase domain in response to a 50-fold increase in ATP concentration (5 μM to 250 μM), and a 3-fold change in Src-optimal peptide concentration (100 μM to 300 μM). (B) Dissociation constant (K_D) of fluorescein-labeled MC1, MC2, and MC9 for Src kinase domain in the presence of 250 μM AMP-PNP or 300 μM Src-optimal peptide. (C) IC_{50} values of MC4b, and MC25b for Src kinase domain in response to a 10-fold increase in ATP (250 μM to 2.5 mM), and a 3-fold change in Src-optimal peptide (300 μM to 900 μM). All experiments were performed in triplicate, and data represent mean values ± standard deviation.

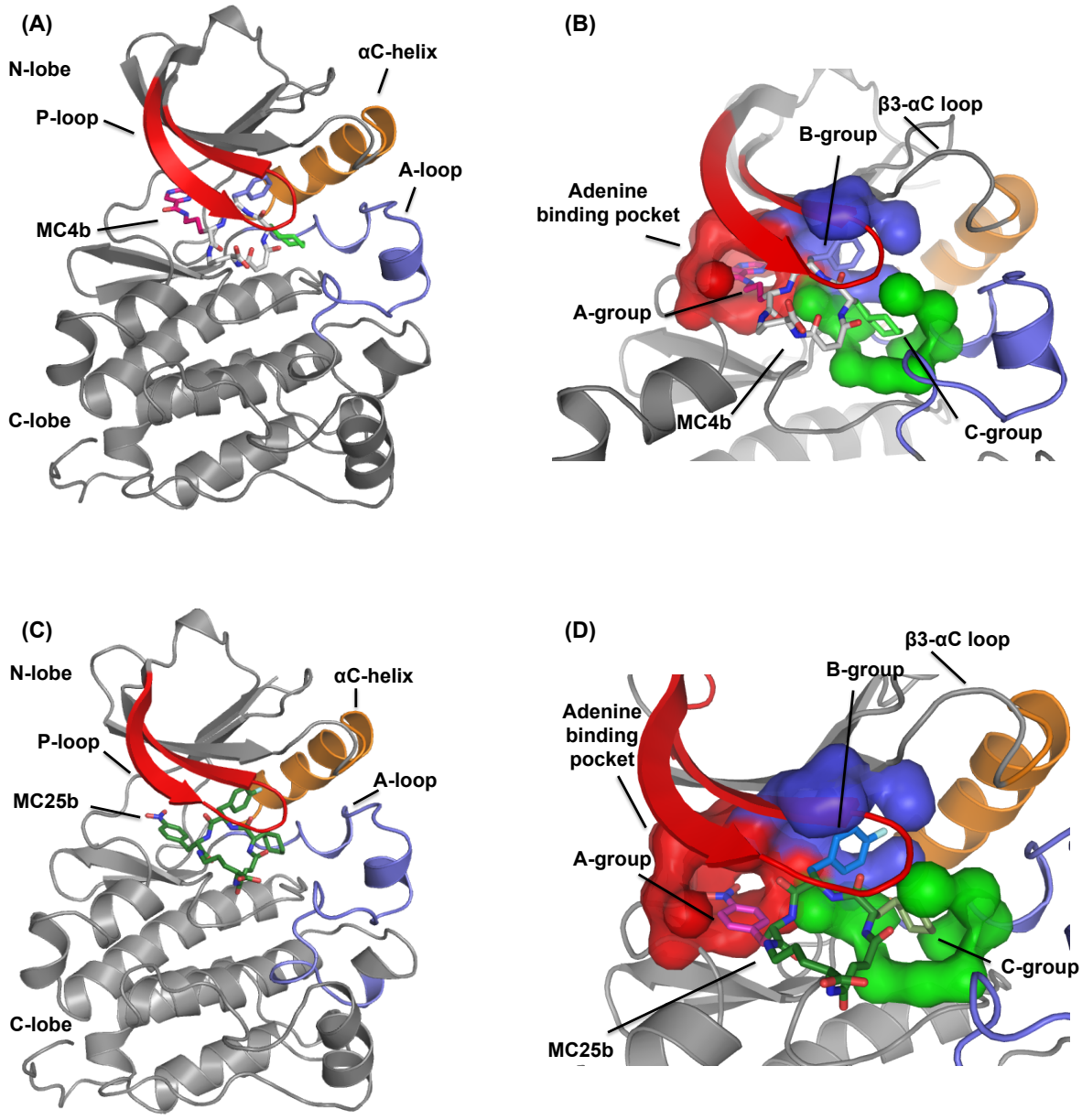


Figure 3.5 – The three-dimensional structure of Src kinase domain bound to macrocyclic inhibitors. (A, C) MC4b and MC25b bind to the active site of the kinase underneath the phosphate-binding P-loop (shown in red). The kinase adopts the Src/CDK-like inactive conformation characterized by the outward orientation of helix aC (orange) and the disruption of the peptide-binding patch by the activation loop (blue). (B, D) The three building block positions of the macrocycle compounds occupy three distinct binding pockets: the A-building block binds to the adenine binding pocket (red), the B-building block binds to a hydrophobic pocket underneath the b3-aC loop, and the C-building block occupies a binding pocket (green) facing the Asp-Phe-Gly motif of the activation loop.

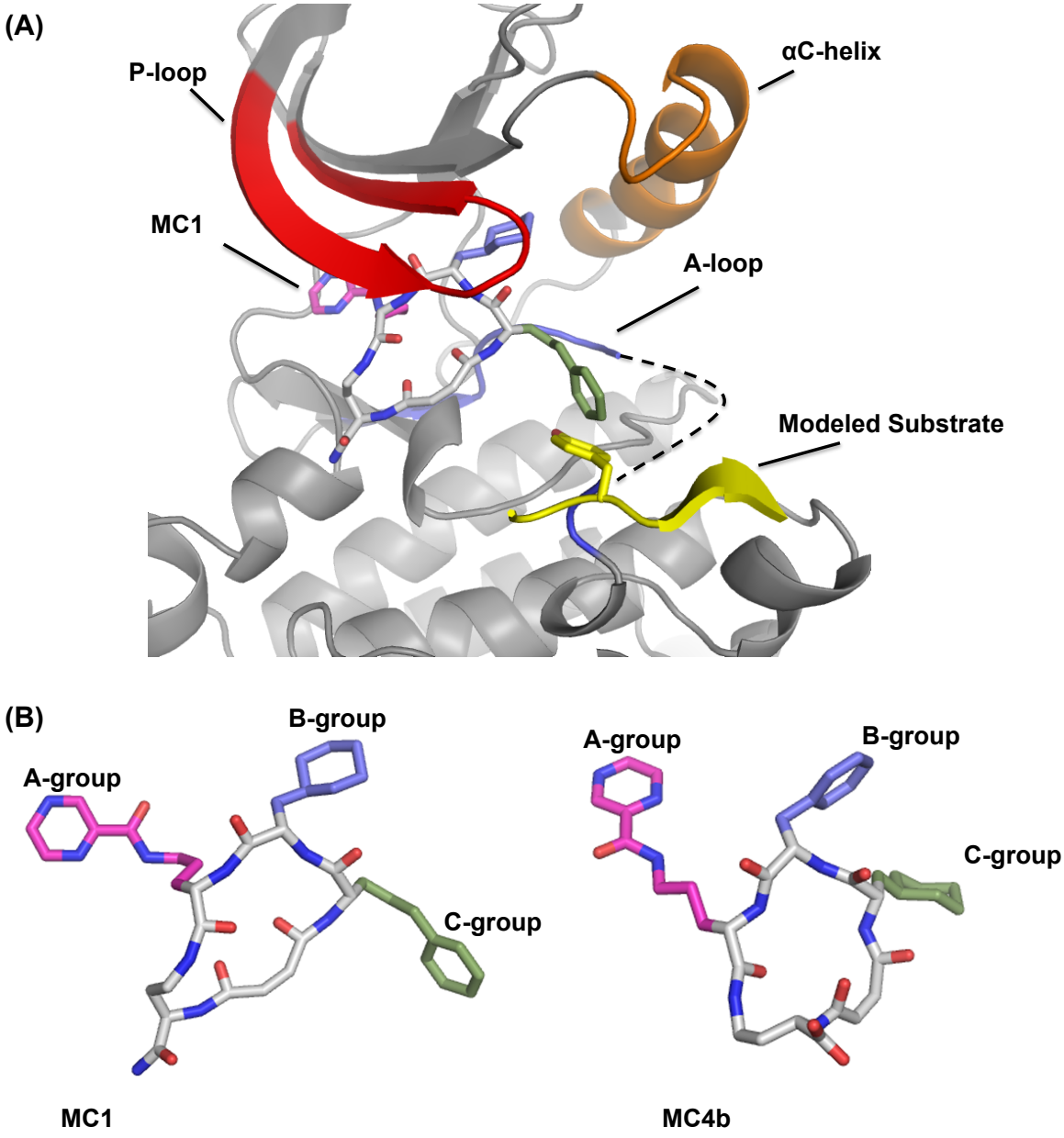


Figure 3.6 – Structural basis for substrate peptide competitive behavior of MC1. (A) Superposition of the experimental X-ray crystal structure of Src•MC1 with the structure of the substrate peptide (yellow) from the complex with IRK (PDB entry 1IR3)[50]. (B) Comparison of the structures of MC1 and MC4b when complexed with Src kinase domain. The macrocycle structures are shown from a perspective that fixes the kinase domains (not shown) in the same orientation.

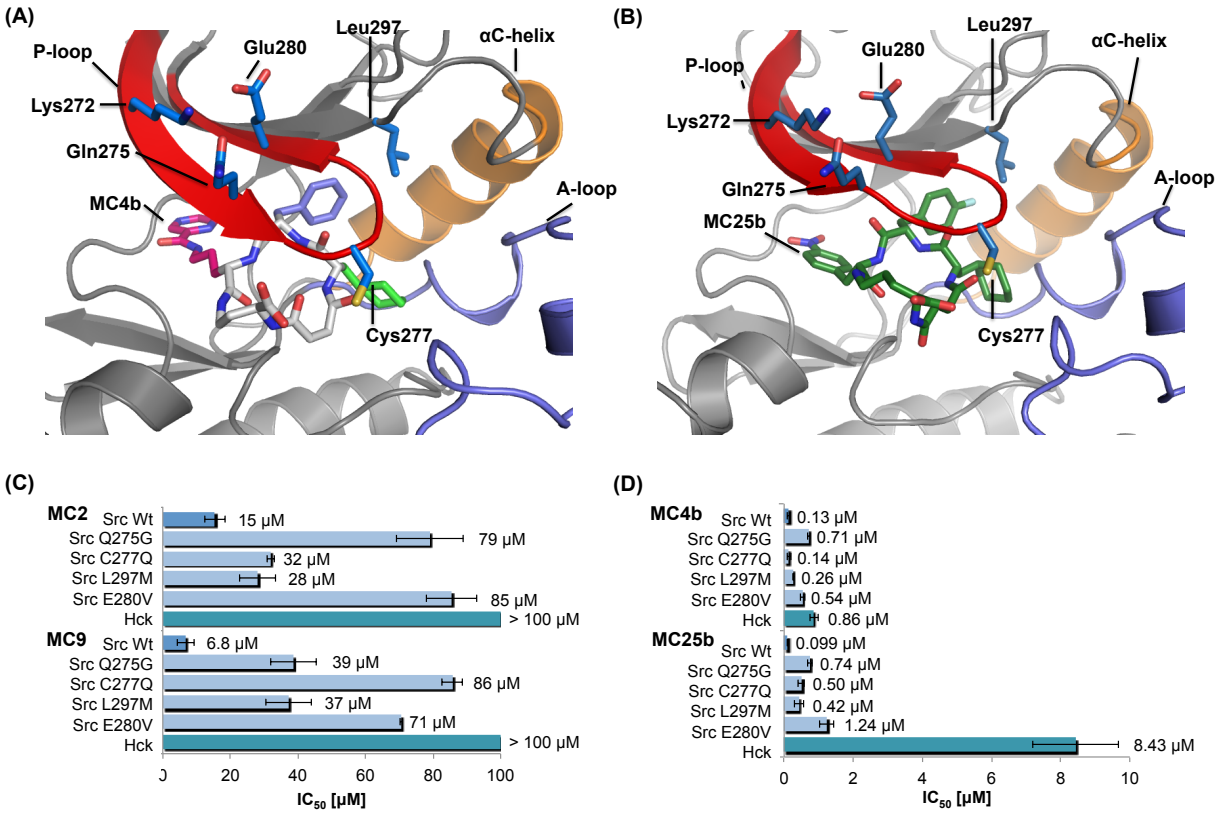


Figure 3.7 – Determinants of macrocycle specificity. (A, B) In the structural models of the Src•MC4b and Src•MC25b complexes, the three amino acids within 5 Å of the macrocycle that differ between Src and Hck/Lck (Src Gln275, Cys277, and Leu297) are rendered as sticks, as well as the salt bridge forming Src residues Lys272 and Glu280. (C, D) IC_{50} of MC2, MC9, MC4b, and MC25b for Src wt, Src Q275G, Src C277Q, Src E280V, Src L297M, and Hck kinase domains. IC_{50} values for MC2 and MC9 were determined in the presence of 5 μM ATP and 100 μM Src-optimal peptide. IC_{50} for MC4b and MC25b were determined in the presence of 250 μM ATP and 300 μM Src-optimal peptide. All experiments were performed in triplicate, and data represent mean values ± standard deviation.

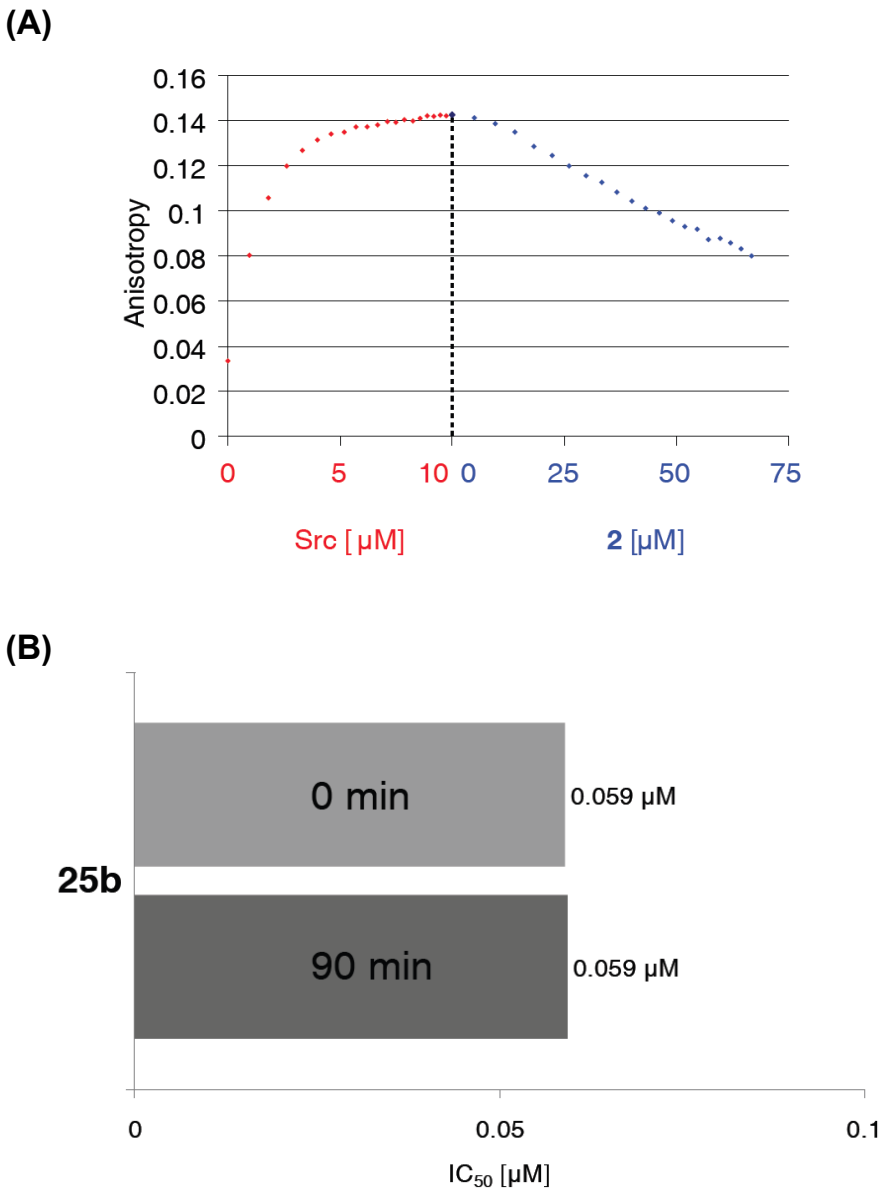


Figure 3.8 - Macrocycles do not chemically react with Cys277. (A) We verified that macrocycle binding was reversible through fluorescence anisotropy. Src kinase domain was added to 0.5 μM fluorescein-MC2. The anisotropy of fluorescein (left half of the plot, concentration referring to Src kinase domain added to the fluorescently labeled compound) was followed as Src concentrations increased. The resulting increase in anisotropy indicated binding of Src to fluorescein-MC2. Following saturation of fluorescein-MC2 anisotropy, we added increasing amounts of unlabeled MC2 to the mixture of Src and fluorescein-2. The decrease in anisotropy indicates the replacement of fluorescein-2 with unlabeled 2 and confirms the reversibility of 2-Src binding. (B) Time-dependent inhibition study shows the macrocycles do not react with Cys277. Src_{83-533} (100 nM) was incubated with 0-640 nM 25b and kinase activity was measured either immediately ("0 min") or after 90 min incubation at 30 ºC ("90min") using the spectrophotometric assay with Src optimal peptide at 300 μM and 250 μM ATP.

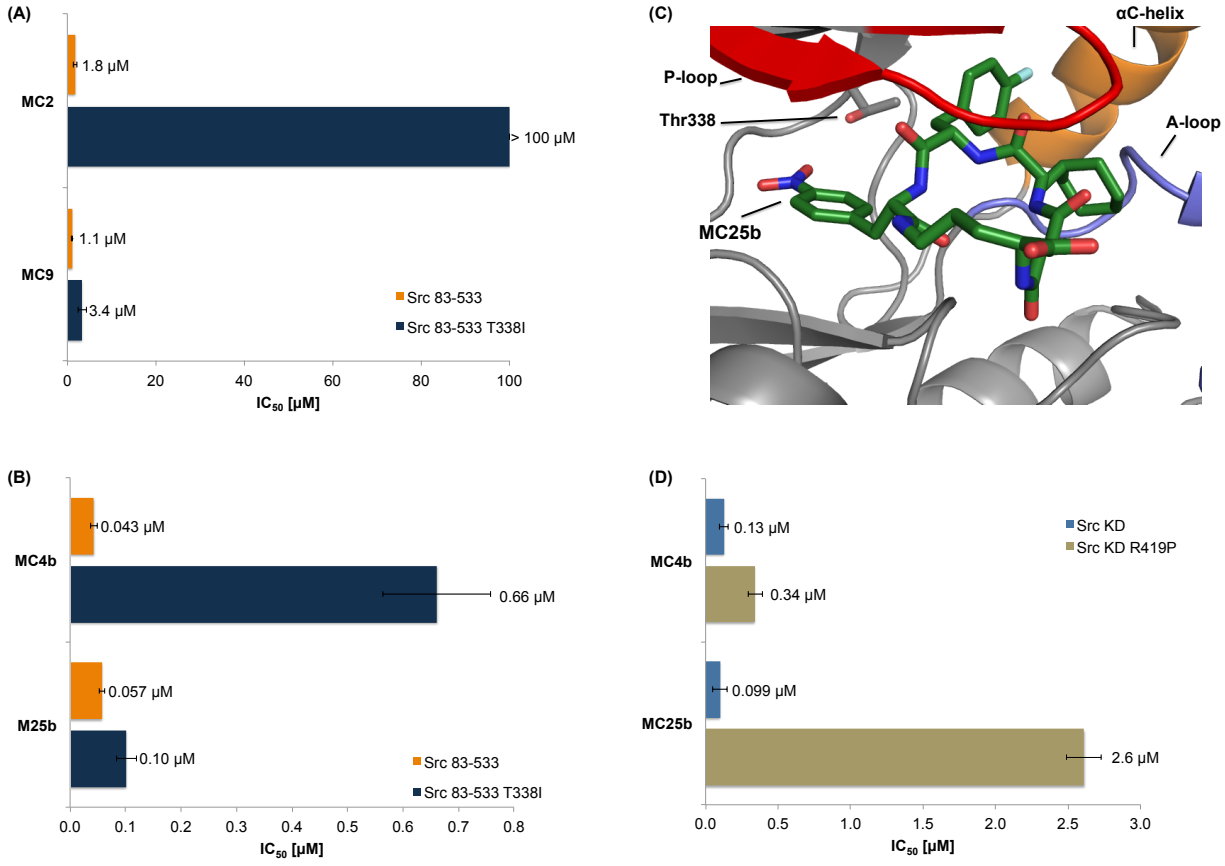


Figure 3.9 - Activity of the macrocycles against resistance mutations modeled in Src kinase. (A) IC_{50} of MC2 and MC9 for Src_{83-533} and Src_{83-533} T338I. IC_{50} values for MC2 and MC9 were determined in the presence of 5 μM ATP, 100 μM Src-optimal peptide. (B) IC_{50} of MC4b, and MC25b for Src_{83-533}, and Src_{83-533} T338I. IC_{50} values for MC4b and MC25b were determined in the presence of 250 μM ATP and 300 μM Src-optimal peptide. (C) The Src•MC25b structure shows how MC25b does not bind to the gatekeeper threonine. A mutation to isoleucine would not clash with the nitrophenylalanine group. (D) IC_{50} of MC4b, and MC25b for Src kinase domain (KD), and Src KD R419P. IC_{50} values for MC4b and MC25b were determined in the presence of 250 μM ATP and 300 μM Src-optimal peptide. All experiments were performed in triplicate, and data represent mean values ± standard deviation.

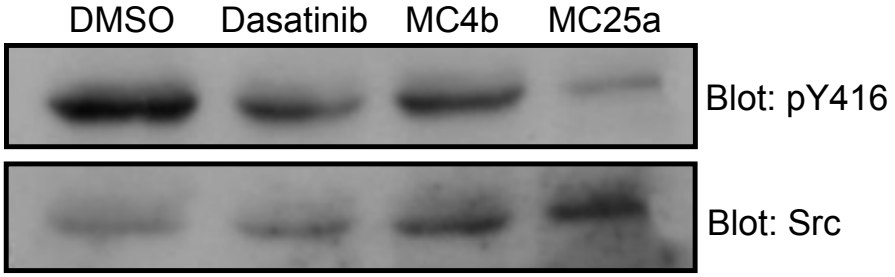
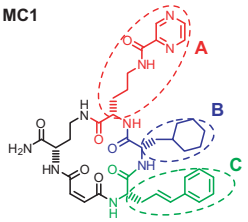
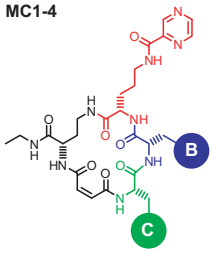


Figure 3.10 – Inhibition of endogenous Src autophosphorylation in HEK 293T cells. Cells were treated with 5 μM dasatinib, 20 μM MC4b/MC25a and a vehicle control. Remarkably, MC25a had a significant success in inhibition of Src autophosphorylation. Further studies must be done to validate and further expand on these preliminary results.

MC1

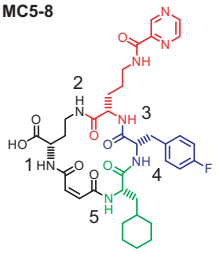


MC1-4



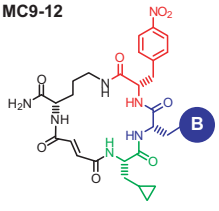
Compound	B building block	C building block	Src IC_{50} (μM)
MC1	cyclohexylalanine	styrylalanine	25
MC2	furylalanine	cyclopropylalanine	0.96
MC3	phenylalanine	cyclopropylalanine	0.19
MC4a	phenylalanine	cyclohexylalanine	0.004

MC5-8



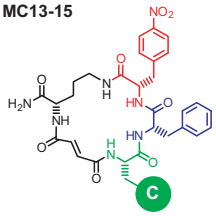
Compound	*N-methyl* amide	Src IC_{50} (μM)
MC5	position 2	0.004
MC6	position 3	3.0
MC7	position 4	>5.0
MC8	position 5	1.0

MC9-12



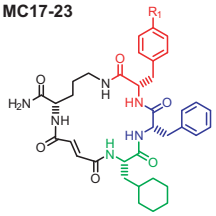
Compound	B building block	Src IC_{50} (μM)
MC9	furylalanine	0.25
MC10	phenylalanine	0.08
MC11	cyclohexylalanine	2.0
MC12	pentafluorophenylalanine	>10

MC13-15



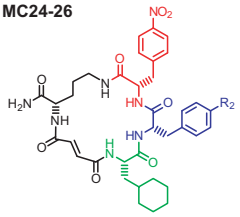
Compound	C building block	Src IC_{50} (μM)
MC13	phenylalanine	0.22
MC14	diphenylalanine	1.3
MC15	1-naphtylalanine	0.37
MC16	cyclohexylalanine	0.006

MC17-23



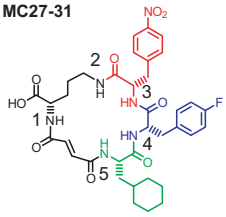
Compound	R_1	Src IC_{50} (μM)
MC17	CH_3	0.27
MC18	Cl	0.047
MC19	Br	0.030
MC20	CF_3	0.18
MC21	CN	0.005
MC22	$CONH_2$	0.0036
MC23	*tert*-butyl	>10

MC24-26



Compound	R_2	Src IC_{50} (μM)
MC24	OH	0.79
MC25a	F	0.004
MC26	CH_3	0.37

MC27-31



Compound	*N-methyl* amide	Src IC_{50} (μM)
MC27	position 1	0.004
MC28	position 2	0.036
MC29	position 3	0.16
MC30	position 4	5.0
MC31	position 5	0.57

Table 3.1 - Chemical structures and kinase inhibition activities of macrocycles described in this work. The concentration of compound that inhibited 50% of Src_{83-533} kinase activity (IC_{50}) was determined using the Z'-LYTE kinase assay[217]. All experiments were performed in triplicate, and data represent mean values ± standard deviation.

Table 3.2 - Data collection and refinement statistics for Src crystallography

	Src•MC1	Src•MC4b	Src•MC25b	
Data collection				
Space group	$P2_1$	P321	$P2_1$	
Cell dimensions				
a, *b*, *c* (Å)	42.2, 117.3, 62.7	143.6, 143.6, 41.5	60.8, 118.6, 42.6	
α, β, γ (°)	90.0, 90.1, 90.0	90.0, 90.0, 120.0	90 90.1 90	
			Before Anisotropic Correction	*After Anisotropic Correction*
Resolution (Å)	50-2.24 *	41.5-1.9	42.6−2.86	42.6−2.5
R_{sym} or R_{merge}	0.069 (0.30)	0.098 (0.42)	0.222 (0.881)	0.252 (0.670)
I / σI	14.2 (3)	13.7 (4.1)	9.7 (2.7)	8.2 (3.0)
Completeness (%)	97.9 (85.7)	100 (99.9)	100 (100)	83.7 (26.6)
Redundancy	3.1	6	7.4 (7.5)	7.4 (6.7)
Refinement				
Resolution (Å)	2.24	1.9	2.5	
No. reflections	29,870	38,763	17,499	
R_{work} / R_{free}	0.1929 / 0.2496	0.1622 / 0.1849	0.2182 / 0.2459	
No. atoms				
Protein	4140	2223	4928	
Ligand/ion	106	51	103	
Water	107	324	151	
B-factors				
Protein	Chain A: 48.4 Chain B: 48.9	17.03	Chain A: 36.78 Chain B: 36.13	
Ligand/ion	Chain A: 53.2 Chain B: 53.4	24.08	Chain A: 41.93 Chain B: 38.10	
Water	39.2	27.44	24.60	
R.m.s. deviations				
Bond lengths (Å)	0.002	0.007	0.003	
Bond angles (°)	0.687	1.117	0.831	
PDB Ascension Code	3U51	3U4W		

*Highest-resolution shell is shown in parentheses.

Chapter 4: The molecular mechanism of DDR1's promiscuous binding of small molecule kinase inhibitors

The work in this chapter is credited to several people:

George Georghiou was responsible for the protein expression, biochemical assays, protein crystallography and designing cellular experiments.

Dr. Leah Ethridge (Hitchcock Lab) was responsible for performing the cellular based assay

Dr. Ian Hitchcock oversaw and advised on the cellular experiments.

Dr. W. Todd Miller provided resources and advised on *in vitro* kinase assays.

Dr. Markus A. Seeliger advised and reviewed all experiments.

4.1 Introduction

Small molecule ATP competitive kinase inhibitors have demonstrated great clinical efficacy in treating diseases like cancer[3,4,148,149,230,231]. The most famous example of this is the BCR-Abl inhibitor imatinib, which increased the five year survival rate of chronic myleogenous leukemia patients from 20% to 90%[160,165]. However, due to the high conservation of the ATP binding pocket, ATP-competitive kinase inhibitors typically bind several kinases, some of which may not be medically relevant targets[163]. Imatinib, for example, binds to several kinases besides BCR-Abl, such as PDGFR and c-Kit[163,171,232,233]. Dasatinib, a second-generation inhibitor of BCR-Abl, can bind an even larger number of kinases compared to imatinib. These include PDGFR, c-Kit, Src family kinases, VEGFR, Tie2, Eph family kinases, and BMX[148,161]. Treating cancer with inhibitors like dasatinib or imatinib can affect both normal and cancer cells[146,162]. In normal cells, off target kinases binding inhibitors like dasatinib can cause adverse effects that lower a patient's quality of life while undergoing treatment[147,234]. This can potentially lead to the discontinuation of treatment using small molecule kinase inhibitors[146].

Small molecule ATP competitive inhibitors are classified into two groups: Type I and Type II inhibitors[148]. Type I inhibitors, such as dasatinib, bind to the active kinase conformation[148,169]. Inhibitors that bind the Abl/c-Kit-like inactive conformation, such as imatinib or nilotinib, are classified as Type II inhibitors[148,170,235]. Type II inhibitors have the added benefit of being more selective for the kinases they inhibit compared to Type I inhibitors due to binding a poorly conserved hydrophobic pocket only accessible in the Abl/c-Kit-like inactive conformation[148].

A recent study assessed the specificity of 38 chemically different ATP-competitive small molecule kinase inhibitors by screening them against over 300 human kinases[163]. From this emerged several findings, including previously unknown off targets of these inhibitors. Curiously, several kinases were found to promiscuously bind many of these inhibitors. The majority of these promiscuous kinases are medically relevant targets that inhibitors were designed to inhibit, such as PDGFR and VEGFR. Interestingly, one of these promiscuous kinases was DDR1, which bound to 11 out of the 38 inhibitors screened with an affinity of 100 nM or less. Overall, DDR1 bound to 23 out of the 38 inhibitors screened with an affinity of less than 5 μM[163]. These inhibitors included dasatinib, imatinib, sunitinib, sorafenib, BIRB-796, VX-680, VX-745, CI-1033, and linifanib (ABT-869). None of the inhibitors screened were designed to target DDR1 and DDR1 has not been clearly established a medically interesting target. DDR1 was found to bind imatinib, dasatinib, and nilotinib more tightly than Abl kinase, making it the most tight binding target for those inhibitors[123]. How DDR1 promiscuously binds such chemically diverse inhibitors is unknown.

DDR1 is a member of the DDRs, a unique two member RTK family that binds and is activated by extracellular collagen[98,99]. The DDRs act as ECM sensors for the cell, allowing the cell to detect whether it is adhered to the basement membrane or to adjacent cells[100]. DDR1 is involved in numerous signaling pathways, such as ECM remodeling, cell differentiation, proliferation, and migration; these processes are dependent on sensing the extracellular environment[102,105]. There are five known DDR1 isoforms (DDR1a-e); three of which are active RTKs (DDR1a, b, and c), and two that are inactive RTKs (DDR1d and e)[100]. The most commonly expressed isoforms are DDR1a and DDR1b[105].

DDR1 is structurally organized like other RTKs, possessing an extracellular region connected to a cytoplasmic kinase domain by a single α helical transmembrane region[15,105,107,108,113,117]. The N-terminal extracellular region of DDR1 includes the collagen binding discoidin homology (DH) domain, a discoidin like (DL) domain, and a heavily glycosylated extracellular juxtamembrane region[110,119]. The transmembrane region links to the intracellular juxtamembrane region and the cytoplasmic kinase domain. The intracellular juxtamembrane region contains several phosphotyrosine motifs needed for SH2/PTB domain containing adaptor/signaling proteins to bind to DDR1[60]. This region differs in length for the active DDR1 isoforms. The length of the juxtamembrane domain for DDR1b and c is 171 amino acids, while DDR1a's juxtamembrane is only 134 amino acids[100,105]. The missing amino acids in DDR1a result in the loss of a phosphotyrosine motif that the SH2 domain of ShcA binds to interact with DDR1, which illustrates the involvement of specific DDR1 isoforms in different signaling pathways[60,117]. C-terminal of the intracellular juxtamembrane region is the kinase domain, which is 37% sequentially homologous to TrkA, MuSK and IRK[123,124].

When compared to other RTKs, DDR1 exhibits unusually slow and sustained activation[98,99]. Upon collagen binding to DH domain, DDR1 can take several hours to become fully autophosphorylated[98,99]. This is shockingly slow with respect to other RTKs, such as EGFR and IRK, which undergo trans-autophosphorylation on the timescale of seconds to minutes[15,100]. Following its full activation, DDR1 can remain in an autophosphorylated state for up to 4 days[98]. This sustained activation may allow DDR1 to serve as a delayed feedback mechanism for other signaling pathways that it interacts with, such as the NF-κB or PI3K signaling pathways[100]. Why DDR1 has such slow activation kinetics with respect to other RTKs is not understood.

In this study, we set out to determine why DDR1 promiscuously binds ATP-competitive small molecule kinase inhibitors and the cause of its slow activation. Recently, structures of DDR1b bound to imatinib and ponatinib revealed how DDR1 binds Type II inhibitors[124]. These structures showed the kinase in an inactive conformation, which provided insight into how the kinase may be regulated, as well as potential interactions that cause it to promiscuously bind Type II inhibitors. Since the inactive conformation is known, we sought to determine the structure of DDR1a in the active conformation to better understand why the kinase has such slow activation kinetics, as well as how it promiscuously binds Type I inhibitors. We solved two co-crystal structures of the kinase domain of DDR1a bound to the pan-tyrosine kinase inhibitor dasatinib and the S/TK Aurora kinase inhibitor VX-680. The biochemical and structural data generated in this study reveal why DDR1 has slow activation kinetics and how its promiscuous behavior is achieved.

4.2 Results

4.2.1 Structure determination of DDR1a•dasatinib and DDR1a•VX-680

We sought to elucidate the mechanism DDR1's promiscuous binding of small molecule kinase inhibitors, as well as its slow activation kinetics. By binding DDR1a to Type I inhibitors dasatinib and VX-680, we hoped to determine the active conformation of the kinase by X-ray protein crystallography. Structural comparison to the DDR1b structures would provide insights into DDR1 promiscuity and slow activation.

The purified kinase domain of DDR1a was complexed with dasatinib at 10 mg/mL protein, 423 μM inhibitor, in buffer containing 20 mM Tris pH 8.5, 125 mM NaCl, 5 mM β-mercaptoehanol and 10% glycerol. DDR1a•dasatinib crystals were grown using the hanging drop vapor diffusion method and micro seeding in a mother liquor of 18% PEG 3350 and 0.1 M Bis-Tris pH 5.5. Micro seeds were generated from DDR1a•dasatinib crystals grown previously in a mother liquor of 22% PEG 3350, and 0.1 M Bis-Tris pH 5.5. Crystals were cryoprotected in mother liquor plus 20% ethylene glycol and stored in liquid nitrogen.

The complex of DDR1a kinase domain with VX-680 was formed at 10 mg/mL protein and 423 μM VX-680 in an identical buffer to the DDR1a•dasatinib complex. The hanging drop vapor diffusion method and DDR1a•dasatinib micro seeds were used to grow crystals in a mother liquor of 18% PEG 3350, 0.05 M NH_4I and 0.1 M Bis-Tris pH 5.5. DDR1a•VX-680 crystals were cryoprotected in mother liquor plus 20% glycerol, and stored in liquid nitrogen.

DDR1a•VX-680 and DDR1a•dasatinib crystals diffracted to a resolution of 2.2 Å and 2.6 Å respectively (Table 4.1). Data for both complexes were processed in space group $P2_12_12_1$ using XDS[215] and Aimless[216] as implemented in the autoPROC pipeline[179]. The structures were solved by molecular replacement using the kinase domain of DDR1b bound to imatinib (PDB entry 4BKJ; residues 599-913) with the αC-helix (residues 660-680), A-loop (residues 775-814), and ligand removed as a search model in Phaser[218].

The structures show the conserved bi-lobal kinase domain, with the P-loop and αC-helix fully resolved in both structures. The A-loop was fully resolved in our structure of DDR1a bound to VX-680, while the DDR1a•dasatinib A-loop was missing Met750 and Ser751. Two disordered loops were unresolved in both structures: the first loop is an insert between β2 and β3

of the N-lobe and the second is the 12 amino acid kinase insert domain in between helices αD and αE.

4.2.2 DDR1a inactive conformation resembles inactive IRK

In our structures, DDR1a bound to both inhibitors in the Abl/c-Kit-like inactive conformation (Figure 4.1)[56]. In this inactive conformation, the side chain of Asp747 of the DFG motif faces out the active site and the side chain of Phe748 faces into the active site. The αC-helix is rotated inwards, maintaining the conserved salt bridge between Lys618 (catalytic lysine) and Glu635 of the αC-helix. The inactive conformation of DDR1a resembles the structures of inactive IRK, MuSK, and TrkA (PDB entries 1IRK, 4GT5, and 1LUF), three kinases which are homologous to DDR1[49,51,54] (Figure 4.1). In the homologous kinase structures, amino acids on the A-loop occlude the ATP and peptide binding regions of the active site. The conformation of the A-loop in both these kinases is stabilized by interactions with the N- and C-lobe[51]. In DDR1a, the A-loop also forms similar interactions between the P-loop, the αC helix, and the αD helix of the C-lobe. These interactions are described below using the structure of DDR1a•VX680.

Starting with the DFG motif, Asp747 faces out of the active site of inactive DDR1a (Figure 4.1). The sidechain of the DFG Asp is hydrogen bonded via a water molecule to the carbonyl of Met639 of the αC-helix. The same water molecule also coordinates an additional hydrogen bond to the carbonyl of Ile648. This water-mediated hydrogen bond network helps stabilize the αC-helix in the Abl/c-Kit-like inactive conformation, preventing the outward rotation of the αC-helix seen in the Src/Cdk-like inactive conformation. The side chain of Phe748 faces into the active site and is stabilized by hydrophobic interactions with the side

chains of Met639 of the αC-helix, Met662, and Met750 of the A-loop. Gly749 of the DFG motif hydrogen bonds directly with side chain of Glu635 of the αC-helix and could potentially form an additional hydrogen bond with Lys618 (catalytic lysine). The interaction of Gly749 with Lys618 would help stabilize the salt bridge between Lys618 and Glu635 of the αC-helix, similar to the contribution of Gln275 to the stability of the conserved Lys272-Glu280 (chicken c-Src numbering) salt bridge in the P-loop of Src kinase[25,217]. Immediately following the DFG motif is Met750, which is a conserved amino acid in IRK[51]. In the inactive conformation of both DDR1 and IRK, the side chain of this methionine occludes the ATP binding pocket, preventing ATP from binding (Figure 4.1)[49,51]. The inactive conformation is further stabilized by Ser751 hydrogen bonding to the amide of Glu585 on the P-loop. This hydrogen bond is unique to the DDR1a•VX-680 structure. When compared to the other inactive DDR1 structures (PDB entries 3ZOS, 4BKJ, and 4CKR), the N- and C-lobes in DDR1•VX-680 are closer together in proximity, allowing this bond to form. This suggests that the Glu585-Ser541 hydrogen bond stabilizes a more 'closed' kinase domain[124,236]. DDR1's A-loop features a unique salt bridge, formed between Arg752 and Asp671 on the αD-helix (Figure 4.1). DDR1 is the only RTK known to contain a salt bridge at this position. In MuSK, which has an arginine at the equivalent position on the A-loop, the arginine does not form a salt-bridge and the side chain is facing out into the solvent[54]. This salt-bridge could make the inactive conformation of DDR1 more stable than the inactive conformations of IRK, TrkA, and MuSK.

Within the A-loops of MuSK, TrkA, and IRK, there is a conserved autophosphorylation motif that is also found in DDR1[123]. These kinases contain an YxxxYY motif, which requires all three tyrosines to be autophosphorylated to reach maximal kinase activity and is part of the autoregulatory mechanism of these kinases[49-53]. In DDR1a, the YxxxYY motif (Tyr755, 759,

760) is also part of its cis-autoinhibition mechanism (Figure 4.1). The side chain of Tyr755 faces towards the C-lobe of the kinase, and hydrogen bonds with the side chain of Asp671 (mentioned previously). Tyr759 acts as a pseudo-substrate, positioning itself where the tyrosine of a protein substrate would bind, thus occluding the substrate-protein binding site[49-51]. The positioning of Tyr759 in this site is coordinated by a hydrogen bond with Asp729 on the C-lobe of the kinase. Unlike the other two tyrosines of the YxxxYY motif, Tyr760 faces out into the solvent and does not mediate any of the cis-autoinhibitory interactions of the A-loop (Figure 4.1). The conformation of the YxxxYY motif in DDR1 is identical to the YxxxYY motif found in the inactive structures of IRK, TrkA, and MuSK (Figure 4.1)[49,51,54]. This suggests that DDR1 may also be autophosphorylated in a similar fashion as those kinases, and potentially has a similar A-loop active conformation[50].

4.2.3 Structural basis for DDR1a inhibition by VX-680 and dasatinib

We solved two structures of DDR1a bound to the Type I inhibitors VX-680 and dasatinib. VX-680 is an inhibitor of the S/TK Aurora kinase family, Abl kinase and most recently was found to inhibit DDR1[163,237]. Dasatinib is a pan-tyrosine kinase inhibitor, inhibiting several kinase families such as Abl family kinases and SFKs[169]. Curiously, the kinase bound both inhibitors in the Abl/c-Kit-like inactive conformation (Figure 4.2). This was unexpected since VX-680 and dasatinib are Type I inhibitors and typically bind in the active kinase conformation.

In previously solved structures of VX-680 in complex with Aurora A, Aurora B and Abl2, VX-680 bound to these kinases in the active conformation (PDB entries 4AF3, 3E5A, 2XYN, and 4B8M)[238-240]. This makes our structure of inactive DDR1a bound to VX-680 unique since it

is the first structure of a VX-680 bound to a non-mutated kinase in the Abl/c-Kit-like inactive conformation. Compared to previous structures of VX-680 bound to Aurora and Abl kinases, VX-680 forms similar interactions with DDR1a[238-240]. In particular, VX-680 forms a total of four hydrogen bonds with DDR1a: two direct hydrogen bonds with Met667 and Asp665, and two water mediated hydrogen bonds with Glu577 and Arg733. Furthermore, VX-680 occupies a rather hydrophobic section of the active site, outlined by the side chains of amino acids along the P-loop, the hinge region, and the A-loop of the kinase. Most interestingly is how the cyclopropane group of VX-680 is protected from the solvent by the side chains of Phe748 of the DFG motif and Met750 (Figure 4.2). This satisfies the requirement for the cyclopropane group to bind in a hydrophobic region of the active site[238-240]. In the Aurora A/B and Abl structures, the P-loop of both Aurora and Abl kinases is able to bend over the active site, forming a hydrophobic cage around the inhibitor[228,238-240]. The P-loop of DDR1a is unable to adopt this kinked conformation due to a series of interactions within the N-lobe of the kinase, mediated by the side chains of Lys758, Glu581, Glu586, His588, and Arg621. These interactions stabilize the extended conformation of the P-loop, preventing it from adopting the kinked conformation. DDR1a manages to compensate for the inability to form a hydrophobic cage with its P-loop by forming a hydrophobic pocket via interactions between the A-loop and the kinase only accessible in the Abl/c-Kit like inactive conformation (Figure 4.2, and 4.3).

Similar to VX-680, dasatinib binds to DDR1a in the Abl/c-Kit-like inactive conformation (Figure 4.2). Interestingly, the DFG motif differs from the DFG motif of the DDR1a•VX-680 structure. In DDR1a•dasatinib, the DFG motif is in a transitionary state between the Abl/c-Kit like inactive conformation and the active conformation. Asp747 and Phe748 of the DFG motif appear to be undergoing the 180° flip to the active conformation. This is noted by difference in

the positioning of the backbone of the DFG motif compared to the DDR1a•VX-680 structure, as well as by the side chains of Asp747 and Phe748. Both Asp748 and Phe749 have undergone a 48°/80° and 41°/2.2° change in their φ/ψ angles respectively when compared to the DDR1a•VX680 structure (Figure 4.2). The remainder of the A-loop appears to be in the same inactive conformation as the DDR1a•VX-680 structure.

When compared to the other dasatinib bound kinase structures, the interactions between DDR1a and dasatinib are similar[169]. DDR1a makes a total of two hydrogen bonds with dasatinib. The first is formed between dasatinib and the carbonyl of Met667 on the hinge region of the kinase. The second hydrogen bond is between dasatinib and the gatekeeper residue Thr664. Dasatinib binding is further stabilized by a series of electrostatic interactions with Tyr666, Glu668, and Gln674. In addition to electrostatic interactions, dasatinib has numerous hydrophobic interactions with the side chains of Leu579, Val587, Ala616, Lys618, Met639, Ile648, Met662, Leu736, and Ala746. The side chain of Met750 of the A-loop also contributes to the hydrophobic region dasatinib occupies, though this was not built in the structure. Positive electron density is present for the side chain of Met750 suggesting that it faces into the active site, similar to the side chain of Met750 in the DDR1a•VX680 structure; however there is little density for the backbone of Met750.

4.2.4 The molecular mechanism for DDR1 promiscuity

Recent studies have shown that DDR1 promiscuously binds ATP-competitive small molecule inhibitors[123,163]. How DDR1 achieves this promiscuous behavior is unknown and could not be addressed until structures of the kinase domain were solved. Recently, two structures of

the kinase domain of DDR1b bound to imatinib and ponatinib were solved (the kinase domains of DDR1a and b are sequentially identical) (PDB entries 3ZOS, 4BKJ)[124]. The structures showed that the inhibitors bound to DDR1 in the Abl/c-Kit-like inactive conformation. The inhibitors binding to this conformation was expected since imatinib and ponatinib target the Abl/c-Kit-like inactive conformation[235,241].

In the DDR1b•imatinib and DDR1b•ponatinib structures, DDR1b is able to bind these inhibitors due to the hydrophobic pocket formed by the interactions between the A-loop and the kinase (Figure 4.3)[124]. In Abl, imatinib and ponatinib both require the P-loop to fold over the active site to protect the more hydrophobic core of the drug from the solvent[235,241]. Since DDR1b cannot fold its P-loop over its active site, the hydrophobic interactions required for binding are met by the formation of a hydrophobic pocket through the salt bridge of Arg789 (Arg752 in DDR1a) in the A-loop with Asp708 (Asp671 in DDR1a) on the αD-helix on the C-lobe of the kinase (Figure 4.3). This hydrophobic pocket is further stabilized by interactions between the A-loop and P-loop.

Interestingly, when comparing the structures of DDR1a bound to VX-680 and dasatinib to the structures of DDR1b bound to imatinib and ponatinib, we see several differences in the structure of the kinase domain, particularly in the conformation and the positioning of the A-loop. Comparing the A-loops of the DDR1b•imatinib and DDR1b•ponatinib (PDB entries 4BKJ, 3ZOS) to the DDR1a•VX-680 structure shows that the A-loop has shifted downwards to accommodate the larger inhibitors and allow them access to the hydrophobic pocket underneath the αC-helix (Figure 4.3)[124]. Most noticeable is how the A-loop near the DFG motif shifts towards the C-lobe by 4.2 Å and how Met787 (Met750 in DDR1a) faces away from the active site. The shifting of the A-loop results in a larger hydrophobic pocket, which is only accessible in

DDR1 by the Abl/c-Kit-like inactive conformation DDR1. The pocket remains intact due to the salt bridge between Arg752 and Asp671 (DDR1a numbering) anchoring the A-loop at that position[124] (Figure 4.3). Similarly, our structures of DDR1a bound to VX-680 and dasatinib show how these inhibitors also bind in the same hydrophobic pocket as imatinib and ponatinib (Figure 4.3). We therefore propose that the flexibility of the A-loop and the formation of this large hydrophobic pocket in the inactive DDR1 kinase domain is how DDR1 is able to promiscuously bind so many different small molecule kinase inhibitors.

4.2.5 Molecular basis for DDR1a's low kinase activity

In order to better understand how the autoregulation of DDR1 relates to the slow activation and low kinase activity of DDR1, we performed *in vitro* substrate peptide phosphorylation assays using [γ-^{32}P] ATP. We used this assay instead of the continuous spectrophotometric assay because of its increased sensitivity for detecting low kinase activity[184]. To assess whether the kinase has low kinase activity due to low affinity for ATP, we measured the K_M for ATP for DDR1a$_{526-876}$. We determined the K_M for ATP was 19.5 μM, showing that affinity for ATP was not the cause of DDR1's poor kinase activity. To assess what the cause of this low kinase activity is, we performed several experiments to fully activate the kinase, such as autophosphorylation, dimerization, and site directed mutagenesis.

Previous studies have shown that for DDR1 to reach maximal kinase activity in cell, all three tyrosines of the A-loop must be phosphorylated[98-100]. Unlike other RTKs, DDR1 autophosphorylation is a slow process, taking several hours to reach the fully autophosphorylated state[98,99]. Similar to the cellular studies, the isolated kinase domain also takes several hours to

become autophosphorylated, with di- and tri-phosphorylated species of DDR1a$_{526-876}$ first appearing after 24 hours. Furthermore, the autophosphorylation of the kinase only resulted in a 2-fold increase in kinase activity compared to unphosphorylated DDR1a (24 fmol sec^{-1} to 54 fmol sec^{-1}; Figure 4.4).

To assess whether DDR1's low kinase activity is strictly caused by the kinase domain, or is allosterically regulated by its intracellular juxtamembrane region, we measured the affect of autophosphorylation of DDR1a$_{485-876}$ on kinase activity. Similar to kinase domain, autophosphorylation of DDR1a$_{485-876}$ only resulted in minor change, increasing activity approximately 3-fold compared to the unphosphorylated DDR1a$_{485-876}$ (5 fmol sec^{-1} to 17 fmol sec^{-1}). The overall activity of DDR1a$_{485-876}$ was also lower than the isolated kinase domain of DDR1a by approximately 4-fold. We hypothesize that the intracellular juxtamembrane occludes the active site of the kinase, making it more difficult for substrate peptide and ATP to bind. Secondary structure prediction through the Phyre2 server shows that the juxtamembrane region is disordered, which could suggest that it is flexible. This flexibility could potentially allow it to fold over the kinase domain and block the active site.

RTKs form dimers upon ligand binding, which results in the activation of their kinase domains[15]. In a non-dimerized state, the individual RTK has little kinase activity. This is seen in kinases like EGFR, where it has low kinase activity as a monomer, but upon dimerization, its kinase activity increases significantly[48]. Past studies have shown that DDR1 is expressed as constitutive dimers that are formed by interactions between the extracellular domains of the kinase[105]. To determine whether the intracellular portion of DDR1 form dimers, we performed small angle x-ray scattering (SAXS) using both DDR1a$_{485-876}$ and DDR1a$_{526-876}$. SAXS data collected showed that the samples for DDR1a$_{526-876}$ were aggregated as well as experienced

radiation damage for all three concentrations of protein screened (2.5 mg/mL, 5 mg/mL, and 10 mg/mL). Similarly, DDR1a$_{485-876}$ samples at 2.5 mg/mL and 5 mg/mL also had radiation damage, however the 10 mg/mL sample showed neither soluble aggregation nor radiation damage. Unfortunately the data collected for the 10 mg/mL sample of DDR1a$_{485-876}$ was of poor quality and could not be processed to determine a reliable R_G.

Zhang *et al.* demonstrated that the isolated kinase domain of EGFR has little activity on its own, however, when clustered onto lipid vesicles to simulate dimerization of EGFR on the plasma membrane, the dimerization allosterically activates EGFR, making the kinase several-fold more active than the monomer [48]. Although we could not show that the intracellular regions of DDR1 form dimers via SAXS, we were curious to see whether we could allosterically activate DDR1 by clustering it on lipid vesicles. We clustered His6-tagged DDR1a$_{526-876}$ onto lipid vesicles to measure the affect of clustering on kinase activity. Unlike EGFR, the kinase activity of DDR1a only increased by approximately 2-fold. These studies were not carried out on DDR1a$_{485-876}$ due to inconsistencies in DDR1a$_{485-876}$ kinase activity.

To better understand why DDR1a has such low activity, we examined the structures of DDR1a and DDR1b. Interestingly, all structures of DDR1 have been of the inactive form, even with an inhibitor that is known to bind to the active conformation in other kinases[124]. This suggests that the DDR1 inactive conformation is its most stable form and DDR1 prefers to remain in this state. The additional salt-bridge between Arg752 of the A-loop and Asp671 of the αD-helix may cause the transition of DDR1 from the inactive to active conformation to take longer than other RTKs. This is also suggested by the DDR1a•dasatinib structure, where only the DFG motif has altered its conformation, while the rest of the A-loop has remained in the same conformation as in the DDR1a•VX-680 structure, including the Asp671-Arg752 salt-

bridge (Figure 4.4). Similar to the other YxxxYY motif containing RTKs, Tyr755 and Tyr759 are hydrogen bonded to residues on the C-lobe and both are part of the kinase autoregulatory mechanism. Tyr755 forms a hydrogen bond with Asp671, while Tyr759 hydrogen bonds with Asp729 (Figure 4.4). We hypothesized that by mutating Asp671 to asparagine, we would disrupt the salt-bridge between the A-loop and the C-lobe, but not disrupt the hydrogen bond with Tyr755, resulting in a more active kinase. Furthermore, mutation of Tyr755 and Tyr759 to alanine should also result in destabilization of the inactive conformation.

Introducing the Asp671Asn (D671N) into DDR1a$_{526-876}$ resulted in a 9.3-fold increase in activity of the kinase compared to wild-type DDR1a (activity increased from 24 fmol sec^{-1} to 224 fmol sec^{-1}; Figure 4.4). This increase in activity suggests that the Asp671-Arg752 salt-bridge was critical for keeping DDR1a in its inactive conformation and by mutating it allows the kinase to more readily adopt the active conformation. The Tyr755Ala (Y755A) and Tyr759Ala (Y759A) mutations also resulted in a 7.5-fold and 7.8-fold increase in activity when compared to wild-type DDR1a, further verifying that the stability of the inactive conformation.

4.2.6 Development of a cellular assay to assess effects of mutations on DDR1 activation

Previous studies have shown that DDR1 signaling is required for megakaryocyte migration[188]. In these cells, collagen-activated DDR1 phosphorylates the phosphatase Shp1, activating it and leading to the dephosphorylation of Syk kinase by Shp1. As a consequence of the dephosphorylation of Syk, megakaryocytes can migrate across a matrix of collagen I[188]. Using this as a model for the development of a cellular assay to assess how mutations that destabilize the inactive conformation of DDR1 affect DDR1 signaling, we established a cellular

assay using DDR1a$_{1-876}$ transfected UT-7 human megakaryocyte cells (Figure 4.5). The advantage of using this cell line is that they do not endogenously express any isoform of DDR1, so any effects seen due to collagen stimulation can be directly related to DDR1 activation. We stimulated DDR1 activation with collagens I and IV and harvested cells at time points of 0, 2, 16, and 24 hours to test for DDR1a activation (Figure 4.5). Interestingly, maximal DDR1a autophosphorylation was seen at 2 hours; by 16 and 24 hours DDR1 autophosphorylation levels were equivalent to those seen at 0 hours (Figure 4.5). With further development, this system can be used to probe DDR1a mutant activity in a time course assay to assess how quickly mutations, such as Asp671Asn, can affect the rate of DDR1 activation and the phosphorylation of downstream targets such as Shp1.

4.3 Discussion

DDR1a is a unique RTK that is characterized by its ability to bind extracellular collagen, slow activation, and low kinase activity[98,99]. In this study, we solved two structures of the kinase domain of DDR1a bound to dasatinib and VX-680. The structures showed the kinase binds to dasatinib and VX-680 in the Abl/c-Kit-like inactive conformation. Furthermore, binding VX-680 and dasatinib to an inactive or inactive-like conformation is unusual since both bind to the active conformation in other kinases and reveals that the inactive conformation of the kinase may be its most preferred state[169,238,240].

Our structures, together with the structures of DDR1b, reveal how the inactive conformation of DDR1 allows the kinase to promiscuously bind small molecule kinase inhibitors. It achieves this promiscuity by the A-loop having multiple interactions with C-lobe of the kinase,

resulting in the formation of a large, flexible, hydrophobic pocket that inhibitors can bind to. This pocket allows larger inhibitors, such as imatinib and ponatinib, to bind by shifting the DFG motif and surrounding residues to accommodate the drug. Furthermore, this hydrophobic pocket is anchored around the Asp671-Arg752 salt-bridge[124]. Additional hydrogen bonds between the A-loop and C-lobe distal from the salt-bridge only help stabilize this inactive conformation.

When comparing to other promiscuous kinases identified in the same study as DDR1, such as PDGFR A/B, CSF1R, and c-Kit, DDR1 differs in that it is the only promiscuous kinase that does not fall within the PDGFR family[163]. Many of the inhibitors screened were designed to target kinases closely related to the PDGFR family, such as Abl and EGFR family kinases. Furthermore, the sequence identity conservation of the kinase domains of these promiscuous kinases is rather low when compared to DDR1, having only 33-36% sequence identity to DDR1. When structurally comparing the structures of DDR1b and c-Kit bound to imatinib (PDB entries 4BKJ and 1T46, respectively), DDR1b differs with the presence of the salt bridge between Asp671 and Arg752. In c-Kit, the activation loop forms an α-helix following the DFG-motif that hydrogen bonds to the P-loop. These secondary structure interactions in c-Kit provide the hydrophobic contacts necessary for imatinib binding, and may potentially explain how it promiscuously binds Type II inhibitors. Interestingly, structures of c-Kit bound to Type I inhibitors show the kinase in an active conformation, unlike DDR1, which binds all inhibitors in the Abl/c-Kit-like inactive conformation. The inactive conformation of DDR1 seen in all the DDR1 structures suggests that its promiscuity is related to the inactive state due to the increased hydrophobic surfaces available for inhibitors to interact with in the inactive conformation. This differs from structures of CSF1R and c-Kit where hydrophobic surfaces available for binding Type I inhibitors are reduced in the active state of these kinases.

The slow kinase activity of DDR1a is also explained by the structures we solved. Disruption of the inactive conformation by the D671N mutation resulted in the loss of the stabilizing salt-bridge seen in the structures, however it did not disrupt the hydrogen bonding with Tyr755 of the YxxxYY motif. This mutation resulted in the largest increase in kinase activity compared to wild-type DDR1a, suggesting that this salt-bridge was critical for mediating DDR1's slow activation. Furthermore, we can hypothesize that the loss of this salt bridge will also result in DDR1 being less promiscuous for kinase inhibitors due to the hydrophobic pocket no longer being anchored to the kinase as tightly as it was previously. We further verified that the inactive conformation of DDR1a was also destabilized by the Y755A and Y759A mutations.

Last, we established a cellular assay for DDR1a activity using UT-7 human megakaryocyte cells[188]. While the data concerning the proper time course for examining DDR1a activation is preliminary, it shows promise of being a successful assay for measuring both DDR1 autophosphorylation and potentially the phosphorylation of downstream substrates of DDR1, such as Shp1. This assay will allow us to examine the effect of DDR1 mutations much more quickly than *in vitro* kinase assays using purified protein, as well as provide insights into the global effects of DDR1 mutations.

The data in this study has allowed us to better understand the contribution of the DDR1 kinase domain in its slow activation, suggesting that the delayed signaling seen in cells may be regulated by internal interactions within the kinase domain. Furthermore, DDR1's promiscuous binding of small molecule inhibitors is also related to the inactive conformation of kinase. Future inhibitors will have to overcome this promiscuity, potentially by targeting allosteric sites rather than the active site of the kinase.

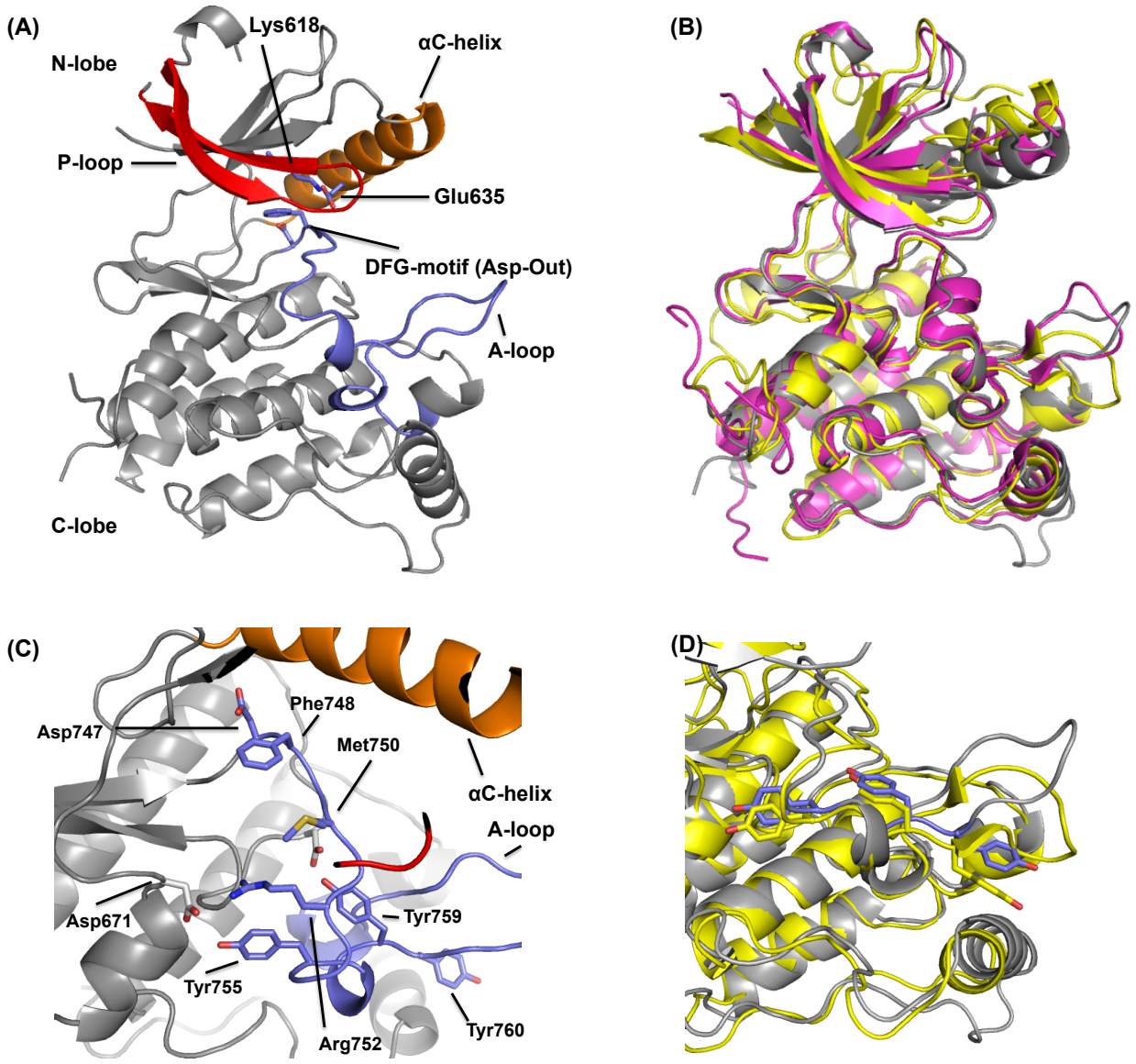


Figure 4.1 – The inactive conformation of DDR1a. (A) Structure of DDR1a•VX-680 shows the kinase in the Abl/c-Kit like inactive conformation. Asp747 faces out of the active site. (B) Structural comparison of the inactive conformations of DDR1a (grey), IRK (yellow), and TrkA (purple) (PDB entries 1IRK, 4GT5)[49,51]. (C) The activation loop of DDR1a forms interactions with the αC-helix and C-lobe in the inactive conformation. Phe748 and Met750 occlude the ATP binding site/ Arg 752 forms a salt-bridge with Asp671 of the C-lobe. Tyr755 and Tyr759 also hydrogen bond to the C-lobe. (D) Comparison of YxxxYY motif between IRK (yellow) and DDR1a (blue).

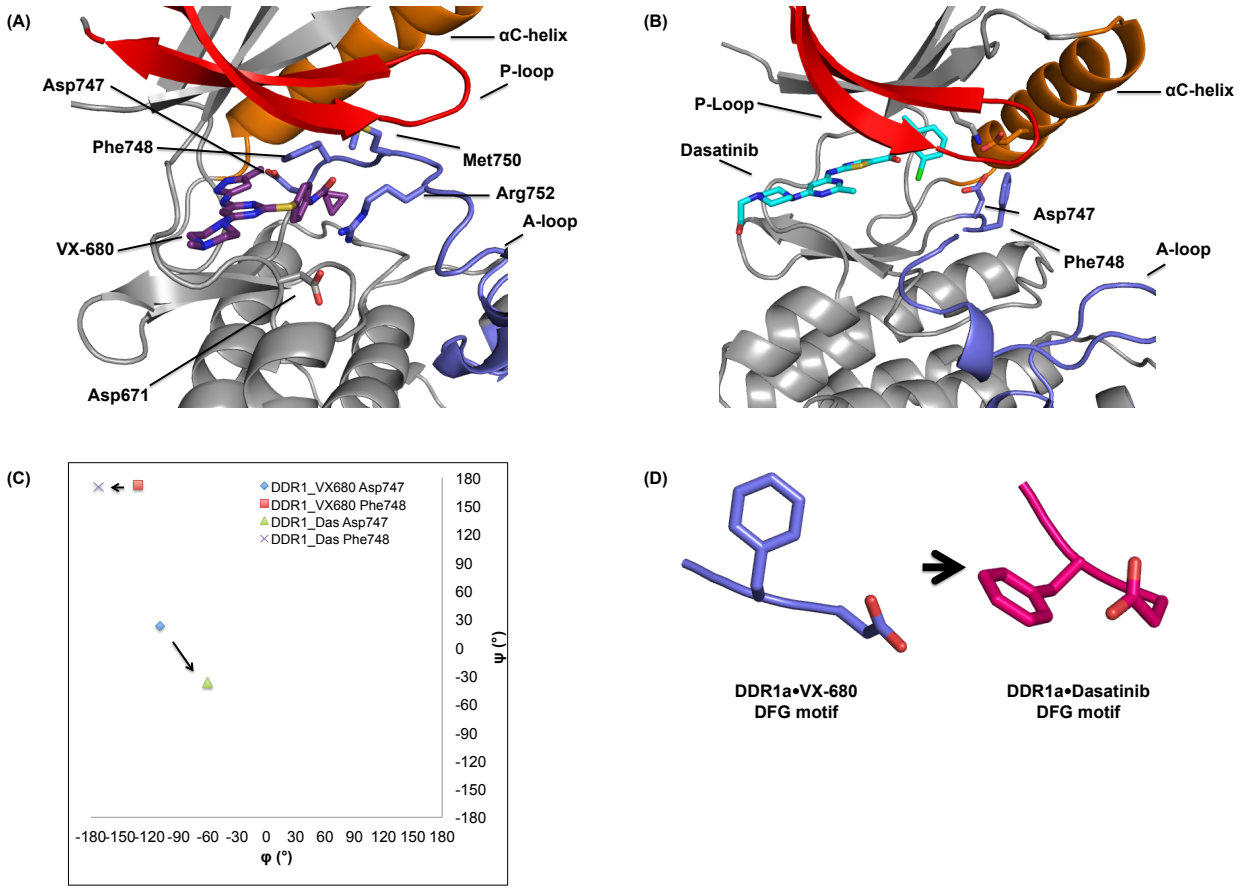


Figure 4.2 – Structural basis for inhibition of DDR1a by VX-680 and dasatinib. (A, B) VX-680 and dasatinib bind to DDR1a in the Abl/c-Kit-like inactive conformation. A hydrophobic pocket is formed by the A-loop that shields the compounds from the solvent. (C) Ramachandran plot comparing the φ/ψ angles of the DFG motifs of DDR1a•VX680 and DDR1a•dasatinib. (D) The DFG motifs of the two DDR1a structures solved show how the DFG motif of dasatinib appears to be transitioning to the active conformation.

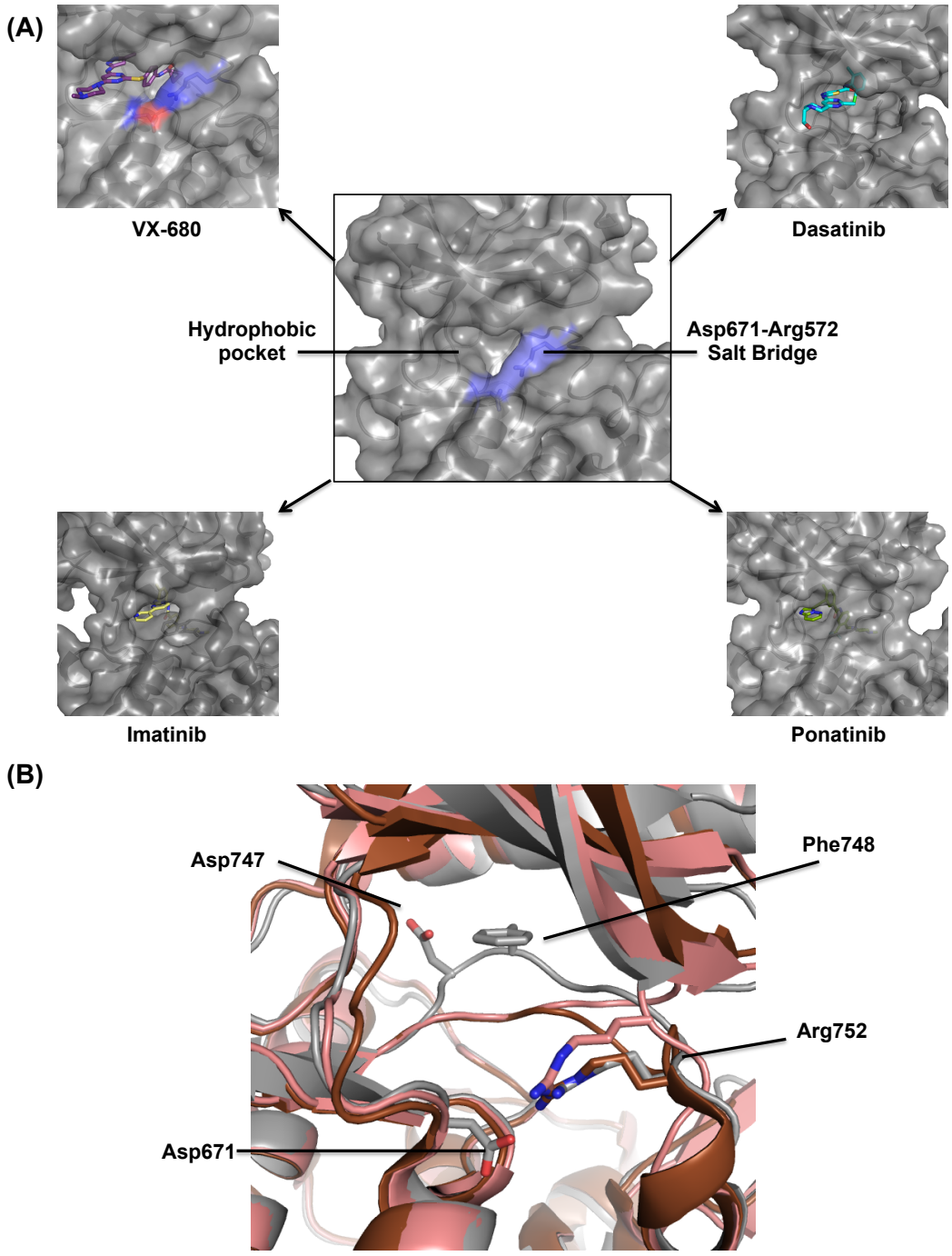


Figure 4.3 – Molecular basis for DDR1's promiscuous binding of small molecule inhibitors. (A) The A-loop of DDR1 forms a flexible hydrophobic pocket in the inactive conformation allowing it to accommodate chemically diverse inhibitors (PDB entries 3ZOS, 4BKJ)[124]. (B) The A-loop of DDR1a•VX-680 (grey) compared to the A-loops of DDR1b•ponatinib (pink) and DDR1b•imatinib (brown). The A-loop around the DFG shifts towards the C-lobe by 4 Å, however the remainder of the A-loop is anchored in place by the Asp671-Arg252 salt bridge.

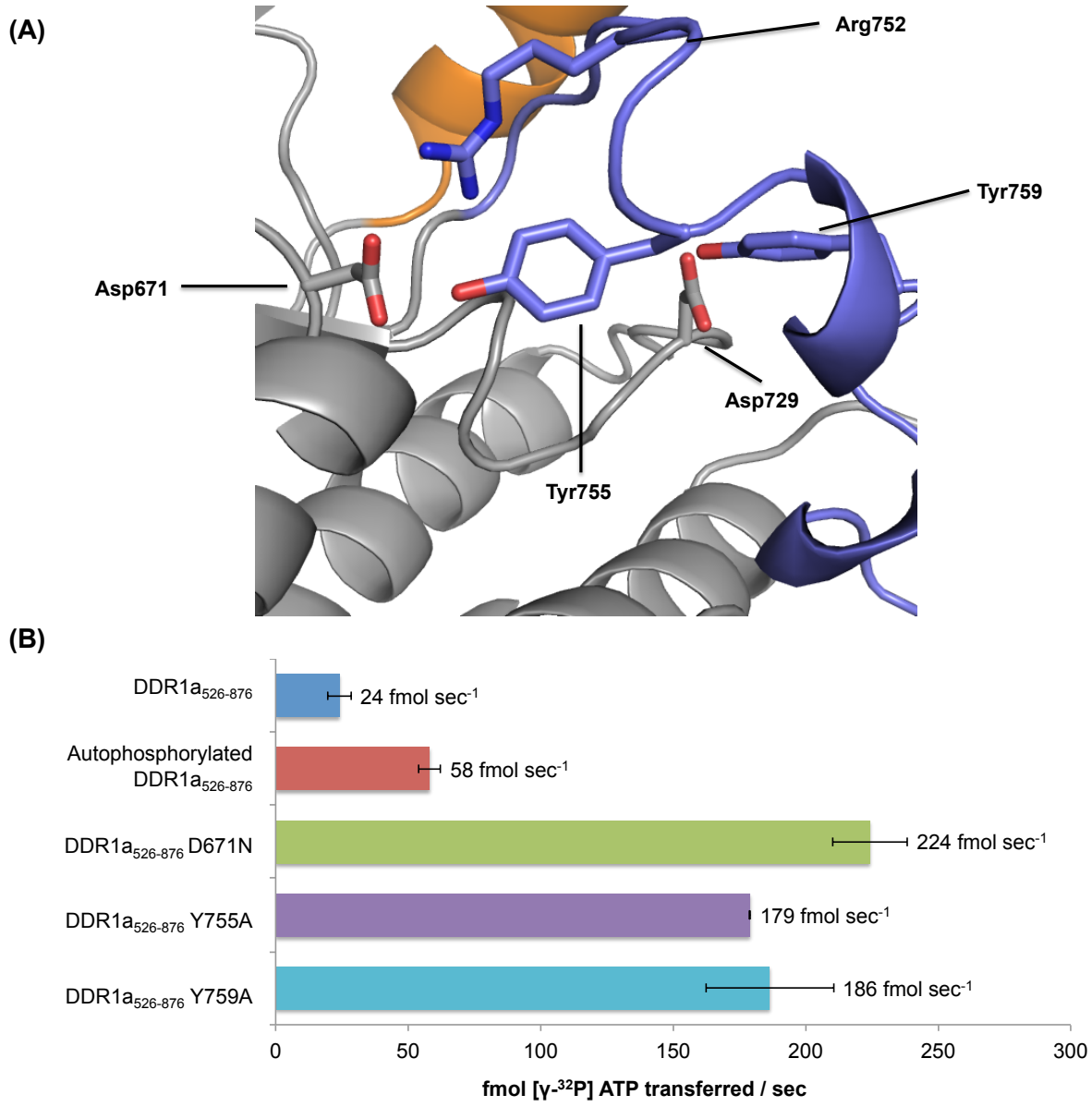


Figure 4.4 – Molecular basis for DDR1's low kinase activity. (A) The inactive conformation of the DDR1's A-loop is stabilized by a series of tyrosine-mediated hydrogen bonds and the Asp671-Arg752 salt bridge. (B) *In vitro* kinase assays of DDR1a show that the kinase can become more active by autophosphorylation or through mutation of residues in the A-loop that interact with the C-lobe of the kinase. Assays were done in duplicate using 0.5 µM kinase, 200 µM ATP, 400 µM Axltide, and 50 cpm/pmol [γ-^{32}P] ATP.

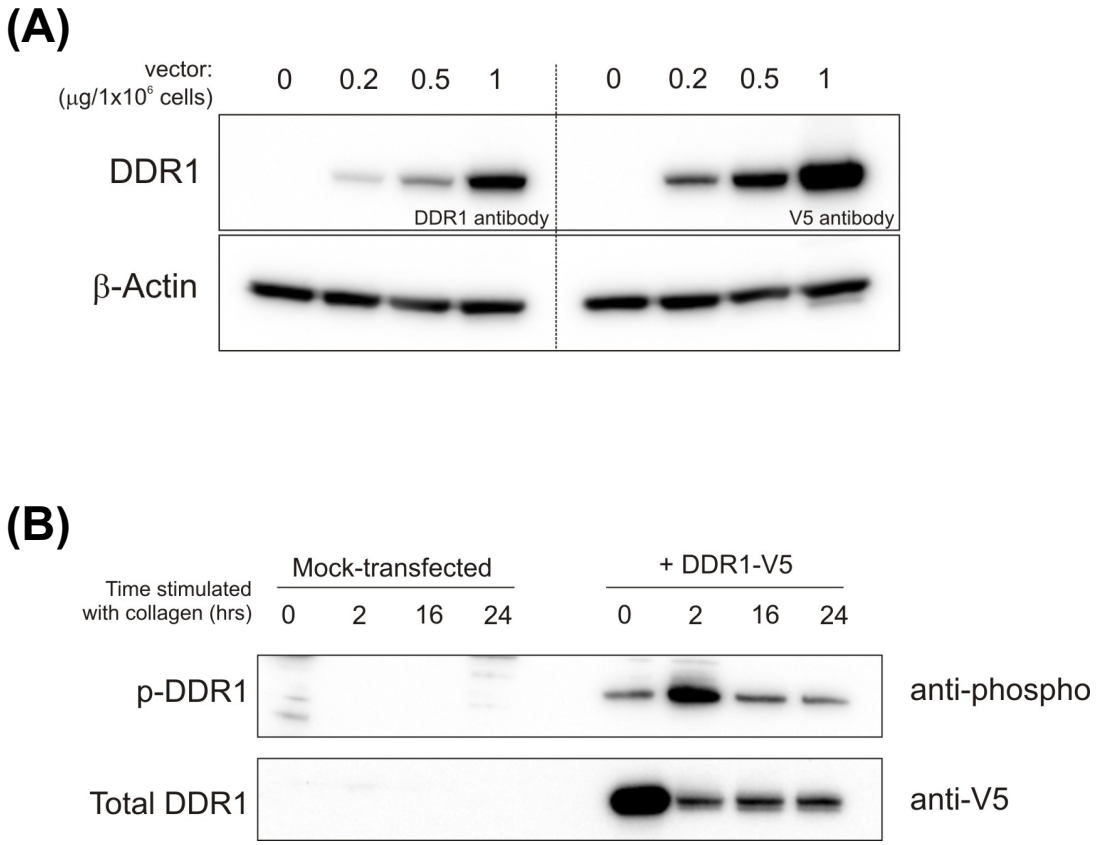


Figure 4.5 – Expression and activation of DDR1a$_{1\text{-}876}$ in UT-7 megakaryocyte cells. (A) pEF1-V5-His A DDR1a$_{1\text{-}876}$ was transfected into UT-7 cells at different concentrations. We determined that the optimal concentration to use for future transfections was 0.2 μg / 1 x 10^6 cells. (B) Time course of collagen stimulation of transfected UT-7 cells. Cells stimulated with 20 μg/mL of collagen. They were harvested and probed for DDR1 autophosphorylated at 0, 2 16 and 24 hours. DDR1 autophosphorylation was maximal at 2 hours, and had returned to basal levels by 16 hours post stimulation.

Table 4.1 Data collection and refinement statistics (molecular replacement)

	DDR1a•Dasatinib	DDR1a•VX-680
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	61.7, 72.3, 74.7	61.5, 75.4, 77.2
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	47.6 - 2.61	48.1-2.23
R_{merge}	0.183 (0.00)	0.115 (0.558)
$I / \sigma I$	13.3 (2.25)	24.4 (2.97)
Completeness (%)	99.8 (98.0)	99.3 (93.04)
Redundancy	13.2 (8.0)	12.2 (4.5)
Refinement		
Resolution (Å)	2.61	2.23
No. reflections	10,658	17,938
R_{work} / R_{free}	0.2038 / 0.2565	0.1876 / 0.2258
No. atoms		
Protein	2218	2153
Ligand/ion	33	33
Water	61	106
B-factors		
Protein	43.1	37.4
Ligand/ion	42.8	41.8
Water	39.7	42.0
R.m.s. deviations		
Bond lengths (Å)	0.002	0.012
Bond angles (°)	0.63	0.91

*Highest-resolution shell is shown in parentheses.

Chapter 5: General Discussion and Future Directions

The discovery of the role of protein kinases in human diseases has been a driving force for the generation of specific small molecule kinase inhibitors by both academics and the pharmaceutical industry for almost 3 decades[3,4]. The success of imatinib in treating BCR-Abl driven CML has further demonstrated how effective small molecule kinase inhibitors can be in the clinic[160]. As of 2013, the United States Food and Drug Administration has approved 23 small molecule kinase inhibitors for the treatment of cancer, Parkinson's disease, and rheumatoid arthritis[4]. More inhibitors are being developed and are undergoing clinical trials, hoping to follow the success small molecule kinase inhibitors have seen thus far.

Many small molecule kinase inhibitors target the ATP binding pocket since it is a large pocket with many hydrophobic surfaces for inhibitors to interact with[148]. The common usage of ATP by all protein kinases has added to the challenge of developing specific inhibitors targeting this site. Imatinib and other type II inhibitors overcome this high structural conservation by targeting an additional hydrophobic pocket in the active site that is accessible in the Abl/c-Kit-like inactive conformation[148]. Here, we demonstrated a similar strategy with the Src specific macrocycles by targeting both the ATP binding site and a hydrophobic pocket that is accessible in the Src/Cdk-like inactive conformation[217]. The low conservation of this pocket, as well as unique residues in Src, allows the macrocycles to distinguish Src from the other SFKs, and the closely related Abl kinase. Furthermore, we demonstrate that the macrocycles overcome another challenge faced by small molecule kinase inhibitors, inhibitor resistance mutations. We successfully showed *in vitro* how the different macrocycle scaffolds target different inhibitor resistance mutations.

There are still several avenues of research to explore with the macrocycles. First, the in cell potency of the macrocycles needs to be determined. Our preliminary experiments show that

the macrocycles inhibit Src autophosphorylation in HEK 293T cells. There is substantial data suggesting that Src inhibition can be efficacious in sensitizing tumor cells to other chemotherapies[90]. The macrocycles could also be screened against different cancer cell lines where Src signaling is dysregulated to determine if they potentially have any therapeutic relevance[90].

While most small molecule kinase inhibitors are designed to be selective for a set of kinases, there are kinases that will promiscuously bind inhibitors. The kinases can bind many chemically diverse kinase inhibitors. The inhibition of these promiscuous kinases can affect the function of normal cells, potentially leading to undesirable side effects reported in the use of small molecule kinase inhibitors during disease treatment. Cardiotoxicity, gastrointestinal bleeding, immune cell suppression, and bone marrow toxicity are examples of how severe these side effects can be[146]. To understand how some kinases achieve this promiscuous behavior, we used DDR1 as a model system.

DDR1 is a unique RTK that binds extracellular collagen[98,99]. It has been implicated in a number of diseases, ranging from fibrosis to cancer, but has never been established as a medically interesting target[108]. Hence, no kinase inhibitor was ever designed to target it. It was through studies examining the selectivity profile of ATP competitive small molecule kinase inhibitors that DDR1 was discovered to promiscuously bind many clinically used inhibitors[163]. In our work, we propose a model that DDR1 achieves this promiscuous behavior through its inactive conformation. The A-loop of inactive DDR1 forms a large hydrophobic pocket that allows DDR1 to bind chemically diverse inhibitors by mimicking the active site of other kinases. Furthermore, we also determined the mechanism of DDR1's slow kinase activity is related to the stability of the inactive conformation.

The work on DDR1 is far from complete and there are several avenues that can be explored. To better define the mechanism of DDR1 promiscuity, DDR1 mutants that disrupt the inactive conformation hydrophobic pocket can be screened against a panel of clinically used kinase inhibitors using both the *in vitro* kinase assay and cellular assay described here. This can provide insight into how the destabilization of the inactive conformation affects DDR1 promiscuity. Additionally, the cellular assay could be adopted to examine how inhibitor resistances can arise in DDR1. Using the system established for BCR-Abl by Azam, et al., random mutations introduced into DDR1 can be transfected into UT-7 cells and screened for inhibitor resistance by using cell migration as a phenotype for inhibitor resistance[168,242]. In a wound healing assay, transfected UT-7 cells treated with collagen will migrate to the wound if DDR1 carries an inhibitor resistance mutation[188]. Potentially, this could identify mutations that reduce the promiscuity of DDR1, allowing particular DDR1 mutants to bind one inhibitor over another.

Another avenue of research to be explored is the activation and regulation of DDR1 outside the kinase domain. Data from our study suggests that DDR1 kinase domain activity is regulated by the salt-bridge between the A-loop and the C-lobe of the kinase domain. The presence of a leucine zipper in the transmembrane region of DDR1 suggests that DDR1 undergoes conformational shifts in the intracellular region upon collagen binding to the extracellular region[104]. It would be interesting to see whether these conformational shifts from collagen binding also contribute to the slow activation of DDR1 or whether the slow activation of DDR1 is entirely based in the kinase domain switching from inactive to active conformation.

Overall, the studies presented here have revealed important aspects of the kinase-inhibitor specificity relationship. The molecular mechanism of DDR1's promiscuous behavior may be a

central theme for how other promiscuous kinases bind chemically diverse inhibitors. The exact mechanisms for promiscuity in other kinases remain to be determined; however, we can expect that like DDR1, they are able to mimic interactions particular inhibitors require for binding. Furthermore, the macrocycles provide a new model for kinase inhibition. Inhibitors designed to target multiple binding pockets in the active site can be more specific than those targeting the ATP binding pocket alone. This concept has already been demonstrated in the selectivity profiles of type I and type II inhibitors[163,232]. By designing new inhibitors that are tailored to the unique binding pockets in the active site of a particular kinase, we increase the specificity of the inhibitor and reduce the chance of it being bound by a promiscuous kinase.

References

1 Cohen, P. The origins of protein phosphorylation. *Nat Cell Biol* **4**, E127-130, doi:10.1038/ncb0502-e127 (2002).

2 Blume-Jensen, P. & Hunter, T. Oncogenic kinase signalling. *Nature* **411**, 355-365, doi:10.1038/35077225 (2001).

3 Cohen, P. Protein kinases--the major drug targets of the twenty-first century? *Nature Reviews Drug Discovery* **1**, 309-315, doi:10.1038/nrd773 (2002).

4 Cohen, P. & Alessi, D. R. Kinase Drug Discovery - What's Next in the Field? *ACS chemical biology* **8**, 96-104, doi:10.1021/cb300610s (2013).

5 Jin, J. & Pawson, T. Modular evolution of phosphorylation-based signalling systems. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **367**, 2540-2555, doi:10.1098/rstb.2012.0106 (2012).

6 Hunter, T. Why nature chose phosphate to modify proteins. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **367**, 2513-2516, doi:10.1098/rstb.2012.0013 (2012).

7 Ciesla, J., Fraczyk, T. & Rode, W. Phosphorylation of basic amino acid residues in proteins: important but easily missed. *Acta Biochim Pol* **58**, 137-148 (2011).

8 Cozzone, A. J. Protein phosphorylation in prokaryotes. *Annual review of microbiology* **42**, 97-125, doi:10.1146/annurev.mi.42.100188.000525 (1988).

9 Attwood, P. V. Histidine kinases from bacteria to humans. *Biochemical Society transactions* **41**, 1023-1028, doi:10.1042/BST20130019 (2013).

10 Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science (New York, NY)* **298**, 1912-1934, doi:10.1126/science.1075762 (2002).

11 Manning, G. *Kinase.com Genomics, evolution and function of protein kinases*, <http://kinase.com/web/current/kinbase/browser/SpeciesID/9606> (2014).

12 Neet, K. & Hunter, T. Vertebrate non-receptor protein-tyrosine kinase families. *Genes to cells : devoted to molecular & cellular mechanisms* **1**, 147-169 (1996).

13 Robinson, D. R., Wu, Y. M. & Lin, S. F. The protein tyrosine kinase family of the human genome. *Oncogene* **19**, 5548-5557, doi:10.1038/sj.onc.1203957 (2000).

14 Duong-Ly, K. C. & Peterson, J. R. in *Current Protocols in Pharmacology* (John Wiley & Sons, Inc., 2001).

15 Lemmon, M. A. & Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **141**, 1117-1134, doi:10.1016/j.cell.2010.06.011 (2010).

16 Scheeff, E. D. & Bourne, P. E. Structural Evolution of the Protein Kinase–Like Superfamily. *PLoS Computational Biology* **1**, e49, doi:10.1371/journal.pcbi.0010049 (2005).

17 Taylor, S. S. & Radzio-Andzelm, E. Three protein kinase structures define a common motif. *Structure (London, England : 1993)* **2**, 345-355 (1994).

18 al-Obeidi, F. A., Wu, J. J. & Lam, K. S. Protein tyrosine kinases: structure, substrate specificity, and drug discovery. *Biopolymers* **47**, 197-223, doi:10.1002/(SICI)1097-0282(1998)47:3<197::AID-BIP2>3.0.CO;2-H (1998).

19 Huse, M. & Kuriyan, J. The conformational plasticity of protein kinases. *Cell* **109**, 275-282 (2002).

20 Hanks, S. K., Quinn, A. M. & Hunter, T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science (New York, NY)* **241**, 42-52 (1988).
21 Hanks, S. K. & Hunter, T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **9**, 576-596 (1995).
22 Endicott, J. A., Noble, M. E. M. & Johnson, L. N. The Structural Basis for Control of Eukaryotic Protein Kinases. *Annual Review of Biochemistry* **81**, 587-613, doi:10.1146/annurev-biochem-052410-090317 (2012).
23 Knighton, D. *et al.* Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science (New York, NY)* **253**, 407-414, doi:10.1126/science.1862342 (1991).
24 Zheng, J. *et al.* Crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with magnesium-ATP and peptide inhibitor. *Biochemistry* **32**, 2154-2161 (1993).
25 Barouch-Bentov, R. *et al.* A Conserved Salt Bridge in the G Loop of Multiple Protein Kinases Is Important for Catalysis and for In Vivo Lyn Function. *Molecular Cell* **33**, 43-52, doi:10.1016/j.molcel.2008.12.024 (2009).
26 Palmieri, L. & Rastelli, G. αC helix displacement as a general approach for allosteric modulation of protein kinases. *Drug discovery today* **18**, 407-414, doi:10.1016/j.drudis.2012.11.009 (2013).
27 Johnson, L. N., Noble, M. E. & Owen, D. J. Active and inactive protein kinases: structural basis for regulation. *Cell* **85**, 149-158 (1996).
28 Kornev, A. P., Haste, N. M., Taylor, S. S. & Eyck, L. F. T. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 17783-17788, doi:10.1073/pnas.0607656103 (2006).
29 Shan, Y. *et al.* A conserved protonation-dependent switch controls drug binding in the Abl kinase. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 139-144, doi:10.1073/pnas.0811223106 (2009).
30 Huang, H., Zhao, R., Dickson, B. M., Skeel, R. D. & Post, C. B. αC Helix as a Switch in the Conformational Transition of Src/CDK-like Kinase Domains. *The Journal of Physical Chemistry B* **116**, 4465-4475, doi:10.1021/jp301628r (2012).
31 De Bondt, H. L. *et al.* Crystal structure of cyclin-dependent kinase 2. *Nature* **363**, 595-602, doi:10.1038/363595a0 (1993).
32 Chao, L. H. *et al.* A mechanism for tunable autoinhibition in the structure of a human Ca2+/calmodulin- dependent kinase II holoenzyme. *Cell* **146**, 732-745, doi:10.1016/j.cell.2011.07.038 (2011).
33 Yadav, S. S., Yeh, B. J., Craddock, B. P., Lim, W. A. & Miller, W. T. Reengineering the signaling properties of a Src family kinase. *Biochemistry* **48**, 10956-10962, doi:10.1021/bi900978f (2009).
34 Taylor, S. S. & Kornev, A. P. Protein kinases: evolution of dynamic regulatory proteins. *Trends in biochemical sciences* **36**, 65-77, doi:10.1016/j.tibs.2010.09.006 (2011).

35 Yadav, S. S. & Miller, W. T. The evolutionarily conserved arrangement of domains in SRC family kinases is important for substrate recognition. *Biochemistry* **47**, 10871-10880, doi:10.1021/bi800930e (2008).
36 Miller, W. T. Determinants of substrate recognition in nonreceptor tyrosine kinases. *Accounts of chemical research* **36**, 393-400, doi:10.1021/ar020116v (2003).
37 Harrison, S. C. Variation on an Src-like theme. *Cell* **112**, 737-740 (2003).
38 Hubbard, S. R., Mohammadi, M. & Schlessinger, J. Autoregulatory mechanisms in protein-tyrosine kinases. *The Journal of biological chemistry* **273**, 11987-11990 (1998).
39 Williams, J. C. *et al.* The 2.35 A crystal structure of the inactivated form of chicken Src: a dynamic molecule with multiple regulatory interactions. *J Mol Biol* **274**, 757-775, doi:10.1006/jmbi.1997.1426 (1997).
40 Filippakopoulos, P., Müller, S. & Knapp, S. SH2 domains: modulators of nonreceptor tyrosine kinase activity. *Current opinion in structural biology* **19**, 643-649, doi:10.1016/j.sbi.2009.10.001 (2009).
41 Levinson, N. M. *et al.* A Src-like inactive conformation in the abl tyrosine kinase domain. *PLoS biology* **4**, e144, doi:10.1371/journal.pbio.0040144 (2006).
42 Nagar, B. *et al.* Structural basis for the autoinhibition of c-Abl tyrosine kinase. *Cell* **112**, 859-871 (2003).
43 Xu, W., Doshi, A., Lei, M., Eck, M. J. & Harrison, S. C. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Molecular Cell* **3**, 629-638 (1999).
44 Miller, W. T. Tyrosine kinase signaling and the emergence of multicellularity. *Biochimica et biophysica acta* **1823**, 1053-1057, doi:10.1016/j.bbamcr.2012.03.009 (2012).
45 Hubbard, S. R. & Till, J. H. Protein tyrosine kinase structure and function. *Annu Rev Biochem* **69**, 373-398, doi:10.1146/annurev.biochem.69.1.373 (2000).
46 Mendrola, J. M., Shi, F., Park, J. H. & Lemmon, M. A. Receptor tyrosine kinases with intracellular pseudokinase domains. *Biochemical Society transactions* **41**, 1029-1036, doi:10.1042/BST20130104 (2013).
47 Chen, H. *et al.* A crystallographic snapshot of tyrosine trans-phosphorylation in action. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 19660-19665, doi:10.1073/pnas.0807752105 (2008).
48 Zhang, X., Gureasko, J., Shen, K., Cole, P. A. & Kuriyan, J. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* **125**, 1137-1149, doi:10.1016/j.cell.2006.05.013 (2006).
49 Artim, S. C., Mendrola, J. M. & Lemmon, M. A. Assessing the range of kinase autoinhibition mechanisms in the insulin receptor family. *The Biochemical journal* **448**, 213-220, doi:10.1042/BJ20121365 (2012).
50 Wei, L., Hubbard, S. R., Hendrickson, W. A. & Ellis, L. Expression, characterization, and crystallization of the catalytic core of the human insulin receptor protein-tyrosine kinase domain. *The Journal of biological chemistry* **270**, 8122-8130 (1995).
51 Hubbard, S. R., Wei, L., Ellis, L. & Hendrickson, W. A. Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* **372**, 746-754, doi:10.1038/372746a0 (1994).
52 Favelyukis, S., Till, J. H., Hubbard, S. R. & Miller, W. T. Structure and autoregulation of the insulin-like growth factor 1 receptor kinase - Nature Structural & Molecular Biology. *Nature structural biology* **8**, 1058-1063, doi:10.1038/nsb721 (2001).

53 Hubbard, S. R. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *The EMBO Journal* **16**, 5572-5581, doi:10.1093/emboj/16.18.5572 (1997).
54 Till, J. H. *et al.* Crystal Structure of the MuSK Tyrosine Kinase. *Structure (London, England : 1993)* **10**, 1187-1196, doi:10.1016/S0969-2126(02)00814-6 (2002).
55 Griffith, J. *et al.* The structural basis for autoinhibition of FLT3 by the juxtamembrane domain. *Mol Cell* **13**, 169-178 (2004).
56 Mol, C. D. *et al.* Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. *The Journal of biological chemistry* **279**, 31655-31663, doi:10.1074/jbc.M403319200 (2004).
57 Wybenga-Groot, L. E. *et al.* Structural basis for autoinhibition of the Ephb2 receptor tyrosine kinase by the unphosphorylated juxtamembrane region. *Cell* **106**, 745-757 (2001).
58 Shewchuk, L. M. *et al.* Structure of the Tie2 RTK domain: self-inhibition by the nucleotide binding loop, activation loop, and C-terminal tail. *Structure* **8**, 1105-1113 (2000).
59 Seet, B. T., Dikic, I., Zhou, M. M. & Pawson, T. Reading protein modifications with interaction domains. *Nature reviews. Molecular cell biology* **7**, 473-483, doi:10.1038/nrm1960 (2006).
60 Lemeer, S. *et al.* Phosphotyrosine mediated protein interactions of the discoidin domain receptor 1. *Journal of proteomics* **75**, 3465-3477, doi:10.1016/j.jprot.2011.10.007 (2012).
61 Furdui, C. M., Lew, E. D., Schlessinger, J. & Anderson, K. S. Autophosphorylation of FGFR1 kinase is mediated by a sequential and precisely ordered reaction. *Mol Cell* **21**, 711-717, doi:10.1016/j.molcel.2006.01.022 (2006).
62 Pawson, T. Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. *Cell* **116**, 191-203 (2004).
63 Schlessinger, J. & Lemmon, M. A. SH2 and PTB domains in tyrosine kinase signaling. *Science's STKE : signal transduction knowledge environment* **2003**, RE12, doi:10.1126/stke.2003.191.re12 (2003).
64 Hurley, J. H., Lee, S. & Prag, G. Ubiquitin-binding domains. *Biochem J* **399**, 361-372, doi:10.1042/BJ20061138 (2006).
65 Jin, J. *et al.* Eukaryotic protein domains as functional units of cellular evolution. *Science signaling* **2**, ra76, doi:10.1126/scisignal.2000546 (2009).
66 Songyang, Z. & Cantley, L. C. Recognition and specificity in protein tyrosine kinase-mediated signalling. *Trends in biochemical sciences* **20**, 470-475 (1995).
67 Filippakopoulos, P. *et al.* Structural coupling of SH2-kinase domains links Fes and Abl substrate recognition and kinase activation. *Cell* **134**, 793-803, doi:10.1016/j.cell.2008.07.047 (2008).
68 McPherson, V. A. *et al.* Contributions of F-BAR and SH2 domains of Fes protein tyrosine kinase for coupling to the FcepsilonRI pathway in mast cells. *Mol Cell Biol* **29**, 389-401, doi:10.1128/MCB.00904-08 (2009).
69 Nagar, B. *et al.* Organization of the SH3-SH2 unit in active and inactive forms of the c-Abl tyrosine kinase. *Molecular Cell* **21**, 787-798, doi:10.1016/j.molcel.2006.01.035 (2006).
70 Grebien, F. *et al.* Targeting the SH2-kinase interface in Bcr-Abl inhibits leukemogenesis. *Cell* **147**, 306-319, doi:10.1016/j.cell.2011.08.046 (2011).

71 Xu, W., Harrison, S. C. & Eck, M. J. Three-dimensional structure of the tyrosine kinase c-Src. *Nature* **385**, 595-602, doi:10.1038/385595a0 (1997).
72 Druker, B. J. *et al.* Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nat Med* **2**, 561-566 (1996).
73 Daley, G. Q., Van Etten, R. A. & Baltimore, D. Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. *Science* **247**, 824-830 (1990).
74 Brauer, P. M. & Tyner, A. L. Building a better understanding of the intracellular tyrosine kinase PTK6 - BRK by BRK. *Biochim Biophys Acta* **1806**, 66-73, doi:10.1016/j.bbcan.2010.02.003 (2010).
75 Yeatman, T. J. A renaissance for SRC. *Nature Reviews Cancer* **4**, 470-480, doi:10.1038/nrc1366 (2004).
76 Martin, G. S. The road to Src. *Oncogene* **23**, 7910-7917, doi:10.1038/sj.onc.1208077 (2004).
77 Takeya, T. & Hanafusa, H. DNA sequence of the viral and cellular src gene of chickens. II. Comparison of the src genes of two strains of avian sarcoma virus and of the cellular homolog. *Journal of virology* **44**, 12-18 (1982).
78 Thomas, S. M. & Brugge, J. S. Cellular functions regulated by Src family kinases. *Annual review of cell and developmental biology* **13**, 513-609, doi:10.1146/annurev.cellbio.13.1.513 (1997).
79 Sicheri, F. & Kuriyan, J. Structures of Src-family tyrosine kinases. *Current opinion in structural biology* **7**, 777-785 (1997).
80 Boggon, T. J. & Eck, M. J. Structure and regulation of Src family kinases. *Oncogene* **23**, 7918-7927, doi:10.1038/sj.onc.1208081 (2004).
81 Parsons, S. J. & Parsons, J. T. Src family kinases, key regulators of signal transduction. *Oncogene* **23**, 7906-7909, doi:10.1038/sj.onc.1208160 (2004).
82 Geahlen, R. L., Handley, M. D. & Harrison, M. L. Molecular interdiction of Src-family kinase signaling in hematopoietic cells. *Oncogene* **23**, 8024-8032, doi:10.1038/sj.onc.1208078 (2004).
83 Gauld, S. B. & Cambier, J. C. Src-family kinases in B-cell development and signaling. *Oncogene* **23**, 8001-8006, doi:10.1038/sj.onc.1208075 (2004).
84 Kalia, L. V., Gingrich, J. R. & Salter, M. W. Src in synaptic transmission and plasticity. *Oncogene* **23**, 8007-8016, doi:10.1038/sj.onc.1208158 (2004).
85 Manning, G., Young, S. L., Miller, W. T. & Zhai, Y. The protist, Monosiga brevicollis, has a tyrosine kinase signaling network more elaborate and diverse than found in any known metazoan. *Proc Natl Acad Sci U S A* **105**, 9674-9679, doi:10.1073/pnas.0801314105 (2008).
86 Li, W., Scarlata, S. & Miller, W. T. Evidence for convergent evolution in the signaling properties of a choanoflagellate tyrosine kinase. *Biochemistry* **48**, 5180-5186, doi:10.1021/bi9000672 (2009).
87 Shupnik, M. A. Crosstalk between steroid receptors and the c-Src-receptor tyrosine kinase pathways: implications for cell proliferation. *Oncogene* **23**, 7979-7989, doi:10.1038/sj.onc.1208076 (2004).
88 Bromann, P. A., Korkaya, H. & Courtneidge, S. A. The interplay between Src family kinases and receptor tyrosine kinases. *Oncogene* **23**, 7957-7968, doi:10.1038/sj.onc.1208079 (2004).

89 Playford, M. P. & Schaller, M. D. The interplay between Src and integrins in normal and tumor biology. *Oncogene* **23**, 7928-7946, doi:10.1038/sj.onc.1208080 (2004).
90 Zhang, S. & Yu, D. Targeting Src family kinases in anti-cancer therapies: turning promise into triumph. *Trends in pharmacological sciences*, doi:10.1016/j.tips.2011.11.002 (2011).
91 Gargalionis, A. N., Karamouzis, M. V. & Papavassiliou, A. G. The molecular rationale of Src inhibition in colorectal carcinomas. *International journal of cancer Journal international du cancer*, doi:10.1002/ijc.28299 (2013).
92 Tsai, P.-C., Chu, C.-L., Chiu, C.-C., Chang, L.-S. & Lin, S.-R. Inhibition of Src activation with cardiotoxin III blocks migration and invasion of MDA-MB-231 cells. *Toxicon : official journal of the International Society on Toxinology*, doi:10.1016/j.toxicon.2013.07.021 (2013).
93 Reynolds, A. B. & Roczniak-Ferguson, A. Emerging roles for p120-catenin in cell adhesion and cancer. *Oncogene* **23**, 7947-7956, doi:10.1038/sj.onc.1208161 (2004).
94 Pene-Dumitrescu, T. & Smithgall, T. E. Expression of a Src family kinase in chronic myelogenous leukemia cells induces resistance to imatinib in a kinase-dependent manner. *J Biol Chem* **285**, 21446-21457, doi:10.1074/jbc.M109.090043 (2010).
95 Zhang, S. *et al.* Combating trastuzumab resistance by targeting SRC, a common node downstream of multiple resistance pathways. *Nature Medicine* **17**, 461-469, doi:10.1038/nm.2309 (2011).
96 Li, C., Iida, M., Dunn, E. F., Ghia, A. J. & Wheeler, D. L. Nuclear EGFR contributes to acquired resistance to cetuximab. *Oncogene* **28**, 3801-3813, doi:10.1038/onc.2009.234 (2009).
97 van Oosterwijk, J. G. *et al.* Src kinases in chondrosarcoma chemoresistance and migration: dasatinib sensitises to doxorubicin in TP53 mutant cells. *British Journal of Cancer*, doi:10.1038/bjc.2013.451 (2013).
98 Vogel, W., Gish, G. D., Alves, F. & Pawson, T. The discoidin domain receptor tyrosine kinases are activated by collagen. *Molecular Cell* **1**, 13-23 (1997).
99 Shrivastava, A. *et al.* An orphan receptor tyrosine kinase family whose members serve as nonintegrin collagen receptors. *Molecular Cell* **1**, 25-34 (1997).
100 Vogel, W. Discoidin domain receptors: structural relations and functional implications. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **13 Suppl**, S77-82 (1999).
101 Playford, M. P. *et al.* The genomic structure of discoidin receptor tyrosine kinase. *Genome research* **6**, 620-627 (1996).
102 Vogel, W. F., Abdulhussein, R. & Ford, C. E. Sensing extracellular matrix: an update on discoidin domain receptor function. *Cellular signalling* **18**, 1108-1116, doi:10.1016/j.cellsig.2006.02.012 (2006).
103 Foehr, E. D. *et al.* Discoidin domain receptor 1 (DDR1) signaling in PC12 cells: activation of juxtamembrane domains in PDGFR/DDR/TrkA chimeric receptors. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **14**, 973-981 (2000).
104 Noordeen, N. A., Carafoli, F., Hohenester, E., Horton, M. A. & Leitinger, B. A transmembrane leucine zipper is required for activation of the dimeric receptor tyrosine kinase DDR1. *The Journal of biological chemistry* **281**, 22744-22751, doi:10.1074/jbc.M603233200 (2006).

105 Leitinger, B. Vol. 310 39-87 (Elsevier, 2014).
106 Heino, J., Huhtala, M., Käpylä, J. & Johnson, M. S. Evolution of collagen-based adhesion systems. *The international journal of biochemistry & cell biology* **41**, 341-348, doi:10.1016/j.biocel.2008.08.021 (2009).
107 Valiathan, R. R., Marco, M., Leitinger, B., Kleer, C. G. & Fridman, R. Discoidin domain receptor tyrosine kinases: new players in cancer progression. *Cancer metastasis reviews* **31**, 295-321, doi:10.1007/s10555-012-9346-z (2012).
108 Borza, C. M. & Pozzi, A. Discoidin domain receptors in disease. *Matrix biology : journal of the International Society for Matrix Biology*, doi:10.1016/j.matbio.2013.12.002 (2013).
109 Leitinger, B. Molecular analysis of collagen binding by the human discoidin domain receptors, DDR1 and DDR2. Identification of collagen binding sites in DDR2. *The Journal of biological chemistry* **278**, 16761-16769, doi:10.1074/jbc.M301370200 (2003).
110 Carafoli, F. *et al.* Structure of the Discoidin Domain Receptor 1 Extracellular Region Bound to an Inhibitory Fab Fragment Reveals Features Important for Signaling. *Structure (London, England : 1993)* **20**, 568-570, doi:10.1016/j.str.2012.03.003 (2012).
111 Carafoli, F. *et al.* Crystallographic insight into collagen recognition by discoidin domain receptor 2. *Structure (London, England : 1993)* **17**, 1573-1581, doi:10.1016/j.str.2009.10.012 (2009).
112 Kiedzierska, A., Smietana, K., Czepczynska, H. & Otlewski, J. Structural similarities and functional diversity of eukaryotic discoidin-like domains. *Biochimica et biophysica acta* **1774**, 1069-1078, doi:10.1016/j.bbapap.2007.07.007 (2007).
113 Carafoli, F. & Hohenester, E. Collagen recognition and transmembrane signalling by discoidin domain receptors. *Biochimica et biophysica acta*, doi:10.1016/j.bbapap.2012.10.014 (2012).
114 Yeung, D., Chmielewski, D., Mihai, C. & Agarwal, G. Oligomerization of DDR1 ECD affects receptor-ligand binding. *Journal of structural biology*, doi:10.1016/j.jsb.2013.06.010 (2013).
115 Phan, T. N. *et al.* Low stability and a conserved N-glycosylation site are associated with regulation of the discoidin domain receptor family by glucose via post-translational N-glycosylation. *Bioscience, biotechnology, and biochemistry* **77**, 1907-1916 (2013).
116 Curat, C. A., Eck, M., Dervillez, X. & Vogel, W. F. Mapping of epitopes in discoidin domain receptor 1 critical for collagen binding. *The Journal of biological chemistry* **276**, 45952-45958, doi:10.1074/jbc.M104360200 (2001).
117 Fu, H.-L. *et al.* Discoidin Domain Receptors: Unique Receptor Tyrosine Kinases in Collagen-mediated Signaling. *The Journal of biological chemistry* **288**, 7430-7437, doi:10.1074/jbc.R112.444158 (2013).
118 Xu, H. *et al.* Normal activation of discoidin domain receptor 1 mutants with disulphide cross-links, insertions or deletions in the extracellular juxtamembrane region: mechanistic implications. *jbc.org* (2014).
119 Fu, H.-L. *et al.* Shedding of discoidin domain receptor (DDR) 1 by membrane type (MT)-matrix metalloproteinases (MMPs). *The Journal of biological chemistry*, doi:10.1074/jbc.M112.409599 (2013).
120 Slack, B. E., Siniaia, M. S. & Blusztajn, J. K. Collagen type I selectively activates ectodomain shedding of the discoidin domain receptor 1: involvement of Src tyrosine kinase. *Journal of cellular biochemistry* **98**, 672-684, doi:10.1002/jcb.20812 (2006).

121 Wang, C.-Z., Su, H.-W., Hsu, Y.-C., Shen, M.-R. & Tang, M.-J. A discoidin domain receptor 1/SHP-2 signaling complex inhibits alpha2beta1-integrin-mediated signal transducers and activators of transcription 1/3 activation and cell migration. *Molecular biology of the cell* **17**, 2839-2852, doi:10.1091/mbc.E05-11-1068 (2006).
122 Iwai, L. K. *et al.* Phosphoproteomics of collagen receptor networks reveals SHP-2 phosphorylation downstream of wild-type DDR2 and its lung cancer mutants. *The Biochemical journal* **454**, 501-513, doi:10.1042/BJ20121750 (2013).
123 Day, E. *et al.* Inhibition of collagen-induced discoidin domain receptor 1 and 2 activation by imatinib, nilotinib and dasatinib. *European journal of pharmacology* **599**, 44-53, doi:10.1016/j.ejphar.2008.10.014 (2008).
124 Canning, P. *et al.* Structural Mechanisms Determining Inhibition of the Collagen Receptor DDR1 by Selective and Multi-Targeted Type II Kinase Inhibitors. *Journal of molecular biology*, doi:10.1016/j.jmb.2014.04.014 (2014).
125 Koo, D. H. H. *et al.* Pinpointing phosphotyrosine-dependent interactions downstream of the collagen receptor DDR1. *FEBS letters* **580**, 15-22, doi:10.1016/j.febslet.2005.11.035 (2006).
126 Castro-Sanchez, L., Soto-Guzman, A., Guaderrama-Diaz, M., Cortes-Reynosa, P. & Salazar, E. P. Role of DDR1 in the gelatinases secretion induced by native type IV collagen in MDA-MB-231 breast cancer cells. *Clinical & experimental metastasis* **28**, 463-477, doi:10.1007/s10585-011-9385-9 (2011).
127 Xu, H. *et al.* Discoidin domain receptors promote α1β1- and α2β1-integrin mediated cell adhesion to collagen by enhancing integrin activation. *PloS one* **7**, e52209, doi:10.1371/journal.pone.0052209 (2012).
128 Huang, Y., Arora, P., McCulloch, C. A. & Vogel, W. F. The collagen receptor DDR1 regulates cell spreading and motility by associating with myosin IIA. *Journal of cell science* **122**, 1637-1646, doi:10.1242/jcs.046219 (2009).
129 Lund, A. W., Stegemann, J. P. & Plopper, G. E. Mesenchymal Stem Cells Sense Three Dimensional Type I Collagen through Discoidin Domain Receptor 1. *The open stem cell journal* **1**, 40-53, doi:10.2174/1876893800901010040 (2009).
130 Rubel, D. *et al.* Collagen receptors integrin alpha2beta1 and discoidin domain receptor 1 regulate maturation of the glomerular basement membrane and loss of integrin alpha2beta1 delays kidney fibrosis in COL4A3 knockout mice. *Matrix biology : journal of the International Society for Matrix Biology*, doi:10.1016/j.matbio.2014.01.006 (2014).
131 Dejmek, J., Dib, K., Jönsson, M. & Andersson, T. Wnt-5a and G-protein signaling are required for collagen-induced DDR1 receptor activation and normal mammary cell adhesion. *International journal of cancer Journal international du cancer* **103**, 344-351, doi:10.1002/ijc.10752 (2003).
132 Neuhaus, B. *et al.* Migration inhibition of mammary epithelial cells by Syk is blocked in the presence of DDR1 receptors. *Cellular and molecular life sciences : CMLS* **68**, 3757-3770, doi:10.1007/s00018-011-0676-8 (2011).
133 Iwai, L. K., Chang, F. & Huang, P. H. Phosphoproteomic analysis identifies insulin enhancement of discoidin domain receptor 2 phosphorylation. *Cell adhesion & migration* **7**, 161-164, doi:10.4161/cam.22572 (2013).
134 Vogel, W. F., Aszódi, A., Alves, F. & Pawson, T. Discoidin domain receptor 1 tyrosine kinase has an essential role in mammary gland development. *Molecular and cellular biology* **21**, 2906-2917, doi:10.1128/MCB.21.8.2906-2917.2001 (2001).

135 Gross, O. *et al.* DDR1-deficient mice show localized subepithelial GBM thickening with focal loss of slit diaphragms and proteinuria. *Kidney international* **66**, 102-111, doi:10.1111/j.1523-1755.2004.00712.x (2004).
136 Al-Kindi, A. *et al.* A novel mutation in DDR2 causing spondylo-meta-epiphyseal dysplasia with short limbs and abnormal calcifications (SMED-SL) results in defective intra-cellular trafficking. *BMC medical genetics* **15**, 42, doi:10.1186/1471-2350-15-42 (2014).
137 Flynn, L. A., Blissett, A. R., Calomeni, E. P. & Agarwal, G. Inhibition of collagen fibrillogenesis by cells expressing soluble extracellular domains of DDR1 and DDR2. *Journal of molecular biology* **395**, 533-543, doi:10.1016/j.jmb.2009.10.073 (2010).
138 Ford, C. E. *et al.* Expression and mutation analysis of the discoidin domain receptors 1 and 2 in non-small cell lung carcinoma. *British Journal of Cancer* **96**, 808-814, doi:10.1038/sj.bjc.6603614 (2007).
139 Kerroch, M. *et al.* Genetic inhibition of discoidin domain receptor 1 protects mice against crescentic glomerulonephritis. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **26**, 4079-4091, doi:10.1096/fj.11-194902 (2012).
140 Guerrot, D. *et al.* Discoidin domain receptor 1 is a major mediator of inflammation and fibrosis in obstructive nephropathy. *The American journal of pathology* **179**, 83-91, doi:10.1016/j.ajpath.2011.03.023 (2011).
141 Gross, O., Girgert, R., Beirowski, B., Kretzler, M. & Kang, H. G. Loss of collagen-receptor DDR1 delays renal fibrosis in hereditary type IV collagen disease. *Matrix Biology* (2010).
142 Ahmad, P. J. *et al.* Discoidin domain receptor-1 deficiency attenuates atherosclerotic calcification and smooth muscle cell-mediated mineralization. *The American journal of pathology* **175**, 2686-2696, doi:10.2353/ajpath.2009.080734 (2009).
143 Franco, C. *et al.* Discoidin domain receptor 1 (ddr1) deletion decreases atherosclerosis by accelerating matrix accumulation and reducing inflammation in low-density lipoprotein receptor-deficient mice. *Circulation research* **102**, 1202-1211, doi:10.1161/CIRCRESAHA.107.170662 (2008).
144 Ren, T., Zhang, J., Zhang, J., Liu, X. & Yao, L. Increased expression of discoidin domain receptor 2 (DDR2): a novel independent prognostic marker of worse outcome in breast cancer patients. *Medical oncology (Northwood, London, England)* **30**, 397, doi:10.1007/s12032-012-0397-3 (2013).
145 Karaman, M. W. *et al.* A quantitative analysis of kinase inhibitor selectivity. *Nat Biotechnol* **26**, 127-132, doi:10.1038/nbt1358 (2008).
146 Eckstein, N. *et al.* Clinical pharmacology of tyrosine kinase inhibitors becoming generic drugs: the regulatory perspective. *Journal of Experimental & Clinical Cancer Research* **33**, 15, doi:10.1200/JCO.2005.02.2194 (2014).
147 Fabbro, D., Cowan-Jacob, S. W., Möbitz, H. & Martiny-Baron, G. Targeting cancer with small-molecular-weight kinase inhibitors. *Methods in molecular biology (Clifton, NJ)* **795**, 1-34, doi:10.1007/978-1-61779-337-0_1 (2012).
148 Zhang, J., Yang, P. L. & Gray, N. S. Targeting cancer with small molecule kinase inhibitors. *Nature Reviews Cancer* **9**, 28-39, doi:10.1038/nrc2559 (2009).
149 Krause, D. S. & Van Etten, R. A. Tyrosine kinases as targets for cancer therapy. *The New England journal of medicine* **353**, 172-187, doi:10.1056/NEJMra044389 (2005).

150 Gazit, A., Yaish, P., Gilon, C. & Levitzki, A. Tyrphostins I: synthesis and biological activity of protein tyrosine kinase inhibitors. *Journal of medicinal chemistry* **32**, 2344-2352 (1989).
151 Levitzki, A. Protein tyrosine kinase inhibitors as novel therapeutic agents. *Pharmacology & therapeutics* **82**, 231-239 (1999).
152 Adrián, F. J. *et al.* Allosteric inhibitors of Bcr-abl-dependent cell proliferation. *Nature Chemical Biology* **2**, 95-102, doi:10.1038/nchembio760 (2006).
153 Pargellis, C. *et al.* Inhibition of p38 MAP kinase by utilizing a novel allosteric binding site. *Nature structural biology* **9**, 268-272, doi:10.1038/nsb770 (2002).
154 Ohren, J. F. *et al.* Structures of human MAP kinase kinase 1 (MEK1) and MEK2 describe novel noncompetitive kinase inhibition. *Nature structural & molecular biology* **11**, 1192-1197, doi:10.1038/nsmb859 (2004).
155 Burgess, M. R., Skaggs, B. J., Shah, N. P., Lee, F. Y. & Sawyers, C. L. Comparative analysis of two clinically active BCR-ABL kinase inhibitors reveals the role of conformation-specific binding in resistance. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 3395-3400, doi:10.1073/pnas.0409770102 (2005).
156 Schindler, T. *et al.* Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science (New York, NY)* **289**, 1938-1942 (2000).
157 Liu, Y. & Gray, N. S. Rational design of inhibitors that bind to inactive kinase conformations. *Nature Chemical Biology* **2**, 358-364, doi:10.1038/nchembio799 (2006).
158 Capdeville, R., Buchdunger, E., Zimmermann, J. & Matter, A. Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug. *Nature Reviews Drug Discovery* **1**, 493-502, doi:10.1038/nrd839 (2002).
159 Nagar, B. c-Abl tyrosine kinase and inhibition by the cancer drug imatinib (Gleevec/STI-571). *The Journal of nutrition* **137**, 1518S-1523S; discussion 1548S (2007).
160 Druker, B. J. *et al.* Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *The New England journal of medicine* **355**, 2408-2417, doi:10.1056/NEJMoa062867 (2006).
161 Rix, U. *et al.* Chemical proteomic profiles of the BCR-ABL inhibitors imatinib, nilotinib, and dasatinib reveal novel kinase and nonkinase targets. *Blood* **110**, 4055-4063, doi:10.1182/blood-2007-07-102061 (2007).
162 Lee, S. J. & Wang, J. Y. J. Exploiting the promiscuity of imatinib. *Journal of biology* **8**, 30, doi:10.1186/jbiol134 (2009).
163 Karaman, M. W. *et al.* A quantitative analysis of kinase inhibitor selectivity. *Nature Biotechnology* **26**, 127-132, doi:10.1038/nbt1358 (2008).
164 Balzano, D., Santaguida, S., Musacchio, A. & Villa, F. A General Framework for Inhibitor Resistance in Protein Kinases. *Chemistry & Biology* **18**, 966-975, doi:10.1016/j.chembiol.2011.04.013 (2011).
165 Krishnamurty, R. & Maly, D. J. Biochemical mechanisms of resistance to small-molecule protein kinase inhibitors. *ACS chemical biology* **5**, 121-138, doi:10.1021/cb9002656 (2010).
166 Shah, N. P. *et al.* Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell* **2**, 117-125 (2002).

167 von Bubnoff, N., Schneller, F., Peschel, C. & Duyster, J. BCR-ABL gene mutations in relation to clinical resistance of Philadelphia-chromosome-positive leukaemia to STI571: a prospective study. *Lancet* **359**, 487-491, doi:10.1016/S0140-6736(02)07679-1 (2002).
168 Azam, M., Latek, R. R. & Daley, G. Q. Mechanisms of autoinhibition and STI-571/imatinib resistance revealed by mutagenesis of BCR-ABL. *Cell* **112**, 831-843 (2003).
169 Shah, N. P. *et al.* Overriding Imatinib Resistance with a Novel ABL Kinase Inhibitor. *Science* **305**, 399-401, doi:10.1126/science.1099480 (2004).
170 Weisberg, E. *et al.* Characterization of AMN107, a selective inhibitor of native and mutant Bcr-Abl. *Cancer Cell* **7**, 129-141, doi:10.1016/j.ccr.2005.01.007 (2005).
171 Manley, P. W. *et al.* Extended kinase profile and properties of the protein kinase inhibitor nilotinib. *Biochimica et biophysica acta* **1804**, 445-453, doi:10.1016/j.bbapap.2009.11.008 (2010).
172 Kleiner, R. E., Dumelin, C. E., Tiu, G. C., Sakurai, K. & Liu, D. R. In vitro selection of a DNA-templated small-molecule library reveals a class of macrocyclic kinase inhibitors. *Journal of the American Chemical Society* **132**, 11779-11791, doi:10.1021/ja104903x (2010).
173 Brandvold, K. R., Steffey, M. E., Fox, C. C. & Soellner, M. B. Development of a Highly Selective c-Src Kinase Inhibitor. *ACS chemical biology* **7**, 1393-1398, doi:10.1021/cb300172e (2012).
174 Seeliger, M. A. *et al.* c-Src binds to the cancer drug imatinib with an inactive Abl/c-Kit conformation and a distributed thermodynamic penalty. *Structure/Folding and Design* **15**, 299-311, doi:10.1016/j.str.2007.01.015 (2007).
175 Seeliger, M. A. *et al.* High yield bacterial expression of active c-Abl and c-Src tyrosine kinases. *Protein science : a publication of the Protein Society* **14**, 3135-3139, doi:10.1110/ps.051750905 (2005).
176 Seeliger, M. A. *et al.* c-Src binds to the cancer drug imatinib with an inactive Abl/c-Kit conformation and a distributed thermodynamic penalty. *Structure* **15**, 299-311, doi:10.1016/j.str.2007.01.015 (2007).
177 Otwinowski, Z. & Minor, W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* **276**, 307-326 (1997).
178 Leslie, A. G. W. Recent changes to the MOSFLM package for processing film and image plate data *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* (1992).
179 Vonrhein, C. *et al.* Data processing and analysis with the autoPROC toolbox. *Acta crystallographica Section D, Biological crystallography* **67**, 293-302, doi:10.1107/S0907444911007773 (2011).
180 Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58**, 1948-1954 (2002).
181 McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* **61**, 458-464, doi:10.1107/S0907444905001617 (2005).
182 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132, doi:10.1107/S0907444904019158 (2004).
183 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica Section D, Biological crystallography* **66**, 12-21, doi:10.1107/S0907444909042073 (2010).

184 Barker, S. C. *et al.* Characterization of pp60c-src tyrosine kinase activities using a continuous assay: autoactivation of the enzyme is an intermolecular autophosphorylation process. *Biochemistry* **34**, 14843-14851 (1995).
185 Songyang, Z. & Cantley, L. C. Recognition and specificity in protein tyrosine kinase-mediated signalling. *Trends Biochem Sci* **20**, 470-475 (1995).
186 Songyang, Z. *et al.* Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature* **373**, 536-539, doi:10.1038/373536a0 (1995).
187 Casnellie, J. E. Assay of protein kinases using peptides with basic residues for phosphocellulose binding. *Methods in enzymology* **200**, 115-120 (1991).
188 Abbonante, V. *et al.* Discoidin Domain Receptor 1 is a novel modulator of megakaryocyte-collagen interactions. *The Journal of biological chemistry*, doi:10.1074/jbc.M112.431528 (2013).
189 Cohen, P. Protein kinases--the major drug targets of the twenty-first century? *Nat Rev Drug Discov* **1**, 309-315, doi:10.1038/nrd773 (2002).
190 Cools, J. *et al.* PKC412 overcomes resistance to imatinib in a murine model of FIP1L1-PDGFR[alpha]-induced myeloproliferative disease. *Cancer Cell* **3**, 459-469 (2003).
191 Bikker, J. A., Brooijmans, N., Wissner, A. & Mansour, T. S. Kinase domain mutations in cancer: implications for small molecule drug design strategies. *J Med Chem* **52**, 1493-1509, doi:10.1021/jm8010542 (2009).
192 Zhang, J., Yang, P. L. & Gray, N. S. Targeting cancer with small molecule kinase inhibitors. *Nature Reviews Cancer* **9**, 28-39, doi:10.1038/nrc2559 (2009).
193 Levitzki, A. Protein tyrosine kinase inhibitors as novel therapeutic agents. *Pharmacol Ther* **82**, 231-239 (1999).
194 Gazit, A., Yaish, P., Gilon, C. & Levitzki, A. Tyrphostins I: synthesis and biological activity of protein tyrosine kinase inhibitors. *J Med Chem* **32**, 2344-2352 (1989).
195 Krishnamurty, R. & Maly, D. J. Biochemical mechanisms of resistance to small-molecule protein kinase inhibitors. *ACS Chem Biol* **5**, 121-138, doi:10.1021/cb9002656 (2010).
196 Soverini, S. *et al.* Contribution of ABL Kinase Domain Mutations to Imatinib Resistance in Different Subsets of Philadelphia-Positive Patients: By the GIMEMA Working Party on Chronic Myeloid Leukemia. *Clinical Cancer Research* **12**, 7374-7379, doi:10.1158/1078-0432.CCR-06-1516 (2006).
197 Azam, M., Seeliger, M. A., Gray, N. S., Kuriyan, J. & Daley, G. Q. Activation of tyrosine kinases by mutation of the gatekeeper threonine. *Nat Struct Mol Biol* **15**, 1109-1118, doi:10.1038/nsmb.1486 (2008).
198 Gartner, Z. J. & Liu, D. R. The generality of DNA-templated synthesis as a basis for evolving non-natural small molecules. *J Am Chem Soc* **123**, 6961-6963 (2001).
199 Gartner, Z. J. *et al.* DNA-templated organic synthesis and selection of a library of macrocycles. *Science* **305**, 1601-1605, doi:10.1126/science.1102629 (2004).
200 Li, X. & Liu, D. R. DNA-templated organic synthesis: nature's strategy for controlling chemical reactivity applied to synthetic molecules. *Angewandte Chemie (International ed in English)* **43**, 4848-4870, doi:10.1002/anie.200400656 (2004).
201 Doyon, J. B., Snyder, T. M. & Liu, D. R. Highly sensitive in vitro selections for DNA-linked synthetic small molecules with protein binding affinity and specificity. *J Am Chem Soc* **125**, 12372-12373, doi:10.1021/ja036065u (2003).
202 Kleiner, R. E., Dumelin, C. E. & Liu, D. R. Small-molecule discovery from DNA-encoded chemical libraries. *Chemical Society reviews*, doi:10.1039/c1cs15076f (2011).

203 Tse, B. N., Snyder, T. M., Shen, Y. & Liu, D. R. Translation of DNA into a library of 13,000 synthetic small-molecule macrocycles suitable for in vitro selection. *J Am Chem Soc* **130**, 15611-15626, doi:10.1021/ja805649f (2008).
204 Knight, Z. A. & Shokat, K. M. Features of selective kinase inhibitors. *Chem Biol* **12**, 621-637, doi:10.1016/j.chembiol.2005.04.011 (2005).
205 Tse, B. N., Snyder, T. M., Shen, Y. & Liu, D. R. Translation of DNA into a library of 13,000 synthetic small-molecule macrocycles suitable for in vitro selection. *Journal of the American Chemical Society* **130**, 15611-15626, doi:10.1021/ja805649f (2008).
206 Biron, E., Chatterjee, J. & Kessler, H. Optimized selective N-methylation of peptides on solid support. *J Pept Sci* **12**, 213-219, doi:10.1002/psc.711 (2006).
207 Abbott, L. *et al.* Discovery of thienopyridines as Src-family selective Lck inhibitors. *Bioorg Med Chem Lett* **17**, 1167-1171, doi:10.1016/j.bmcl.2006.12.035 (2007).
208 Burchat, A. F. *et al.* Pyrazolo[3,4-d]pyrimidines containing an extended 3-substituent as potent inhibitors of Lck -- a selectivity insight. *Bioorg Med Chem Lett* **12**, 1687-1690 (2002).
209 Maly, D. J., Choong, I. C. & Ellman, J. A. Combinatorial target-guided ligand assembly: identification of potent subtype-selective c-Src inhibitors. *Proc Natl Acad Sci U S A* **97**, 2419-2424, doi:97/6/2419 [pii] (2000).
210 Bamborough, P., Drewry, D., Harper, G., Smith, G. K. & Schneider, K. Assessment of chemical coverage of kinome space and its implications for kinase drug discovery. *J Med Chem* **51**, 7898-7914, doi:10.1021/jm8011036 (2008).
211 Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta crystallographica Section D, Biological crystallography* **58**, 1948-1954 (2002).
212 Cowan-Jacob, S. W. *et al.* The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure* **13**, 861-871, doi:10.1016/j.str.2005.03.012 (2005).
213 Xu, W., Doshi, A., Lei, M., Eck, M. J. & Harrison, S. C. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol Cell* **3**, 629-638 (1999).
214 Getlik, M. *et al.* Hybrid compound design to overcome the gatekeeper T338M mutation in cSrc. *J Med Chem* **52**, 3915-3926, doi:10.1021/jm9002928 (2009).
215 Kabsch, W. XDS. *Acta crystallographica Section D, Biological crystallography* **66**, 125-132, doi:10.1107/S0907444909047337 (2010).
216 Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta crystallographica Section D, Biological crystallography* **69**, 1204-1214, doi:10.1107/S0907444913000061 (2013).
217 Georghiou, G., Kleiner, R. E., Pulkoski-Gross, M., Liu, D. R. & Seeliger, M. A. Highly specific, bisubstrate-competitive Src inhibitors from DNA-templated macrocycles. *Nature Chemical Biology*, doi:10.1038/nchembio.792 (2012).
218 McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta crystallographica Section D, Biological crystallography* **61**, 458-464, doi:10.1107/S0907444905001617 (2005).
219 Karplus, P. A. & Diederichs, K. Linking Crystallographic Model and Data Quality. *Science (New York, NY)* **336**, 1030-1033, doi:10.1126/science.1218231 (2012).
220 Sawaya, M. in *Structural Genomics* Vol. 1091 *Methods in Molecular Biology* (ed Yu Wai Chen) Ch. 15, 205-214 (Humana Press, 2014).

221 Chatterjee, J., Gilon, C., Hoffman, A. & Kessler, H. N-methylation of peptides: a new perspective in medicinal chemistry. *Acc Chem Res* **41**, 1331-1342, doi:10.1021/ar8000603 (2008).
222 Gregoret, L. M., Rader, S. D., Fletterick, R. J. & Cohen, F. E. Hydrogen bonds involving sulfur atoms in proteins. *Proteins* **9**, 99-107, doi:10.1002/prot.340090204 (1991).
223 Levinson, N. M. *et al.* A Src-like inactive conformation in the abl tyrosine kinase domain. *PLoS Biol* **4**, e144, doi:10.1371/journal.pbio.0040144 (2006).
224 Hubbard, S. R. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *EMBO J* **16**, 5572-5581, doi:10.1093/emboj/16.18.5572 (1997).
225 Bikker, J. A., Brooijmans, N., Wissner, A. & Mansour, T. S. Kinase domain mutations in cancer: implications for small molecule drug design strategies. *Journal of medicinal chemistry* **52**, 1493-1509, doi:10.1021/jm8010542 (2009).
226 Corbin, A. S., La Rosée, P., Stoffregen, E. P., Druker, B. J. & Deininger, M. W. Several Bcr-Abl kinase domain mutants associated with imatinib mesylate resistance remain sensitive to imatinib. *Blood* **101**, 4611-4614, doi:10.1182/blood-2002-12-3659 (2003).
227 Griswold, I. J. *et al.* Kinase domain mutants of Bcr-Abl exhibit altered transformation potency, kinase activity, and substrate utilization, irrespective of sensitivity to imatinib. *Molecular and cellular biology* **26**, 6082-6093, doi:10.1128/MCB.02202-05 (2006).
228 Young, M. A. *et al.* Structure of the kinase domain of an imatinib-resistant Abl mutant in complex with the Aurora kinase inhibitor VX-680. *Cancer Research* **66**, 1007-1014, doi:10.1158/0008-5472.CAN-05-2788 (2006).
229 Sicheri, F., Moarefi, I. & Kuriyan, J. Crystal structure of the Src family tyrosine kinase Hck. *Nature* **385**, 602-609, doi:10.1038/385602a0 (1997).
230 Wilhelm, S. *et al.* Discovery and development of sorafenib: a multikinase inhibitor for treating cancer. *Nature Reviews Drug Discovery* **5**, 835-844, doi:10.1038/nrd2130 (2006).
231 Quintás-Cardama, A., Kantarjian, H. & Cortes, J. Flying under the radar: the new wave of BCR-ABL inhibitors. *Nature Reviews Drug Discovery* **6**, 834-848, doi:10.1038/nrd2324 (2007).
232 Bamborough, P., Drewry, D., Harper, G., Smith, G. K. & Schneider, K. Assessment of chemical coverage of kinome space and its implications for kinase drug discovery. *Journal of medicinal chemistry* **51**, 7898-7914, doi:10.1021/jm8011036 (2008).
233 Verkhivker, G. M. Exploring sequence-structure relationships in the tyrosine kinome space: functional classification of the binding specificity mechanisms for cancer therapeutics. *Bioinformatics (Oxford, England)* **23**, 1919-1926, doi:10.1093/bioinformatics/btm277 (2007).
234 Yang, X. *et al.* Kinase inhibition-related adverse events predicted from in vitro kinome and clinical trial data. *Journal of biomedical informatics* **43**, 376-384, doi:10.1016/j.jbi.2010.04.006 (2010).
235 Nagar, B. *et al.* Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571). *Cancer Research* **62**, 4236-4243 (2002).
236 Kim, H.-G. *et al.* Discovery of a Potent and Selective DDR1 Receptor Tyrosine Kinase Inhibitor. *ACS chemical biology*, 130813082331005, doi:10.1021/cb400430t (2013).

237 Harrington, E. A. *et al.* VX-680, a potent and selective small-molecule inhibitor of the Aurora kinases, suppresses tumor growth in vivo. *Nat Med* **10**, 262-267, doi:10.1038/nm1003 (2004).
238 Zhao, B. *et al.* Modulation of kinase-inhibitor interactions by auxiliary protein binding: crystallography studies on Aurora A interactions with VX-680 and with TPX2. *Protein Sci* **17**, 1791-1797, doi:10.1110/ps.036590.108 (2008).
239 Salah, E. *et al.* Crystal structures of ABL-related gene (ABL2) in complex with imatinib, tozasertib (VX-680), and a type I inhibitor of the triazole carbothioamide class. *J Med Chem* **54**, 2359-2367, doi:10.1021/jm101506n (2011).
240 Elkins, J. M., Santaguida, S., Musacchio, A. & Knapp, S. Crystal structure of human aurora B in complex with INCENP and VX-680. *J Med Chem* **55**, 7841-7848, doi:10.1021/jm3008954 (2012).
241 O'Hare, T. *et al.* AP24534, a Pan-BCR-ABL Inhibitor for Chronic Myeloid Leukemia, Potently Inhibits the T315I Mutant and Overcomes Mutation-Based Resistance. *Cancer Cell* **16**, 401-412, doi:10.1016/j.ccr.2009.09.028 (2009).
242 Azam, M., Raz, T., Nardi, V., Opitz, S. L. & Daley, G. Q. A screen to identify drug resistant variants to target-directed anti-cancer agents. *Biological procedures online* **5**, 204-210, doi:10.1251/bpo63 (2003).