Global Methods for Discovering and Characterizing Regulatory Mechanisms of Multi-Domain Protein Kinases

Zachary E. Potter

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Dustin J. Maly, Chair

Michael Gelb

Champak Chatterjee

Program Authorized to Offer Degree:

Chemistry

University of Washington

Abstract

Global Methods for Discovering and Characterizing Regulatory Mechanisms of Multi-Domain Protein Kinases

Zachary E. Potter

Chair of the Supervisory Committee:

Dustin J. Maly

Department of Chemistry

Perturbations to cellular phosphorylation levels are highly correlated with a variety of disease states. Because protein kinases are the enzymes responsible for protein phosphorylation, they play a central role in maintaining homeostatic phosphorylation levels, and as such have become attractive drug targets. Consequently, the regulatory mechanisms that govern protein kinase activity have been studied for decades. Roughly half of protein kinases have at least one protein domain in addition to their catalytic kinase domain and in many cases these domains serve as "regulatory domains" by making physical contacts with surfaces on the catalytic domain, disrupting the alignment of

catalytically necessary residues. While the intramolecular regulatory mechanisms of many kinases have been delineated, there are many layers of regulation that lack definition. Specifically, a collaborative effort between the Maly and Fowler labs (University of Washington Department of Genome Sciences) revealed new putative regulatory surfaces on the catalytic domain of the long-studied Src kinase. One central hypothesis of this work is that there are similar but distinct regulatory surfaces on other members of Src Family of Kinases (SFKs), which give rise to differences in kinase substrate specificity, localization, temporal control over kinase activity, and overall mechanisms of regulation. Given the involvement of the SFKs Lck and Fyn in T-cell development and mature thymocyte signaling, we would like to better understand how these regulatory surfaces contribute to productive T cell receptor (TCR) signaling, which has yet to be systematically explored. Therefore, we identified putative inter- and intramolecular regulatory surfaces on Lck—the most centrally involved SFK in TCR signaling—using a series of saturation mutagenesis Deep Mutational Scans (DMS) in yeast. In addition to revealing fundamental information about the roles of Lck in mediating healthy TCR signaling, the methods described herein are general, and can be applied to study any protein of interest.

TABLE OF CONTENTS

LIST OF FIGURES

Introduction:

There are more than 500 protein kinases encoded in the human genome all of which contain a highly conserved bi-lobal kinase domain with an ATP-binding cleft in between them (NCBI Resource Coordinators, 2013; Manning, *et al.*, 2002) . Alignment of key residues within the active site is required for catalytic phosphotransferase activity (Engh, *et al.*, 2002; Akamine, *et al.* 2003; Kornev, *et al.*, 2006). These include the conserved DFG-motif which contains a catalytic aspartate that points into the ATP-binding site to coordinate Mg^{2+}. Additionally, a conserved glutamic acid, which resides on the helix αC, forms a salt bridge with a catalytic lysine that coordinates the α and β phosphates of ATP. Significant movement of either of these features disrupts kinase activity. Analysis of inhibitor bound crystal structures of many protein kinases reveals two distinct inhibitor-stabilized inactive ATP-binding site conformations (Chapman, *et al.*, 1986; Jeffrey, *et al.*, 1995; Huse, *et al.* 2002). In the "DFG-out" conformation, the DFG-motif flips nearly 180° and the aspartate points out of the ATP-binding site. In the "helix αC-out" conformation the helix αC rotates and moves revealing a pocket which is often occupied by inhibitors that stabilize this conformation. The SFKs contain five distinct domains: the kinase, SH2 (phosphotyrosine motif binding), SH3 (polyproline motif binding), unique and SH4 (membrane binding) domains. The intramolecular engagement or disengagement of one or more of these domains has been extensively studied. This conformational equilibrium is defined by two extremes. When all regulatory domains are disengaged, and the SFK is phosphorylated at its activation loop, it is most catalytically active and in the "open" global conformation. On the other hand, when the kinase is dephosphorylated at its activation loop, and phosphorylated at the C-terminal tail, the regulatory domains make physical contact with the kinase domain, perturbing the alignment of the key active site residues, abrogating kinase activity and the kinase is said to be in the "closed" global conformation.

Chapter 1: Parallel Chemoselective Profiling for Mapping Protein Structure

Introduction

Solution-based structural techniques complement high-resolution structural data by providing insight into the often-missing links between protein structure and function. Advances in mass spectrometry (MS) hardware and informatics software have made bottom-up proteomics-style workflows appealing for characterizing protein structure in solution. Bottom-up proteomic methods can be merged with so-called "protein footprinting" methods, which take advantage of a relatively simple concept, that the relative solvent accessibility of protein surfaces is a proxy for structure and dynamics. Two of the most widely used footprinting methods—hydrogen-deuterium exchange MS (HDX-MS) and fast photochemical oxidation of proteins—have been developed and employed successfully to map epitope-binding sites and characterize the conformational dynamics of proteins, among other applications. However, these methods rely on relatively specialized equipment and expertise, limiting their general implementation. I sought to develop a simple and general method that can report on the solvent accessibility of individual residues within a protein of interest. Recently, I published a general MS-based method called Parallel Chemoselective Profiling (Potter, et al. 2020) that does not rely on the use of specialized equipment beyond an LC system and a high-resolution tandem mass spectrometer. This method is particularly useful for mapping the structure and dynamics of proteins that are not well-suited for characterization using traditional structural techniques, including large highly dynamic multi-domain proteins. The method relies on two key components: a pool of protein variants each bearing a single amino acid mutation with a desired chemistry at a defined position and a Parallel Reaction Monitoring (PRM) assay for targeted mass spectrometric quantification of labeled mutant residues. The single amino acid protein variants are produced in a pooled format using a library of expression plasmids, each containing one cysteine mutant per plasmid.

The position of each mutation is guided by the "wild-type-like" mutations identified during a DMS of the protein of interest (**Figure 1**, *left*). Two samples of the pooled protein—for example, treated with two different ligands—are incubated with a chemoselective labeling agent that is employed to label solvent exposed residues (**Figure 1**, *middle*). Samples are taken and quenched at different time points. After tryptic digest, labeled residues are quantified with targeted MS using a PRM assay, which can detect attomole levels of labeled peptide using its fragmentation spectra (**Figure 1**, *right*). The differences in labeling kinetics of a given cysteine residue can be read out through the quantified signal intensity of each mutant containing peptide measured at several timepoints. Taken together, the labeling rates of all observed cysteine residues paint a picture of a protein's dynamic features and/or binding interfaces.

Structural characterization techniques are essential tools for studying protein function. Methods such as X-ray crystallography and cryogenic electron microscopy have become standard techniques for providing high-resolution protein structure. Despite the general utility of these methods, they only capture part of the picture and are best complemented by techniques that characterize dynamics in solution. The ability to assess protein dynamics in solution yields functional insight beyond static structures and provides information about regions of proteins that cannot be characterized using traditional structural methods. A subset of mass spectrometry-based structural characterization techniques—collectively referred to as protein footprinting methods—profile the solvent-exposed surfaces of proteins. These methods take advantage of a relatively simple concept, that the relative solvent accessibility of protein surfaces is a proxy for structure and dynamics. To measure accessibility and report on local residue environment, the solvent-exposed residues of a protein are labeled with a chemical modification that can be quantified with downstream analytical methods. Comparative analysis of protein variants or ligand-bound protein complexes can reveal the influence of such perturbations. Advances in mass spectrometer hardware and informatics software

have made bottom-up proteomic workflows appealing for detecting and quantifying modified proteins. Two methods that utilize a proteomics-style approach, hydrogen-deuterium exchange mass spectrometry (HDX-MS) and fast photochemical oxidation of proteins, have been developed and employed successfully to map ligand-binding sites and characterize the conformational dynamics of proteins, among other applications (Chea & Jones, 2018; Hodkinson, *et al.* 2009). These methods, however, rely on relatively specialized liquid chromatography (LC) equipment, which limits their general applicability (**Table S1**).

Here, we describe the development of a simple and general method that can report on the solvent accessibility of individual residues within a protein of interest (**Figure 1**). Integral to our method is the use of deep mutational scanning (DMS) data to inform the generation of multiple single amino acid variants of a protein of interest with desired side-chain chemistries at specific locations. Chemoselective labeling of a pool of DMS-guided protein variants under comparative conditions is leveraged to report on protein local environment and dynamics, and a parallel reaction monitoring (PRM) assay is used to provide residue-level quantification by liquid chromatography-tandem mass spectrometry (LC-MS/MS). Our "Parallel Chemoselective Profiling" method allows rapid characterization of dynamic protein structure under multiple conditions and does not require the use of highly specialized equipment beyond an LC and MS capable of acquiring high-resolution tandem mass spectra. To demonstrate the utility of parallel chemoselective profiling, we studied how ATP-binding site occupancy affects the local and global structural dynamics of the tyrosine kinase Src. Using our method, we performed an analysis of how stabilizing the ATP-binding site of full-length Src in two different inactive conformations with inhibitors affects its local and global structure and dynamics. In addition to observing expected changes in local and global conformation, our chemoselective profiling results revealed what appears to be an extended allosteric network within the kinase domain of full-length Src, which links the ATP-binding site to a region on the C-terminal lobe that is more

than 20 Å away. These observations corroborate results from HDX-MS experiments that we performed in parallel and elaborate on previously reported nuclear magnetic resonance studies (Tong et al., 2017). Together, our findings support the notion of an extended allosteric network within the kinase domain of full-length Src and highlight the utility of parallel chemoselective profiling for characterizing protein structure and dynamics.

Structural characterization techniques are essential tools for studying protein function. Methods such as x-ray crystallography and cryogenic electron microscopy have become standard techniques for providing atomic resolution into protein structure and function. Despite the general utility of these techniques, they only capture part of the picture and are best complemented by techniques that characterize dynamics in solution. The ability to rapidly assess protein dynamics in solution yields additional functional insight beyond static structures and provides information about regions of proteins that cannot be characterized using traditional structural techniques. A subset of mass spectrometry-based structural characterization techniques—collectively referred to as protein footprinting methods—profile the solvent exposed surfaces of proteins. These methods take advantage of a relatively simple concept, that the relative solvent accessibility of protein surfaces is a proxy for structure and dynamics. To measure accessibility and report on local residue environment, solvent exposed residues of a protein are labeled with a chemical modification that can be quantified with downstream analytical methods. Comparative analysis of protein variants or ligand-bound

protein complexes can reveal the influence of such perturbations. Footprinting experiments have been used to report changes in conformational dynamics of a protein and map

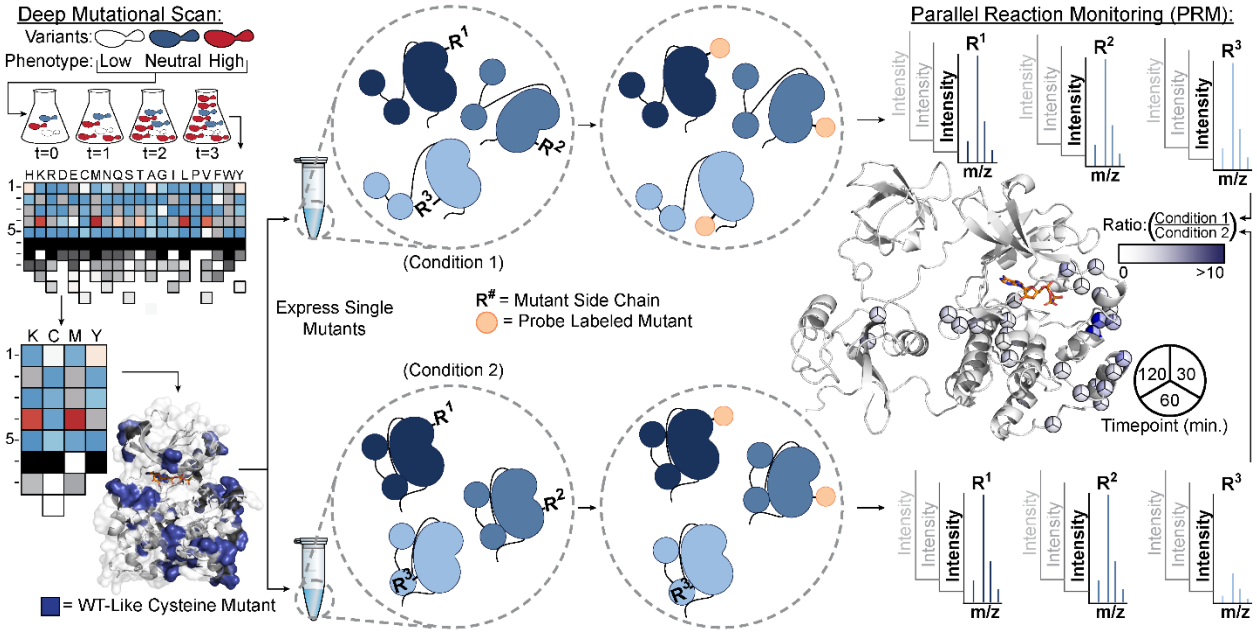


Figure 1. General Schematic for Parallel Chemoselective Profiling
Deep mutational Scanning (DMS) used to determine positions where chemically reactive reporter amino acids (e.g. cysteine) can be installed on a protein of interest without perturbing protein function (*left*). Individual mutants are expressed and purified in a pooled format and are treated under comparative conditions *in vitro*. Following treatment, solvent exposed mutant residues are labeled with a chemoselective reagent (e.g. isotopically-labeled iodoacetamide in the case of cysteine) (*middle*). Unreacted mutant residues are then capped (if necessary) under denaturing conditions, protein is digested using trypsin and labeled residues are quantified using Parallel Reaction Monitoring (PRM) Liquid Chromatography Mass Spectrometry (LC-MS/MS) (*right*). Ratios of labeled residues are compared across timepoints or labeling reagent conditions, which can be visualized and use to infer changes to the local environment of each mutant residue.

physical binding interfaces (Baerga-Ortiz, *et al.* 2002; Coales, *et al* 2009; Chea & Jones, 2018; Hodkinson, *et al.* 2009; Houde, *et al* 2009). Advances in mass spectrometer hardware and informatics software have made proteomics-style workflows appealing for detecting and quantifying modified proteins by way of bottom-up proteomics. Two examples of methods that utilize a proteomics-style approach, Hydrogen-Deuterium Exchange mass spectrometry (HDX-MS) and Fast Photochemical Oxidation of Proteins (FPOP) have been developed and employed successfully to map epitope

binding sites and characterize the conformational states of proteins, among other applications. These methods, however, rely on relatively specialized equipment which limits their general applicability.

We sought to develop a simple and general method that can report the solvent accessibility of individual residues within a protein of interest. To achieve this, Deep Mutational Scanning (DMS) data is leveraged to inform the generation of multiple single amino acid variants of a protein of interest with desired chemistries at specific locations (**Figure 1**, *left*). Following recombinant expression and purification of all DMS-guided variants, a pooled mutant library is subjected to comparative conditions *in vitro*. A chemoselective labeling reagent is then introduced which serves as a mass-encoded label for downstream detection and quantification by liquid chromatography tandem mass spectrometry (LC-MS/MS) (**Figure 1**, *middle*). Samples are quenched at various timepoints and previously unlabeled mutants are capped with a distinguishable mass-encoded label. Differences in the extent of mutant residue labeling compared between conditions provides information about changes in the local environment of a specific residue. While we initially attempted to use a less technically demanding shotgun proteomic workflow to quantify labeled peptides, we found this approach to be suboptimal. To accurately and rapidly detect labeled peptides we turned to a Parallel Reaction Monitoring (PRM) assay (**Figure 1**, *right*). Our "Parallel Chemoselective Profiling" method allows rapid characterization of dynamic protein structure under multiple conditions and does not require the use of highly specialized equipment beyond a MS capable of acquiring high resolution tandem mass spectra.

To demonstrate the utility of parallel chemoselective profiling, we studied how ATP-binding site occupancy affects the local and global structural dynamics of the tyrosine kinase Src. Using parallel chemoselective profiling, we performed the first analysis of how stabilizing the ATP-binding site of full-length Src in two different inactive conformations with inhibitors affects its local and global structure and dynamics. In addition to observing expected changes in local and global conformation, our chemoselective profiling results revealed what appears to be an extended allosteric network within

the kinase domain of full-length Src, which links the ATP-binding site to a region on the C-terminal lobe that is more than 20 Å away. These observations corroborate results from HDX-MS experiments that we performed in parallel and elaborate on previously reported NMR studies of how conformation-selective inhibitors affect the solution dynamics of Src's kinase domain (Tong, *et al.* 2017). Together, our findings support the notion of an extended allosteric network within the kinase domain of full-length Src and highlight the utility of parallel chemoselective profiling for characterizing protein structure and dynamics.

Results

DMS Data as a Resource for Identifying Mutants Appropriate for Chemoselective Profiling

Deep Mutational Scanning has emerged as a powerful tool to study the functional effects of thousands of single and multi-amino acid variants in a single experiment (Fowler & Fields, 2014). In contrast with traditional single amino acid-scanning methods, DMS experiments can be used to quantify the effects of all possible amino acid substitutions at each position in a protein. As this field of protein science grows, public repositories—including MaveDB—have been established to distribute data collected from multiplexed assays of variant effects (https://www.mavedb.org; Esposito *et al.,* 2019). We posit that these datasets provide a rich source of information which can be leveraged to identify sets of mutants that function similarly to the wild-type protein but have sidechain chemistries at predefined locations that can be used to report changes in local environment. Carefully selecting a sidechain chemistry and suitable labeling reagent will allow us to achieve our goal of performing protein footprinting experiments using chemoselective profiling.

To explore the potential utility of our proposed method, we analyzed the DMS data of four proteins—ubiquitin, PTEN, Src and BRCA1—that we felt are representative candidates for the

application of chemoselective profiling and for which high resolution structural data is available (**Table 1**). Ubiquitin, PTEN, Src and BRCA1 span a range of molecular weights and vary from single domain to multi-domain architectures. Importantly, these four proteins are representative of a range of biological functions and diverse parameters were used to evaluate the fitness of their variants for the generation of DMS datasets. Ubiquitin is a small, single domain protein that serves as a multi-functional post-translational modification. A yeast growth assay was used to evaluate how mutations affect ubiquitin structure and function. Phosphatase and tensin homolog (PTEN) is a multi-domain phosphoinositide phosphatase. Relative cellular abundance was the parameter used to assess the effects of PTEN mutations. The tyrosine kinase Src is a multi-domain protein that is subject to multiple layers of regulation. Src mutations were scored based on their relative effect on kinase activity. Breast cancer type 1 susceptibility protein (BRCA1) is a tumor suppressor and E3 ubiquitin-protein ligase. While the effects of mutations to the exons encoding the RING and BRCT domains were assessed in HAP1 cells, we only analyzed BCRT domain variants given the more extensive structural characterization of this region. All of the DMS experiments we analyzed were performed in different laboratories at different times, using different assay conditions and analytical techniques, mitigating the chances of a bias arising from a particular experimental design.

Table 1: Large-scale mutagenesis datasets used in this study.

Data set	Variant length	Mutagenized Residues*	Number of Mutations	Mutational completeness**	Selected phenotype	Citation
***Ubiquitin* Yeast growth**	76 amino acids	2-76	4,800	99%	Yeast growth	Roscoe, *et al.* 2013
PTEN	403 amino acids	14-351	4,112	54%	Protein abundance	Matreyek, *et al.* 2018
***Src* (kinase domain dataset)**	536 amino acids	270-519	3,506	73%	Kinase activity	Ahler, *et al.* 2019
BRCA1 (BRCT dataset)	1,863 amino acids	1,649-1,859	1,317	36%	Cell survival	Findlay, *et al.* 2018

*Residues represented in our analysis in **Figures 2** and **6**

**Proportion of observed mutations relative to all possible single amino acid mutations in mutagenized region in our analyses

Tolerability and Solvent Accessibility of Cys Substitutions

In the first iteration of our chemoselective profiling method, we analyzed the cysteine substitutions contained in the DMS datasets for ubiquitin, PTEN, Src and BRCA1 (**Table 1**). We chose cysteine due to the availability of many electrophilic reagents that can be used to selectively label this reactive residue (Abo, M, *et al.* 2018; Weerapana & Wang, *et al.* 2010). We extracted the mutational effect scores for the ubiquitin, PTEN, Src and BRCA1 datasets separately and binned the cysteine scores using the annotations for each single mutant provided in the original reports. Briefly, function scores are determined by comparing the scored phenotype for each variant relative to the distribution of scores for all synonymous wild-type variants. Therefore, while the specific statistical criteria used to define function scores for each variant were different for each protein and assay type, all function scores are derived from comparisons to the distribution of wild-type variants. This allows us to bin the variants of each protein using three categories for each cysteine mutation—wild type like (WT-like), not wild type like (Not WT-like) and those cysteine mutants that did not have scores reported in the DMS dataset (Not in DMS). We then mapped these bins to the crystal structures of each protein

(**Figure 2A-D**). We quantified the tolerance of cysteine mutations as the proportion of WT-like mutants to the total number of cysteine mutants contained in the dataset and found that in all cases, more than half of all cysteine mutants observed were WT-like (**Figure 2E,G,I,K**). We then used the PyMOL script FindSurfaceResidues to identify the solvent exposed sidechains in each crystal structure. In all cases, a majority of WT-like cysteine mutants are located at amino acid positions that have solvent exposed sidechains. Taken together, these analyses suggest that cysteine mutations are well tolerated at positions that are generally solvent accessible. Therefore, cysteine appears to be a promising choice for developing our chemoselective profiling methodology.

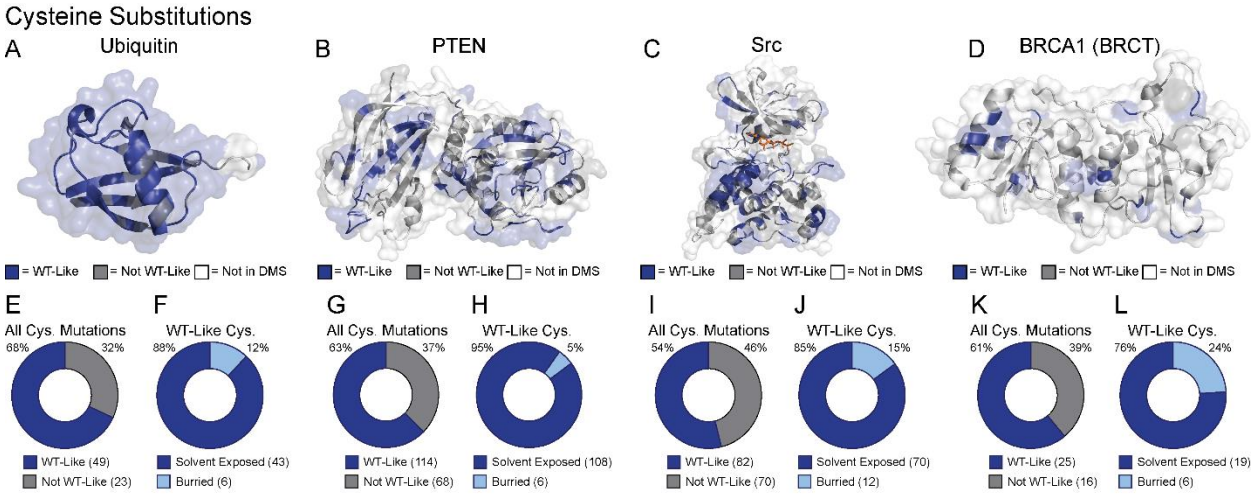


Figure 2. Tolerance and Solvent Accessibility of Cysteine Substitutions
A. Fitness scores for cysteine substitutions (Roscoe, B.P. et al 2013) mapped to the crystal structure of ubiquitin (PDBL 1UBQ). B. VAMP-seq scores for cysteine substitutions (Matreyek, K.A. et al., 2018) mapped to the crystal structure of PTEN (PDB: 1D5R). C. Activity scores for cysteine substitutions (Ahler, E. et al., 2019) mapped to the crystal structure of Src's kinase domain (PDB: 3DQW). D. Function scores for cysteine substitutions (Findlay, G.M., et al 2018) mapped to the crystal structure of BRCA1 BRCT domain (PDB: 1T29). E, G, I, K. Fractions of all cysteine mutants classified as WT-line or not WT-line in the original publications. F, H, J, L. Fraction of WT-line cysteine mutants whose endogenous amino acid contains a solvent-exposed sidechain. E-L. Each subpanel relates to the protein depicted immediately above.

Creation of an Src Cys Mutants Library for Chemoselective Profiling

To demonstrate the utility of our method, we chose to apply it to study the conformational dynamics of Src (**Figure 2C**). We chose Src as a test case for several reasons. Src has been studied for decades, which has helped define the roles of the SH2 and SH3 domains as allosteric regulators of Src's phosphotransferase activity (Boggon & Eck, 2004). Additionally, the rich Src literature includes well-characterized point mutants, truncation constructs and small molecule tools for perturbing Src conformational dynamics. We reasoned that we could employ parallel chemoselective profiling to study the conformational dynamics of full-length Src (Src FL) for the first time. From a DMS performed on the kinase domain of Src in the context of the full-length enzyme (**Figure 3A**), which contains 3,506 single amino acid variants, we extracted the mutational effect scores for all cysteine mutants (**Figure 3B**). Of these 152 cysteine mutants, 82 of them were classified as WT-like indicating that cysteine mutations at these sites do not significantly affect the catalytic activity of Src FL. From this information, we infer that these WT-like mutations do not disrupt the catalytic machinery or interdomain interactions of Src FL, and therefore we selected these mutants to move forward with for developing our parallel chemoselective profiling method.

We employed an inverse PCR strategy to generate individual plasmids for each WT-like cysteine mutant in mammalian expression vectors (**Figure S1A,B**). This approach yielded 74 of the 82 possible point mutants in the first pass without optimization. We verified the expression of each construct individually using western blot analysis of transiently transfected HEK293T cells (data not shown). DNA concentrations of the 74 mutant constructs were then normalized and plasmids were pooled and transiently transfected into suspension 293F cells. Resulting recombinant Src FL cysteine mutant proteins were dephosphorylated *in vitro* with the phosphatase YopH and purified as a pool of mutants—hereafter referred to as the Src Cysteine Library (Src^{CysLib}) (**Figure S1C, S2**).

Targeted MS Allows Quantification of Cys Mutant Peptides

To generate alkylated, cysteine-containing peptides for mass spectrometric analysis, we prepared samples of Src^{CysLib} under denaturing conditions with a concentration of iodoacetamide that should yield quantitative alkylation of all cysteines present. We first attempted to identify the labeled mutant peptides using a traditional shotgun proteomic workflow operating the MS in Data-Dependent Acquisition (DDA) mode, however this yielded few mutant peptide identifications (**Figure 3C**). Given the low number of cysteine mutants identified from the Src^{CysLib} it became apparent that our parallel chemoselective profiling method would not be amenable to a DDA workflow. We reasoned that in the ideal case, the Src^{CysLib} contains a pool of mutants expressed in equimolar ratios relative to one another. Therefore, each Src cysteine mutant peptide is diluted with 73 copies of the wild-type Src (Src*WT* peptide. Perfect chromatographic separation of all peptides is not achieved with the run length and consequent organic gradient needed to perform the number of analyses required for this method in a practical amount of LC-MS/MS run time. Therefore, multiple abundant Src *WT* peptide analytes, including all various non-cysteine-containing peptides, compete with the mutant cysteine-containing peptides for ionization and identification. This issue is exacerbated by the fact that we anticipate sub-stoichiometric alkylation regimes are required for surface profiling experiments. We wanted to accurately quantify differentially labeled cysteine peptides despite the challenges of low stoichiometry of labeling and the presence of co-eluting Src*WT* peptides. To address these challenges, we pursued a targeted proteomic workflow.

We implemented a Parallel Reaction Monitoring (PRM) assay to allow rapid quantification of our Src^{CysLib} peptides. We began by constructing a preliminary inclusion list using the FASTA sequence of the 74 individual Src mutants with variable modifications specified for oxidation of methionine and the products of iodoacetamide (light) and isotopically-labeled iodoacetamide (heavy) alkylation on cysteine (**Figure 3D**, *top left*). The list was refined to ~800 precursors by removing short tryptic

peptides (<6 amino acids), allowing for up to two missed tryptic cleavages and charge states of +2 through +5, resulting in a list of precursors for 67 mutants for each mass label (**Figure 3D**, *top right*). We prepared samples by separately alkylating cysteines under denaturing conditions with light or heavy iodoacetamide and acquired fragmentation spectra of these Src mutants using our targeted precursor inclusion lists. To implement a Parallel Reaction Monitoring (PRM) assay for accurate quantification of low abundance cysteine mutants, these results were aggregated in Skyline to produce a list consisting of precursors for each mass label (light and heavy) for 50 cysteine mutants. Additionally, using Pierce Peptide Retention Time Standards (PRTC, iRT) allowed us to schedule the PRM method, reducing the number of targeted precursors at any given time, and increasing the quality of quantification (**Figure 3D**, *bottom right*). This final scheduled PRM target list with 229 members allowed us to robustly quantify 41 cysteine mutants, 4 Src*WT* cysteine-containing peptides, 14 Src*WT* non-cysteine-containing standards and the 15 iRT peptides (**Figure 3D**, *bottom left*).

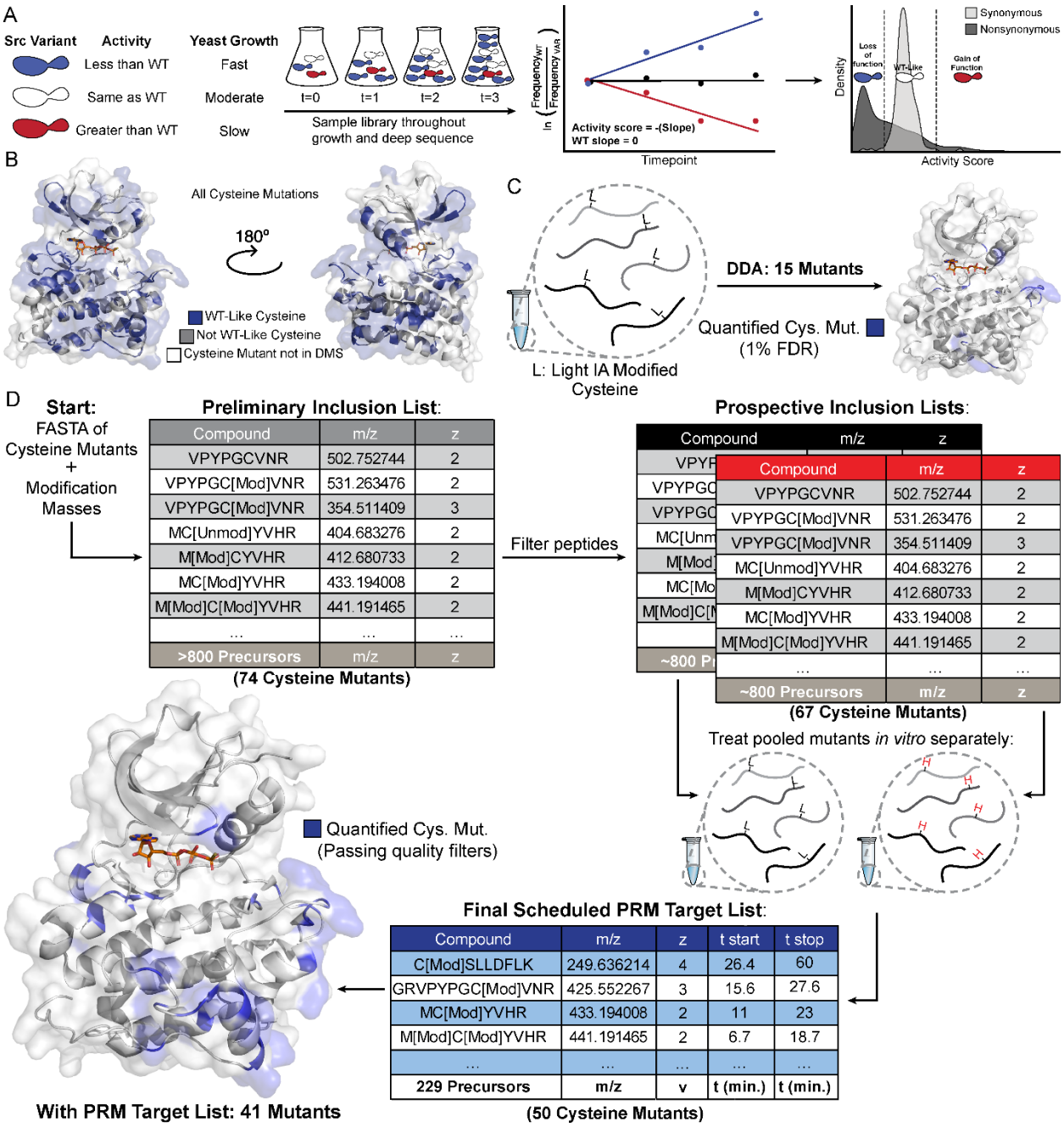


Figure 3. Parallel Reaction Monitoring Assay Development

A. Schematic of the yeast growth-based deep mutational scan (DMS) of Src's kinase domain (Figure panel adapted from Ahler et al. 2019). B. Binned activity scores (Ahler et al. 2019) for all cysteine substitutions mapped onto the Src kinase domain (PDB: 3DQW). C. Initial Data Dependent Acquisition (DDA) attempt to identify cysteine mutants without enrichment or targeting yielded 15 mutant IDs with mutant peptide FDR<1%. D. Schematic of the workflow used to develop spectral libraries for robustly quantifying Src cysteine mutants. Candidate precursors were identified for all 74 mutant cysteines by inputting mutant FASTA sequences and modification masses of heavy and light isotopically-labeled iodoacetamide into Skyline (*Top left*). Precursors were filtered to remove potentially

problematic peptides, including those shorter than 6 amino acids, resulting in prospective inclusion lists containing precursors for 67 cysteine mutants, with one mass list for each mass label (*Top right*). Each prospective DDA inclusion list was run in triplicate, resulting spectra were aggregated using Skyline to generate a spectral library and a corresponding Scheduled PRM Target List containing 50 empirically identified cysteine mutants. Operating the MS in PRM mode yielded robust quantification of 41 mutant cysteine-containing peptides, 4 cysteine-containing Src*WT* peptides, and 14 non-cysteine-containing Src*WT* peptides (*Bottom left*).

Biochemical Modulation and Characterization of Src's Global Conformation

We applied our new workflow to map the kinase domain of Src bound to two different conformation-selective ATP-competitive inhibitors. Previous studies have demonstrated that conformation-selective inhibitors can induce large structural changes in regions that are both proximal and distal to the ATP-binding site (Kung & Jura, 2016). Numerous co-crystal structures of inhibitor-bound kinase complexes have revealed that inhibitors can stabilize large conformational movements in regions lining the ATP-binding pocket (Tong & Seeliger, 2015; Ung, *et al.* 2018; Zhao, *et al.* 2014). In particular, certain pharmacophores promote large displacements of helix αC—which contains a conserved glutamate residue that forms a salt bridge with the catalytic lysine residue—or the DFG-motif of the activation loop, from their active conformations in a number of kinases (**Figure S4**). Furthermore, we and others have demonstrated that conformation-selective inhibitors are capable of inducing large structural changes that are distal to the ATP-binding site in multi-domain kinases (Aguis *et al.* 2019; Ahler *et al.* 2019; Chakraborty *et al.* 2019; Fang *et al.* 2020; Foda, *et al.* 2015; Krishnamurty *et al.* 2013; Kwarcinski *et al.* 2016; Leonard *et al.* 2014; Register *et al.* 2014; Tong *et al.* 2017). For example, conformation-selective inhibitors modulate the level of SH2 and SH3 domain engagement with the kinase domain of Src by exploiting the allosteric networks that are utilized to control catalytic activity. We felt that a comparison of cysteine alkylation rates between two conformation-selective inhibitor-bound Src complexes would provide an ideal test case for determining the utility of parallel chemoselective profiling for mapping protein structure.

For our study, we selected conformation-selective inhibitors **1** and **2**, which contain pharmacophores that stabilize the helix αC-out and DFG-out conformations of the ATP-binding site of Src, respectively (**Figure 4A**). While we have not obtained co-crystal structures of **1** or **2** bound to Src, their interactions with the ATP-binding site can be inferred from structures of electrophilic analogs of these inhibitors complexed to a cysteine-containing mutant of Src (**Figure S4**; Ahler, *et al.* 2019). Inhibitor **1** contains a 4-phenoxyphenyl substituent at the C-3 position of the pyrazolopyrimidine scaffold which occupies the hydrophobic pocket created by the outward rotation of helix αC. This promotes the helix αC-out conformation of Src and stabilizes a closed global conformation that is characterized by autoinhibitory engagement of the SH2 and SH3 domains with the backside of the kinase domain. Inhibitor **2** contains a 3-trifluoromethylbenzamide substituent at the C-3 position—connected to the pyrazolopyrimidine scaffold through a rigid (2-methylphenyl)-propargyl linker—that occupies the hydrophobic pocket created by the ~180° "flip" of the DFG-motif at the base of the activation loop from an active conformation. The 3-trifluoromethylbenzamide substituent stabilizes the DFG-out conformation of Src and promotes an open global conformation that is characterized by SH2 and SH3 domains that are disengaged from the kinase domain. Because both inhibitors share an identical pyrazolopyrimidine scaffold and only vary in the pharmacophore responsible for stabilizing different ATP-binding site conformations, most observed differences in our parallel chemoselective profiling experiments will likely be due to structural changes in the kinase domain.

Prior to performing parallel chemoselective profiling experiments with inhibitor-bound complexes, we benchmarked the effects of inhibitors **1** and **2** on the global conformation of Src. First, we characterized how each inhibitor modulates the global conformation of Src with a limited proteolysis assay using a Src construct consisting of the SH3, SH2 and kinase domains (Src 3D).

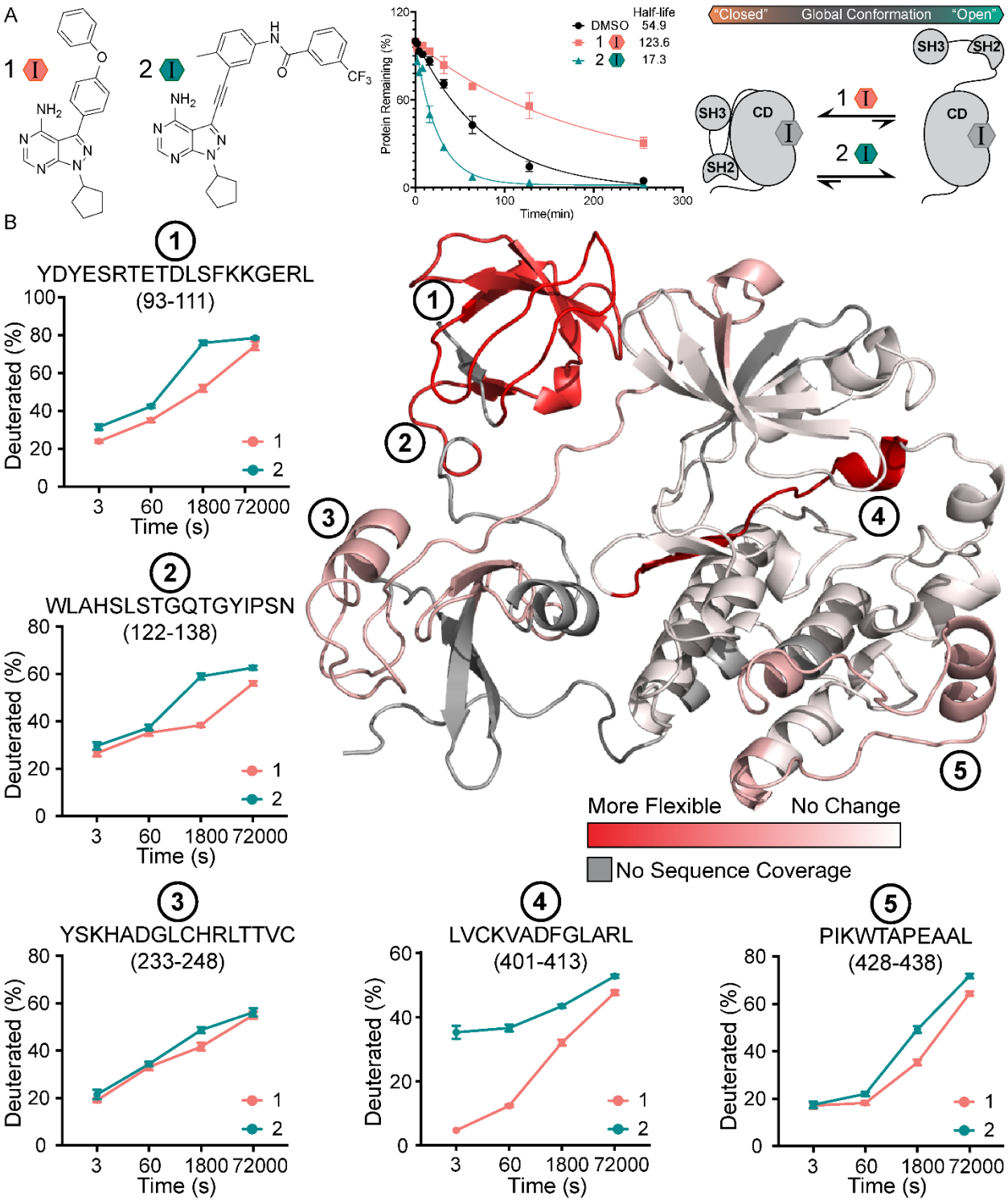


Figure 4. Biochemical Characterization of Inhibitor-Bound Src 3D Complexes

A. Structures of inhibitors 1 and 2 (Krishnamurty et al., 2013; Fang et al., 2020). Quantification of the limited proteolysis experiments performed with inhibitor-bound complexes of Src 3D. Points represent mean ± SEM. B. HDX-MS analysis of inhibitor-bound Src 3D complexes. Deuteration differences between the inhibitor 1-Src and 2-Src 3D complexes are plotted on the crystal structure

of Src 3D (PDB: 2SRC) from no change (white) to the largest change (red). Deuterium uptake plots for peptides at regions of interest are shown on the periphery. Points represent mean ± SEM.

Previously, it has been shown that the protease thermolysin can cleave the linker between the SH2 and kinase domains of Src (SH2-KD-linker), with the rate of cleavage, which we report as a half-life, correlating to Src's global conformation (Agius *et al.* 2019; Fang *et al.* 2020; MacAuley & Cooper 1989). Consistent with inhibitor **2** stabilizing the DFG-out conformation of Src and promoting a more open global conformation, we found that thermolysin more rapidly cleaved the SH2-KD-linker of the inhibitor **2**-Src complex than of *apo* Src (DMSO) (**Figure 4A**, *middle*). We observed the opposite trend for inhibitor **1**, with the inhibitor **1**-Src complex showing less susceptibility to cleavage by thermolysin than *apo* Src. Thus, inhibitors **1** and **2** promote the expected global conformations of Src.

To further benchmark how inhibitors **1** and **2** influence the structure of Src, we performed Hydrogen Deuterium Exchange-Mass Spectrometry (HDX-MS) experiments with the inhibitor **1**-Src 3D and inhibitor **2**-Src 3D complexes. To ensure that quantitively-bound inhibitor-Src 3D complexes were formed, we used saturating concentrations of both inhibitors. Following generation of each inhibitor-Src 3D complex, we initiated deuterium exchange by adding buffered D_2O (containing 90% deuterated buffer). At specific timepoints exchange reactions were quenched by lowering the pH to 2.5 and immediately flash freezing samples to lock exchanged deuterium in place and denature the protein. Samples were then processed using standard protocols (see "Methods"). After filtering out peptides showing weak signal or partial overlap, 32 unique peptic peptides remained, which cover 73% of Src 3D's sequence. In determining which peptic peptides demonstrate a difference between the two inhibitor-bound complexes, we excluded peptides with overlapping sequences that demonstrate different levels of deuteration in order to reduce the possibility of interpreting noise. This yielded a relatively conservative set of peptides for further analysis. A summary of the deuteration exchange

kinetics from the two inhibitor-bound Src 3D complexes are represented on the crystal structure shown in **Figure 4B**.

Comparison of the HDX-MS results for the inhibitor **1**-Src 3D and inhibitor **2**-Src 3D complexes shows that there are five regions, represented by peptides 1 through 5 in **Figure 4B**, that demonstrate the largest differences in deuterium exchange kinetics. Consistent with inhibitor **2** generally increasing the solvent accessibility of Src 3D, we observe that all five of these regions underwent faster exchange kinetics in the inhibitor **2**-Src 3D complex relative to the inhibitor **1**-Src 3D complex. The largest overall difference in exchange kinetics was observed for the peptic peptide that contains the DFG-motif in the activation loop (peptide 4), which is consistent with inhibitor **2**'s stabilization of an inactive form of the ATP-binding site where the DFG-motif is "flipped" ~180° relative to the active conformation (DFG-out conformation). Additionally, our HDX-MS experiments recapitulate the observations made using limited proteolysis experiments regarding the global conformation of Src 3D (**Figure 4B**). We found that peptic peptides 93-111 (peptide 1), 122-138 (peptide 2), and 233-248 (peptide 3), which are located on the regulatory SH2 and SH3 domains, demonstrated faster exchange kinetics in the inhibitor **2**-Src 3D complex. This result is congruent with inhibitor **2** promoting a more open global conformation of Src 3D. Interestingly, deuteration level differences on the complementary surfaces of the kinase domain that directly contact the SH2 and SH3 domains were not observed, which is likely due to the kinetic regime used in this experiment (Hamuro, 2017). Finally, in the inhibitor **2**-Src 3D complex, we also observed an unexpected increase in deuteration rate for the region represented by peptic peptide 428-438 (peptide 5), which is located at the bottom of the C-terminal lobe of the kinase domain. While the overall measured difference in exchange rate in this region is modest, this structural perturbation which is likely the result of a long-range allosteric network is an intriguing result worthy of further inquiry. Together, these HDX-MS experiments provide a basis for comparison with our parallel chemoselective profiling method.

Parallel Chemoselective Profiling of Inhibitor-Bound Src Complexes

We next performed analogous mapping experiments using our parallel chemoselective profiling workflow. Because our method can be performed with lower concentrations of input protein relative to HDX-MS, these experiments were performed with Src FL, which contains Src's membrane-interacting N-terminus, rather than Src 3D. To gain residue level insight into the changes in surface accessibility and dynamics of Src's kinase domain upon complexation with inhibitors **1** and **2**, we performed parallel chemoselective profiling experiments using our Src^{CysLib}. In a similar manner that we prepared inhibitor-Src 3D samples for HDX-MS, we separately formed quantitative inhibitor-bound Src FL complexes by incubating Src^{CysLib} with saturating concentrations of inhibitors **1** or **2**, followed by the addition of three different concentrations of heavy iodoacetamide (**Figure 5A**, *left)*. Heavy iodoacetamide concentrations were selected based on their likelihoods of providing sub-stoichiometric labeling over the timecourse of the experiment, which we reasoned would allow us to measure differences in cysteine alkylation rates over multiple timepoints. We then removed and quenched aliquots of each reaction after 30, 60, or 120 min after the addition of heavy iodoacetamide (**Figure 5A**, *right*). Quenching was achieved by denaturing Src^{CysLib} in guanidinium buffer containing a sufficiently high concentration of light iodoacetamide to outcompete any background labeling from residual heavy iodoacetamide. We then heated reactions to 37 °C and digested with mass spectrometric grade trypsin overnight. Following peptide desalting, peptides were analyzed using our PRM workflow and spectra were imported directly into Skyline for quantification.

For our analysis, we only considered residues that had been quantified in two-thirds of our experimental samples, yielding a relatively conservative set of peptides that were robustly quantified in the majority of samples (see "Methods"). This allowed us to utilize 23 cysteine mutants and 4 wild-

type cysteines to probe the structural differences between the inhibitor **1**-Src^{CysLib} and inhibitor **2**-Src^{CysLib} complexes.

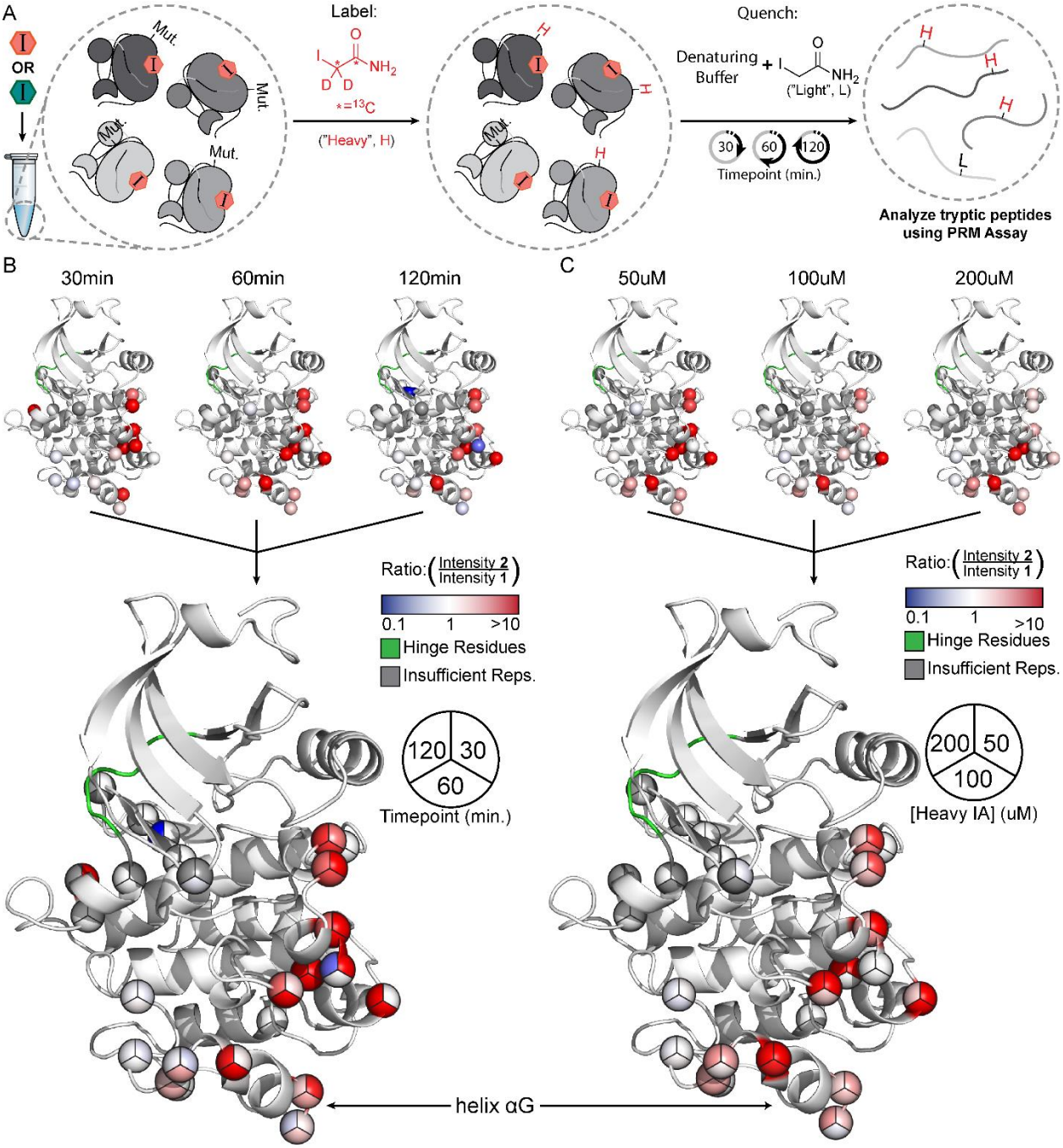


Figure 5. Parallel Chemoselective Profiling of Conformation-Selective Inhibitor Bound Full-Length Src

A. Schematic of the Parallel Chemoselective Profiling experiment. Conformation-selective inhibitors **1** or **2** were added to the Src Cysteine Library (SrcCysLib) at saturating concentrations to form inhibitor-Src FL complexes. After complexation, heavy iodoacetamide was added to a final concentration of 50, 100 or 200 μM. At the indicated timepoints, aliquots of the reaction were removed and quenched. Protein was then digested using MS grade trypsin and peptides were analyzed

using our PRM workflow. B-C. **2**-SrcCysLib/**1**-SrcCysLib intensity ratios from select conditions mapped on the kinase domain of Src (residues 264-536, human Src numbering) extracted from PDB ID: 2Src. Crystal structures are superimposed to generate circle charts. Hinge region residues are shown in green to highlight the ATP-binding site.

These cysteine residues line a significant portion of the ATP-binding site, including multiple residues on the activation loop, which allowed us to probe local structural differences of the inhibitor-bound Src complexes. Additionally, we quantified a number of residues on the C-terminal lobe of the kinase domain, including in positions that significantly overlap with the region (peptide 5, **Figure 4B**) that demonstrated unexpected dynamics in our HDX-MS experiments. While our coverage of residues that line the SH2/SH3 regulatory interface in the kinase domain were poor, we were able to quantify a wild-type cysteine residue that lines the C-terminal tail-binding interface of the SH2 domain. Finally, these are several mutant cysteine residues on the kinase domain that are not solvent exposed. These positions served as controls for interpreting our differential alkylation results.

To visualize our parallel chemoselective profiling data, we first generated (**2**-Src^{CysLib}/**1**-Src^{CysLib}) intensity ratios of the heavy iodoacetamide labeled peptides for each cysteine residue at each timepoint and concentration of heavy iodoacetamide, which provided nine conditions in total. We then mapped these ratios as a heatmap onto the kinase domain from a crystal structure of Src (**Figure 5B-C**, *top*). To facilitate data interpretation, we superimposed the crystal structures from each condition and represented the heatmaps for each residue as a pie chart (**Figure 5B-C,** *bottom*).

The insights provided from our parallel chemoselective profiling experiments are highly consistent with our HDX-MS results. First, for almost all cysteines that we observe a difference in labeling between the two inhibitor-bound complexes, the (**2**-Src^{CysLib}/**1**-Src^{CysLib}) intensity ratios are >1. This is consistent with inhibitor **2** increasing the overall solvent accessibility of Src FL. Furthermore, we found that the wild-type cysteine residue located in the C-terminal tail-binding interface of the SH2 domain underwent more rapid labeling in the presence of inhibitor **2**, consistent with **2** promoting an

open global conformation of Src FL (**Figure S6A**). We also found high (**2**-Src^{CysLib}/**1**-Src^{CysLib}) intensity ratios for a number of cysteine mutant residues in the activation loop, indicating an increase in solvent exposure for this region when Src FL is bound to inhibitor **2**. While one of these residues (L413C) overlaps with peptic peptide 4 from our HDX-MS experiments, five of these cysteine mutations (E415C, T420C, A421C, R422C and Q423C) are C-terminal to the DFG-motif and lie in a region of the activation loop that is highly solvent exposed and dynamic (**Figure S6B**). Unlike the DFG-motif-containing peptic peptide, this region of the activation loop was highly deuterated at early timepoints and did not show a difference in deuteration levels in our HDX-MS experiments, likely due to the flexibility and intrinsic solvent exposure of this region. Our chemoselective profiling data shows that flipping of the DFG-motif affects the dynamics of distal parts of the activation loop and highlights how our method can complement other structural characterization techniques.

In addition to characterizing local and global structural differences between inhibitor-bound Src FL complexes, we also observed high (**2**-Src^{CysLib}/**1**-Src^{CysLib}) intensity ratios in regions of the kinase domain distal to the ATP-binding site. Most notably, we found that inhibitor **2** led to a dramatic increase in the labeling of mutant cysteine residues that lie in an area of the kinase domain's C-terminal lobe that overlaps with the region represented by peptic peptide 5 from our HDX-MS experiments relative to inhibitor **1**. This region of the C-terminal lobe includes residues R472C, E478C and R479C, which reside on the αG helix (**Figure 5B**, *bottom*). It is intriguing that one of these C-terminal lobe residues (R472C) demonstrates intensity ratios that are comparable to a residue (L413C) that is directly adjacent to the flipped DFG-motif in the activation loop, despite being more than 20 Å away from the ATP-binding site. R472C shows higher (**2**-Src^{CysLib}/**1**-Src^{CysLib}) intensity ratios than the two residues at the opposite end of the helix αG, suggesting that the solvent accessibility of this region of helix αG is most influenced by ATP-binding site occupancy. Both our HDX-MS and chemoselective profiling results are consistent with a hypothesized allosteric coupling between the ATP-binding site and the

helix αG (Tong, *et al.* 2017). Given the spatial connectivity of the residues that show increased labeling in the presence of inhibitor **2** from the activation loop to helix αG, our chemoselective profiling method suggests the potential composition of the allosteric network that connects these two spatially separated regions. Our ability to obtain residue level information into what appears to be an extended allosteric network provides complementary insight into an intriguing trend that we observed with HDX-MS.

Broader Application of Parallel Chemoselective Profiling

We anticipate that parallel chemoselective profiling will be amenable to amino acid substitutions beyond cysteine. To explore this possibility, we considered other amino acids that can be labeled with selective reagents. Recently, a chemoselective labeling reagent for methionine was reported (Li *et al.* 2017). We analyzed the mutational effect scores of methionine substitutions for the same DMS datasets listed in **Table 1**. We applied the same three bins for methionine mutational effect scores that were used when analyzing the cysteine mutational effect scores and mapped them to the crystal structures of ubiquitin, PTEN, Src and BRCA1 (**Figure 6A-D**). We also performed the same tolerance analysis, which revealed trends similar to those observed for cysteine. With the exception of Src, more than half of all methionine substitutions surveyed in the DMS data were classified as WT-like (**Figure 6E,G,I,K**). Additionally, the solvent exposure analysis conducted with Pymol revealed that a majority of WT-like substitutions are on solvent exposed sidechains (**Figure 6F,H,J,L**).

Combined, these analyses suggest that methionine is also a promising candidate for use in parallel chemoselective profiling.

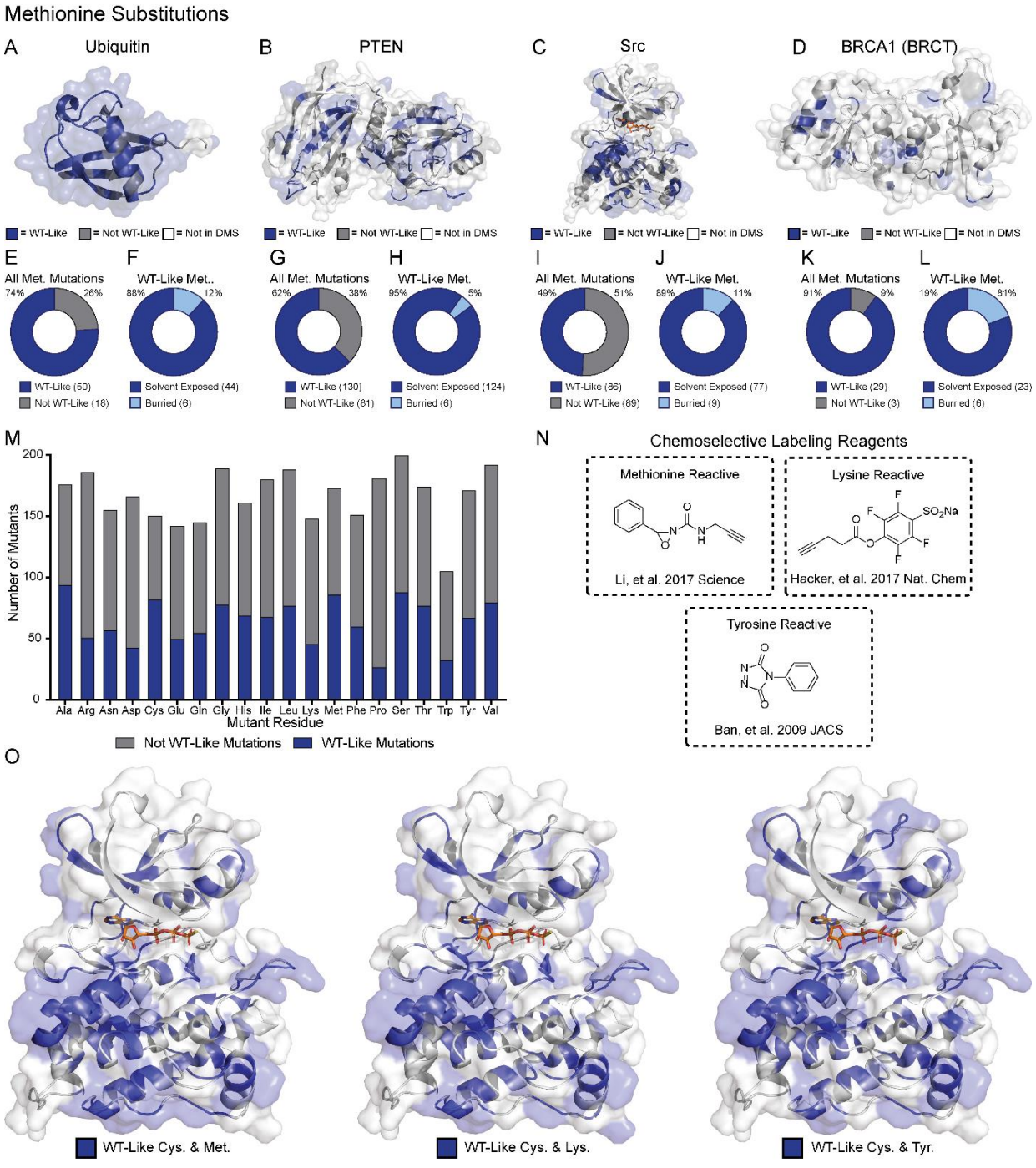


Figure 6. Extending Parallel Chemoselective Profiling Beyond Cysteine

A. Fitness scores for methionine substitutions (Roscoe, B.P. et al. 2013) mapped to the crystal structure of ubiquitin (PDB: 1UBQ). B. VAMP-seq scores for methionine substitutions (Matreyek, K.A. et al. 2018) mapped to the crystal structure of PTEN (PDB: 1D5R). C. Activity scores for methionine substitutions (Ahler, E. et al. 2019) mapped to the crystal structure of Src's kinase domain

(PDB: 3DQW). D. Function scores for methionine substitutions (Findlay, G.M. et al. 2018) mapped to the crystal structure of BRCA1 BRCT domain (PDB: 1T29). E-L. Fraction of all WT-like methionine substitutions that are solvent exposed in each crystal structure E-L. Each subpanel relates to the protein depicted immediately above. M. Tabulation of all single amino acid substitutions from the DMS of Src kinase domain (Ahler, E. et al. 2019) and the fraction of those that result in WT-like substitutions. N. A selection of chemoselective labeling reagents that could be used to extend this method beyond cysteine residues. O. WT-like methionine, lysine or tyrosine residues mapped with WT-like cysteine residues demonstrate the potential additional coverage that would be possible by combining chemoselective labeling of multiple residues.

Methionine does not appear to be a special case. For example, if we consider all 3,506 variants from the Src kinase domain DMS dataset, there are other amino acid substitutions that are well tolerated as indicated by the fraction of each amino acid substitution characterized as WT-like (**Figure 6M**). Similar ratios of WT-like substitutions relative to the total number of substitutions quantified are observed for the other DMS datasets contained in **Table 1** (**Figure S7**). Among these well-tolerated amino acid substitutions are several examples of residues with that can be labeled with selective reagents (**Figure 6N**).

Discussion

In this work we present parallel chemoselective profiling, a solution-based method for studying protein structure and dynamics. Chemoselective profiling leverages DMS data to install amino acids with desired chemistries at defined locations on a protein surface and differential labeling of mutant residues—measured in parallel using targeted mass spectrometry—is used to infer structural changes. Our method has the potential to be generalized to any protein target of interest and, importantly, does not require specialized instrumentation beyond a MS capable of obtaining high-resolution tandem mass spectra. To demonstrate the utility of our method, we applied it to the tyrosine kinase Src. By performing a comparative chemoselective profiling experiment with conformation-selective inhibitor-bound Src complexes, we were able to characterize how ATP-binding site occupancy influences the structure and dynamics of full-length Src for the first time. By combining

our chemoselective profiling results with complementary HDX-MS experiments, we were able to observe what appears to be an extended allosteric network within the kinase domain of Src FL, linking the ATP-binding site, activation loop and helix αG, which resides more than 20 Å away from the ATP-binding site.

A majority of our parallel chemoselective profiling data significantly overlapped with the HDX-MS experiments that we performed, allowing us to cross-validate the insight we obtained into how conformation-selective ATP-competitive inhibitor binding affects Src structure and dynamics. In some areas, parallel chemoselective profiling and HDX-MS proved to be complementary. The lower concentration of input protein required for parallel chemoselective profiling compared to HDX-MS allowed us to profile a library of Src FL constructs, which contain a membrane-interacting N-terminal region, rather than the solubility-optimized truncation variant Src 3D that was used in our HDX-MS experiments. Furthermore, parallel chemoselective profiling provided residue-level insight into Src's structure and dynamics in some areas of Src's kinase domain where HDX-MS showed more modest effects. We did, however, find that there are regions in the kinase domain that we were not able to profile using our chemoselective profiling method but that were probed with HDX-MS. This leaves room for future improvements in our parallel chemoselective profiling workflow, which can be tuned for specific applications. Two regions that were particularly under-represented in our parallel chemoselective profiling data were the αF pocket and N-terminal lobe of Src's kinase domain. We speculate that despite these areas possessing WT-like and solvent exposed cysteine mutations that poor coverage can be attributed to unfavorable properties of the tryptic peptides in this region (for example, high charge state and/or suboptimal length). One potentially useful avenue that could be explored for increasing the coverage of these areas would be to redistribute the trypsin cut-sites on peptides that contain WT-like cysteine mutations. The addition or elimination of trypsin cut-sites can be guided by DMS data, which can provide suitable lysine and arginine replacements or suggest

positions where WT-like lysine or arginine mutations can be introduced. The optimal combination of trypsin cut-sites could then be determined empirically. If implemented as a modification to the general approach, cut-site engineering could have the potential to dramatically increase the coverage of regions of interest that are poorly characterized in first pass attempts at quantifying mutants for the PRM assay due to the unfavorable properties of certain tryptic sequences.

Another potential strategy for increasing the solvent accessible surface available for profiling is the use of multiple separate single amino acid-mutant libraries in combination. For example, if we consider the mutational effect scores for methionine, lysine and tyrosine—which all can be selectively labeled with suitable reagents—in addition to cysteine, we can increase the solvent-exposed surface of Src's kinase domain sampled by WT-like mutants (**Figure 6O**). For any protein of interest, it should be possible find complementary combinations of WT-like amino acids that provide maximal coverage of solvent exposed surfaces. Taken together, the parallel chemoselctive profiling method presented here introduces a useful addition to the relatively small toolkit available for studying protein structure and dynamics in solution.

Materials and Methods

Src WT-like Cysteine Library (Src^{CysLib})

Positions of cysteine mutants were informed by the deep mutation scanning (DMS) data previously reported (Ahler *et al.,* 2019). Briefly, interactions between the regulatory domains of Src and the catalytic kinase domain of Src serve negative regulatory roles. Disruption of these interfaces releases this inhibition and increases Src kinase activity. Site saturation mutagenesis was employed to generate single amino acid variants of Src kinase domain in the context of full-length enzyme (3,506, ~70% of possible mutants were made and assayed). This mutant library was adapted to a yeast growth assay (Kritzer *et al.* 2018). Once transformed into *S. Cerv.* samples were taken at different timepoints

and deep-sequenced to quantify each mutant. "WT-like" mutations were defined as those mutants with activity scores within one standard deviation of the activity scores for all synonymous *wt* Src variants.

There are 82 individual WT-like cysteine mutants in the DMS dataset, of which 74 were constructed using an inverse PCR approach with no optimization (**Figure S1A**). Expression and purification of full-length Src from bacterial expression systems has proven to be inefficient in our hands compared to yields of recombinant protein from truncated variants of Src. To address this, we made several modifications to previously used expression protocols. First, we selected the FreeStyle 293F (ThermoScientific) as our protein production host to take advantage of the endogenous co-factors and chaperones to increase the yield of folded protein. Because the N and C termini of Src are known to be important for regulating the conformational equilibria of Src, we desired recombinant Src cysteine mutants to have near-native termini. To achieve this, we expressed Src cysteine mutants as His6-SUMO* fusions for two reasons. First, expressing proteins as SUMO fusions has been shown to increase fusion protein solubility and expression levels (Butt *et al.*, 2005). Endogenous mammalian SUMO proteases are able to cleave SUMO fusions, so we employed the complementary SUMO protease Ulp1* which, like wild-type Ulp1 cleaves C-terminal to a Gly-Gly motif releasing the fused protein with a "scarless" or native N-terminus (**Figure S1B**) (Liu *et al.* 2008). These modifications allowed us to purify full-length Src

Mutagenesis primers were designed by centering each primer at the mutagenic position (**Figure S#B**). The TGC codon was used to encode the cysteine mutation at each position. Selections with T_m 55°C were selected 5' and 3' of each mutation. The combined sequence constituted the forward mutagenic primer. The reverse primer is the reverse complement of the 5' T_m 55°C portion of the forward primer.

PCR and thermocycler conditions were as follows:

7.5uL 2x Q5 Master Mix (New England Bioscience)
0.45uL 10uM Forward primer
0.45uL 10uM Reverse Primer
2uL 10ng/uL template
5.05uL H_2O

Temperature	Time
98°C	20sec
98°C	20sec
55°C	30sec
72°C	1min/kb
Go to Step 2	X10
72°C	5min

After amplification, digest template by adding 1uL FastDigest Dpn1 (ThermoScientific) to each reaction, incubate at 37 °C for 30 min. Transform 1uL reaction mixture into 10uL high efficiency NEB5α using heat shock, 120uL SOC snf s 30min outgrowth at 37°C, 300RPM shaking. The entire reaction was plated on LB+Carb. plates and incubated overnight at 37°C.

Solvent exposure analysis of Ubiquitin, PTEN, Src (kinase domain) and BRCA1 (BRCT domain)

Crystal structures for each protein were imported into PyMOL for analysis. The script FindSurfaceResidues was downloaded from the PyMOL Wiki (https://pymolwiki.org/index.php/FindSurfaceResidues) and used with its default settings (2.5 Ang^2 cutoff). After residues were identified by the script, they were visually inspected to ensure that the sidechain and not the backbone of each residue were solvent exposed. Numbers of residues were tabulated and reported as fractions of the total length of crystalized protein.

HDX-MS of 3D Src in presence of conformation selective inhibitors

0.2 mg/mL 3D Src was preincubated with 20 mM conformation selective inhibitor **1** or **2** in protein dilution buffer (50 mM HEPES, pH 7.8, 150 mM NaCl, 1mM DTT, 5% glycerol) at room temperature for 30 min to prepare a kinase-ligand complex. 10 mL of this was then added to 90 mL of buffered D_2O (prepared 5 mL with 4.5 mL of D20 and 0.5 mL of 10x protein dilution buffer and 0.2 ug/mL of Gu-1-Fibrino peptide) to initiate deuteration at 22°C. Deuterium exchange was quenched after 3 s, 1 min, 30 min, and 20 hr by adding the reaction to 100 mL of ice-cold quench buffer (0.2% formic acid, 8M Urea, 0.1% trifluoroacetic acid) for a final pH of 2.5. All time points for each inhibitor were collected in triplicate. Samples were immediately frozen in a dry ice/ethanol and stored at −80°C. Undeuterated samples were prepared the same except using buffered H_2O instead of D_2O.

Frozen samples were thawed on a 5 °C block for 4 minutes prior to injection onto a loading loop. The loaded sample was passed over a custom packed pepsin column (Porcine pepsin

immobilized on POROS 20-AL resin; 2.1 x 50 mm column) (PMID 12096131) kept at 12 °C with a flow of 0.1% trifluoroacetic acid (TFA) and 2% acetonitrile (ACN) at 200 μL/min.

Digested peptic fragments were trapped onto a Waters XSelect CSH C18 XP VanGuard Cartridge (2.1 x 5 mm, 2.5 μm). After 5 minutes of loading, digestion, and trapping, peptides were resolved on an analytical column (Waters BEH 1 x 100 mm, 1.7μm, 130Å) using a gradient of 3 % to 40 % solvent B for 9 minutes (A: 0.1 % FA, 0.025 % TFA, 2 % ACN; B) 0.1 % FA in ACN). The LC system was coupled to a Thermo Orbitrap performing full scans over the m/z range of 300 to 1500 at a resolution of 30,000. The MS source conditions were set to minimize loss (PMID 22965280). Undeuterated samples were run prior to and at the end of all the LC-MS runs.

During the analytical separation step, a series of 250 μL injections were used to clean the pepsin column: 1) 0.1 % Fos-12 with 0.1 % TFA; 2) 2 M GndHCl in 0.1 % TFA; 3) 10 % acetic acid, 10 % acetonitrile, 5 % isopropanol (PMID 22993047 and 29299838). After each gradient the trapping column was washed with a series of 250 μL injections: 1) 10 % FA; 2) 30 % trifluoroethanol; 3) 80 % methanol; 4) 66 % isopropanol, 34 % ACN; 5) 80 % ACN. During the trap washes the analytical column was cleaned with three rapid gradients (PMID 21643454).

Peptic peptides were identified from data-dependent acquisition (DDA) experiments on undeuterated samples by exact mass and tandem mass spectrometry (MS/MS) spectra using Protein Prospector (PMID 24591702) filtering with a score cutoff of 15. Mass shifts were determined using HD-Examiner v2 (Sierra Analytics). The GluFib (CAS: 103213-49-6, Sigma) internal standard peptide was checked in all samples to verify that back-exchange levels were consistent in all experiments (PMID 22571272).

We binned peptides based on their fractional deuteration levels. We assigned the deepest shade of red to the peptide KAVDFGLARL which showed the maximal differential fractional deuteration

level of 50.02% between the two inhibitors. We then normalized color intensity from red to white for the rest of the peptides and mapped these to the crystal structure of 3D Src (**Figure 4B**).

Half-life characterization of 3D Src conformation selective inhibitor bound complexes

Limited proteolysis of Src in presence of inhibitors was modified from (PMID: 31287657). Briefly, 1 mM of Src was preincubated with 20 mM of the inhibitors **1**, **2** or DMSO (4%) in proteolysis buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.5 mM $CaCl_2$) at room temperature for 30 min to prepare a kinase-ligand complex. Proteolysis was initiated by adding 3.8 mM Thermolysin (Promega, catalog number: V4001) stock solution to the kinase-ligand complex to a final concentration of 60 nM. 20 μL of this mixture was then added to 10 μL of 50 mM EDTA to terminate proteolysis at various time points (0, 2, 4, 8, 16, 32, 64, 128, 256). The quenched samples were analyzed by SDS-PAGE (12 % Bis-Tris gel in SDS running buffer, and stained with SYPRO Ruby ThermoFisher Scientific: catalog number S12000). Band intensities were analyzed by ImageStudioLite imaging software. Percent protein remaining was computed based on band intensity at 0 min and was plotted against time on GraphPad Prism 8. The curve was fit to an exponential decay equation using GraphPad Prism 8 software to obtain half-lives of each kinase-ligand complex.

Parallel Reaction Monitoring (PRM) of Src Cysteine Mutants

Samples were prepared for spectral libraries by denaturing Src^{CysLib} in buffer containing isotopically light or heavy (Sigma 721328) iodoacetamide; final concentrations: 860nM pooled cysteine mutant protein, 100mM Tris pH 7, 1.6M guanidinium HCl, 1mM $CaCl_2$, 1.3% DMSO, 10mM iodoacetamide in 75uL. Samples were heated to 37 °C for 10 min, then 1uL of 0.5ug/uL MS grade trypsin was added. pH was adjusted to 8.5 using 1N NaOH and samples were incubated overnight at 37°C with 500rpm orbital shaking. Samples were acidified with 15uL 88% formic acid and 2uL of 1.5pmol/uL Pierce retention time calibration mix was added to each sample (ThermoScientific 88321).

Samples were desalted on C-18 stagetips made in house by washing the stagetip with 50uL Buffer B (80% acetonitrile in water, 0.1% TFA), equilibrating the stagetip with 50uL Buffer A (5% acetonitrile in water, 0.1% TFA), applying the sample, washing with 50uL Buffer A and eluting into a Lo-Bind microcentrifuge tube using 50uL Buffer B. Peptides were then loaded on a self-pulled 360 μm OD x 100 μm ID 15 cm column with a 7 μm tip packed with 3 μm Reprosil C18 resin (Dr. Maisch, Germany). Peptides were analyzed by nanoLC-MS in a 60 minutes linear gradient from 10% to 35% buffer B (buffer A: 0.1% acetic acid; buffer B: 0.1% acetic acid, 80% acetonitrile) (Thermo EASY nLC 1200) on an Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer. Top Speed data-dependent acquisition with 3 second cycle time was used with each inclusion list. The method consisted of Orbitrap FTMS spectra (R = 60 000 at 200 m/z; m/z 350–1600; 7e5 target; max 20ms ion injection time) for MS1 and HCD MS/MS spectra (R = 30 000 at 200 m/z; 29% CE; 5e4 target; max 100 ms injection time) were collected with an intensity filter set at 2.5e4 and dynamic exclusion for 15 second. Inclusion lists for each spectral library (heavy or light) were generated by importing a FASTA sequence for each of the 74 cysteine mutant full length Src sequences into Skyline. Precursors were identified by filtering all peptides for unique 6 amino acid or longer peptides, allowing up to two missed cleavages and charge states +2 through +5 with variable modifications on cysteine for each mass label, and variable oxidation of methionine. To generate spectral libraries, resultant DDA RAW files were converted to the mzML file format using MSConvert and search using MSGF+ against the combined Src cysteine mutant FASTA database which contained the Pierce iRT peptide sequences. The resulting mzID files were imported into Skyline for analysis and generation of spectral libraries.

PRM assay samples were prepared as master mixes in triplicate for each heavy iodoacetamide concentration—50, 100 and 200uM. First, Src^{CysLib} was diluted to 1.3uM in 100mM Tris pH 7.6, 2mM $MgCl_2$, 100mM NaCl with 20uM of **1** or **2** added from a DMSO stock (2% DMSO final). Inhibitor-Src^{CysLib} complexes were formed at RT for 30 min after which heavy iodoacetamide (Sigma 721328)

was added from a 50X stock. At the indicated timepoints (30, 60 or 120min) 50uL was taken from each master mix and quenched in 25uL 3X denaturing buffer: 4.5M guanidinium HCl, 3mM $CaCl_2$, 15mM TCEP, 30mM iodoacetamide. Samples were then processed like library samples (above). Spectra were collected for each sample as described above with the following modifications. The Orbitrap spectra were collected using the following settings: HCD CE at 29%, R = 15 000 at 200 m/z, m/z 350-1100, 5e4 target, max injection time at 22 ms, with the scheduled PRM target list

After collecting data, raw spectra were imported into Skyline for peptide-centric analysis and intensity values for each peptide were exported for further processing in Microsoft Excel. We analyzed peptides that were contained intensity values in either heavy or light channel in two thirds of experimental samples. This included several peptides that did not have heavy intensity values in the experimental samples. These cysteines did not have an appreciable amount of heavy alkylation product and we therefore imputed intensity values of 1. We normalized the heavy intensity values to internal standards (iRT and Src*WT* peptides). Ratios of heavy labeled peptides from inhibitor **2**/**1** were then calculated and mapped to the crystal structure of Src (PDB IS: 2Src). The crystal structures were then overlaid in Photoshop to generate pie charts for each residue.

Supplemental Information for Chapter 1

Table S1: Components of existent protein footprinting methods. Related to Figure 1.

Method	Label	Necessary Instrumentation	Resolution Level	Additional Considerations
HDX-MS	Deuterium	High-resolution MS + advanced LC for chromatography at ~0° C	Peptide (Backbone amides)	Generally, buffers pH 6-8
FPOP*	Hydroxyl radical	High-resolution MS + standard LC + KrF laser or high-power UV light source	Residue (hydroxyl radical-reactive sidechains)	Choose buffer components that do not unintentionally quench hydroxyl radicals
Parallel Chemoselective Profiling	Residue dependent (e.g. iodoacetamide with cysteine)	High-resolution MS + standard LC	Residue (Installed/defined location)	Buffer conditions suitable for labeling reagent

*FPOP: Fast Photo-Chemical Oxidation of Proteins (Hambly and Gross, 2005)

Supplemental Figure 1

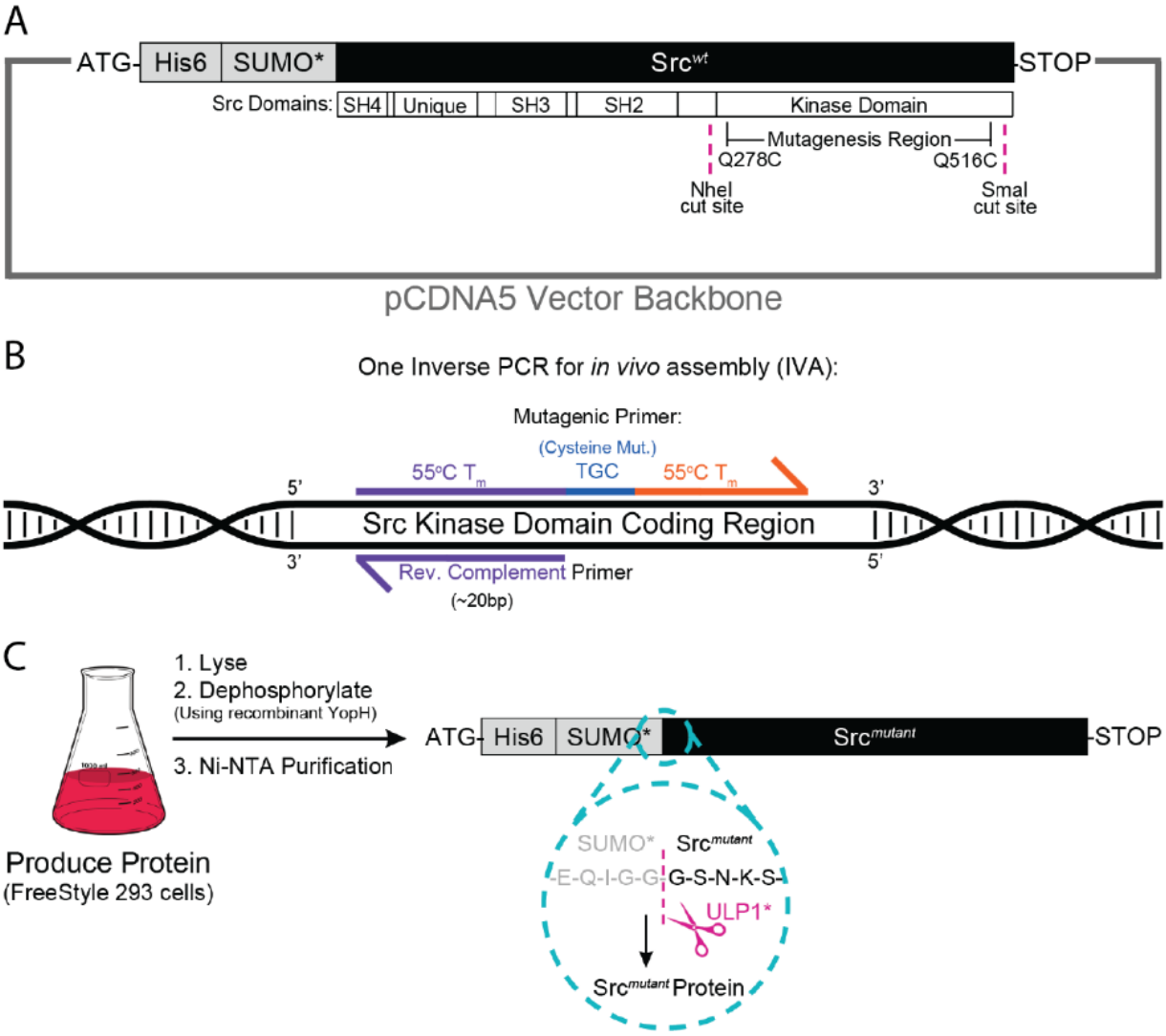


Figure S1. Key Features and Methods of Full-Length Src Mutant Protein Expression and Purification

A. Mammalian expression plasmid with Src domain architecture and mutagenesis region noted. Silent NheI cut site was introduced to allow for the facile assembly of pooled mutant constructs with varying domain architecture (i.e. truncation mutants). B. Primer design strategy for inverse PCR followed by transformation of linear DNA product into NEB5alpha bacteria for *in vivo* assembly. C. Full-length Src mutants were expressed in a pooled format in suspension FreeStyle 293F cells. As the N- and C-termini of Src are important for its regulation, protein was expressed as a direct SUMO* fusion. Following standard Ni-NTA purification procedures, protein was cleaved yielding a pool of individual cysteine mutants with near native N-termini.

Supplemental Figure 2

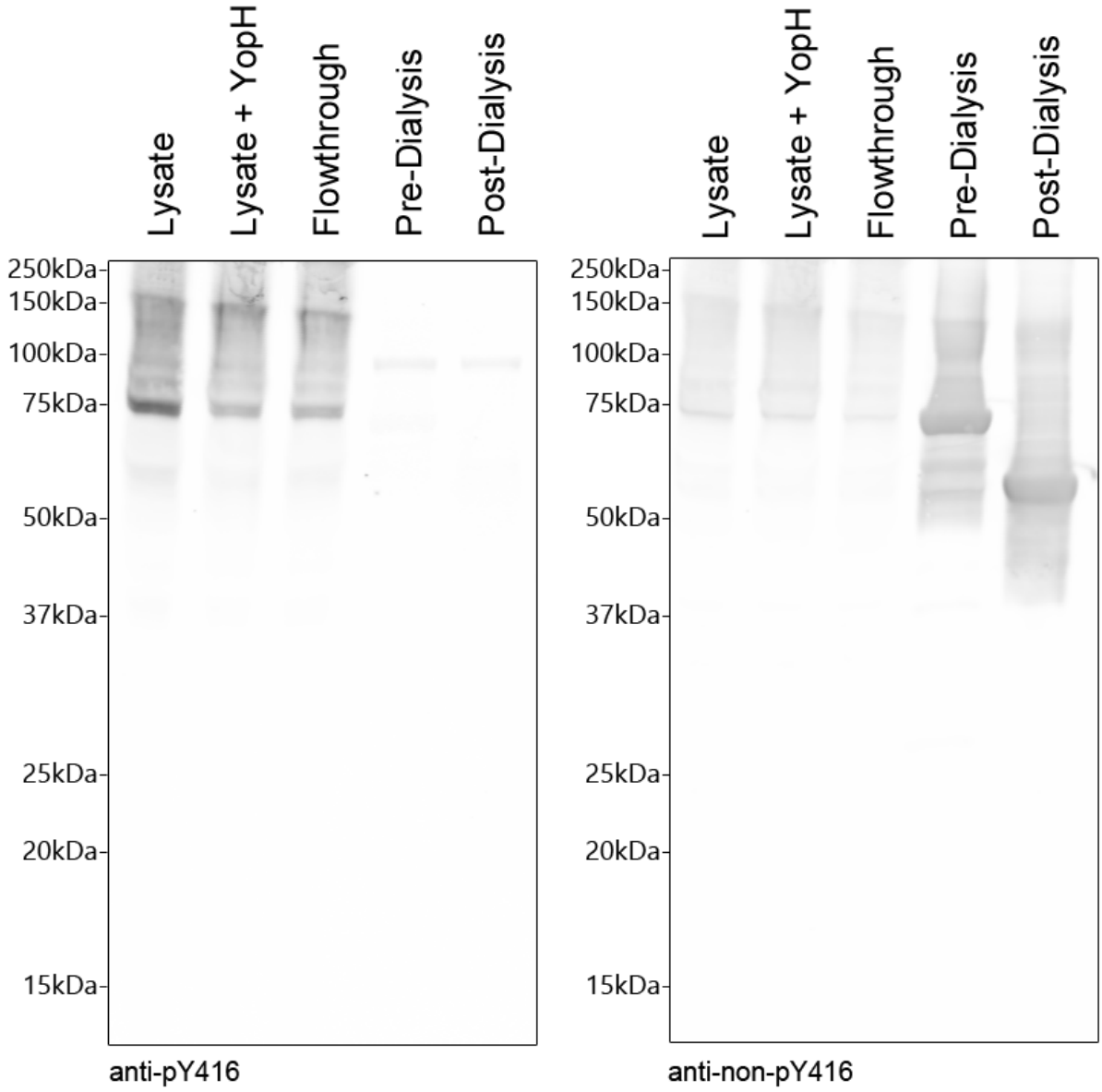


Figure S2. Purification of Src Cysteine Library
Protein was purified using standard Ni-NTA procedures with the inclusion of protease inhibitors in lysis buffer. Western blots show dephosphorylation of activation loop tyrosine 416.

Supplemental Figure 3

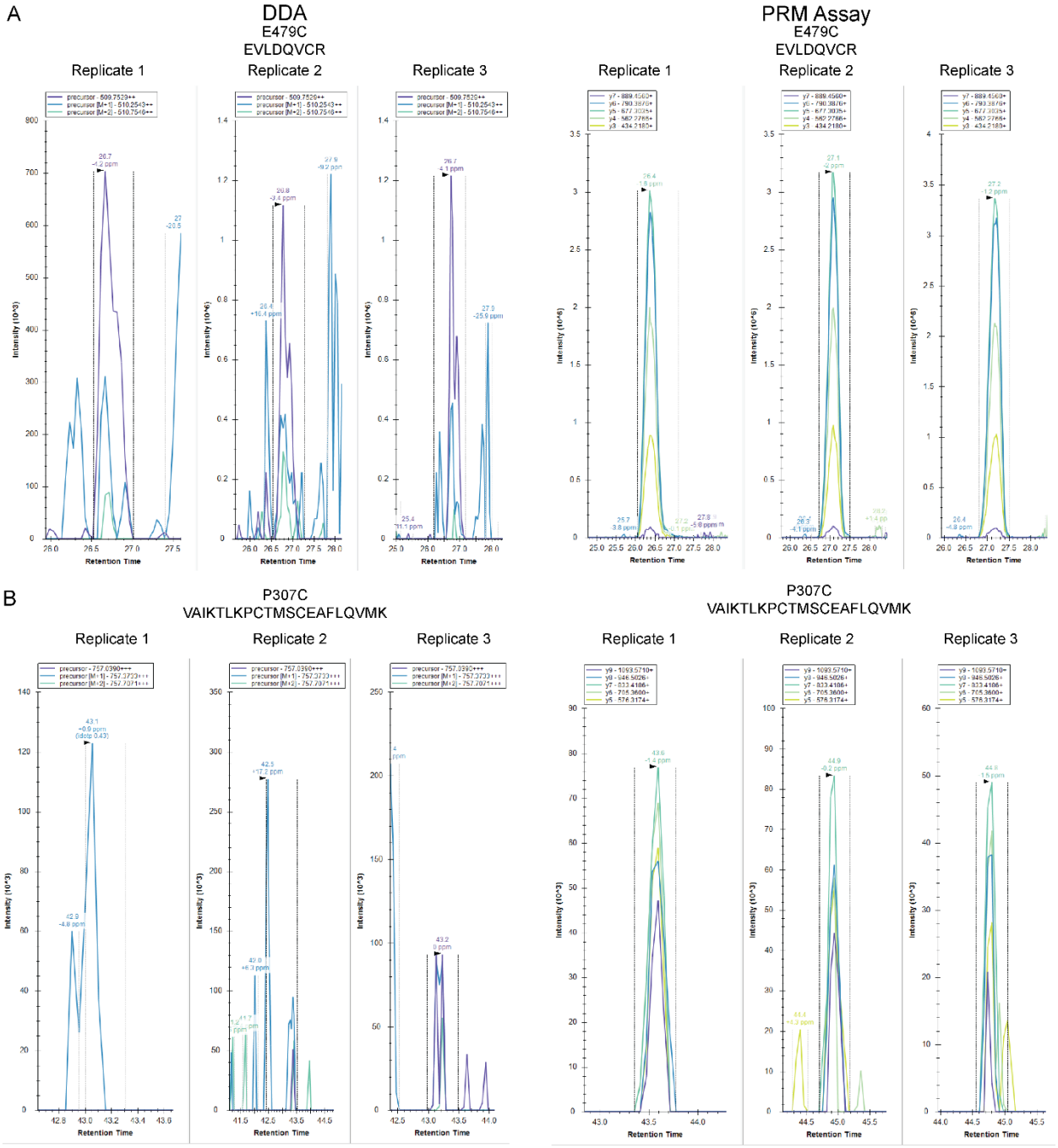


Figure S3. Comparison of extracted ion chromatograms from DDA and PRM experiments
Ion chromatograms extracted using Skyline (A) Parent/precursor masses of EVLDQVCR peptide from E479C Src mutant from DDA experiment (left) compared to traces for fragment ions of the same peptide from the PRM assay (right). (B) Same analysis for the VAIKTLKPCTMSCEAFLWVMK peptides from P307C mutant of Src.

Supplemental Figure 4

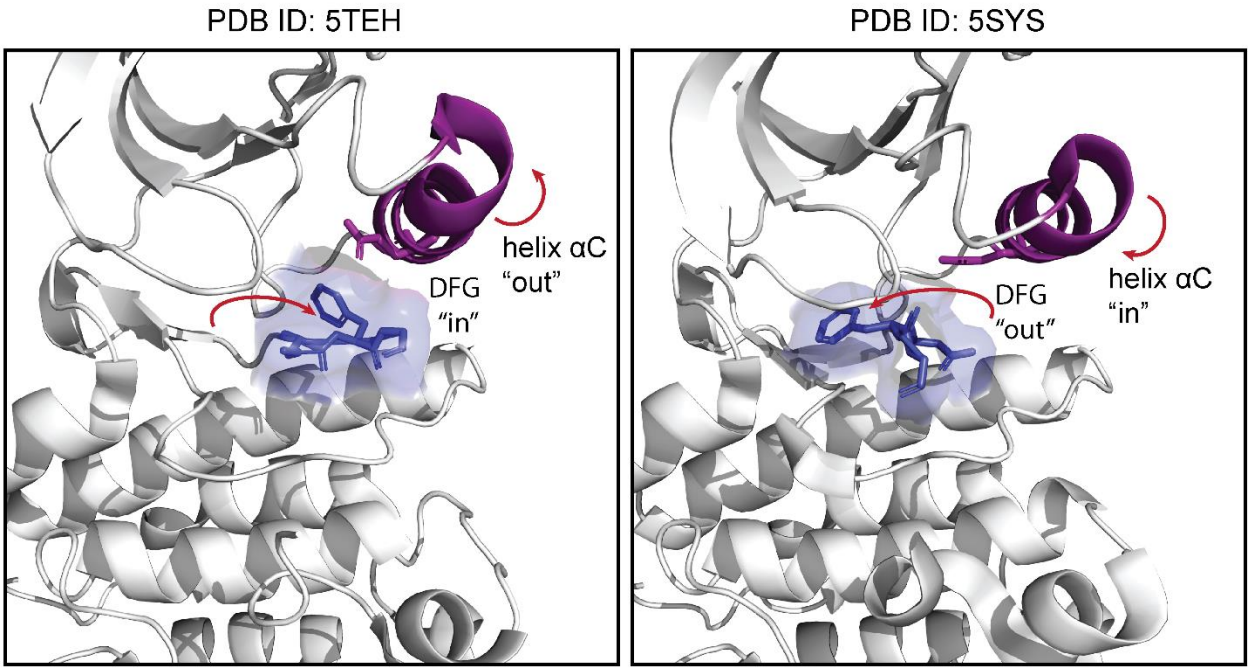


Figure S4. Active site conformations of Src stabilized by ATP-competitive conformation-selective inhibitors

Highlighted are the catalytically important helix αC and DFG-motif. The helix αC-out conformation is characterized by the rotation of the αC out of the active site and the DFG-motif in the active "in" conformation. Alternatively, the DFG-out conformation is characterized by the rotation of the helix αC into the active site, while the DFG-motif flips ~180° out of the active site.

Supplemental Figure 5

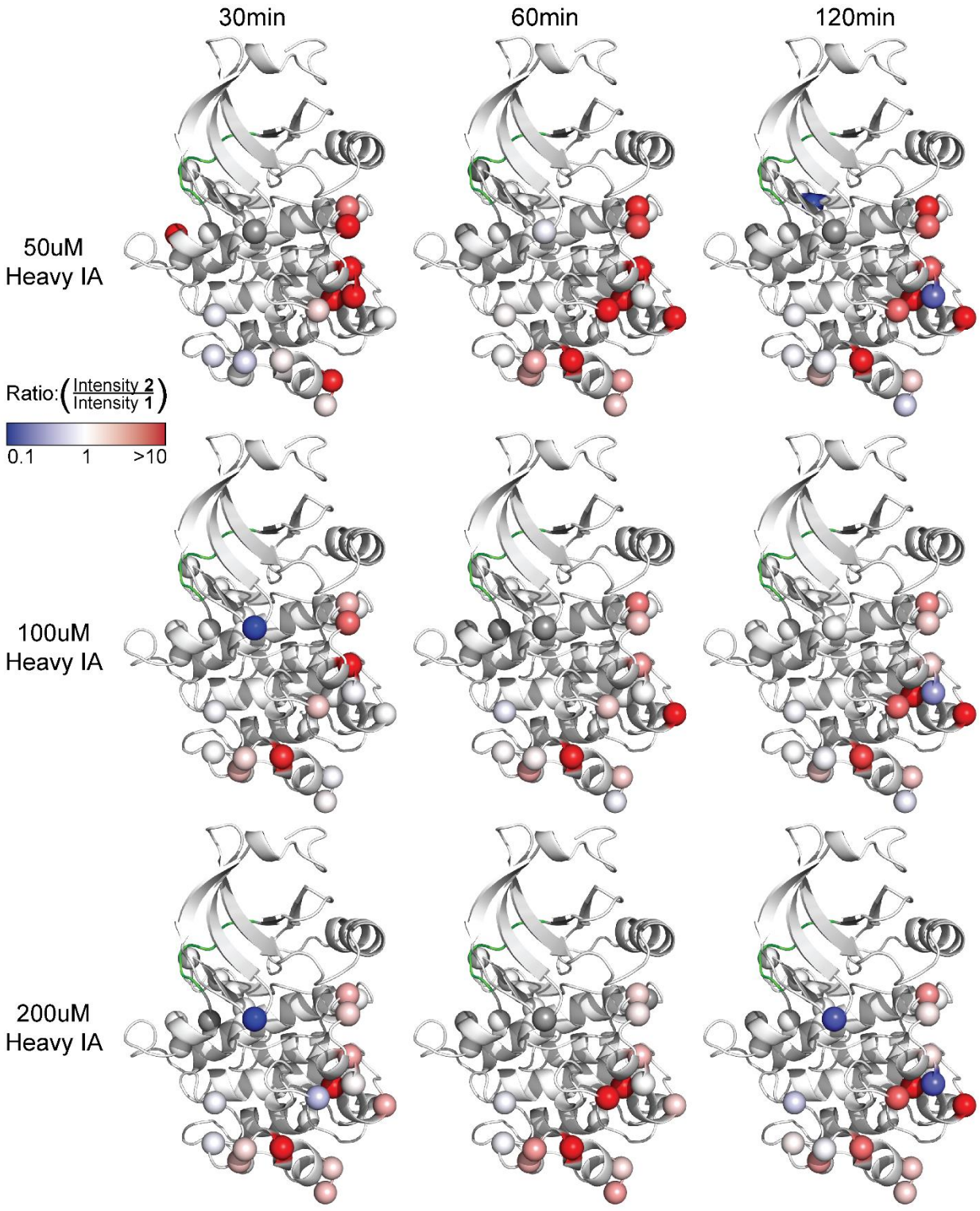


Figure S5. Parallel Chemoselective Profiling of conformation selective inhibitor bound Src
Complete dataset represented as ratios of heavy labeled peptides

Supplemental Figure 6

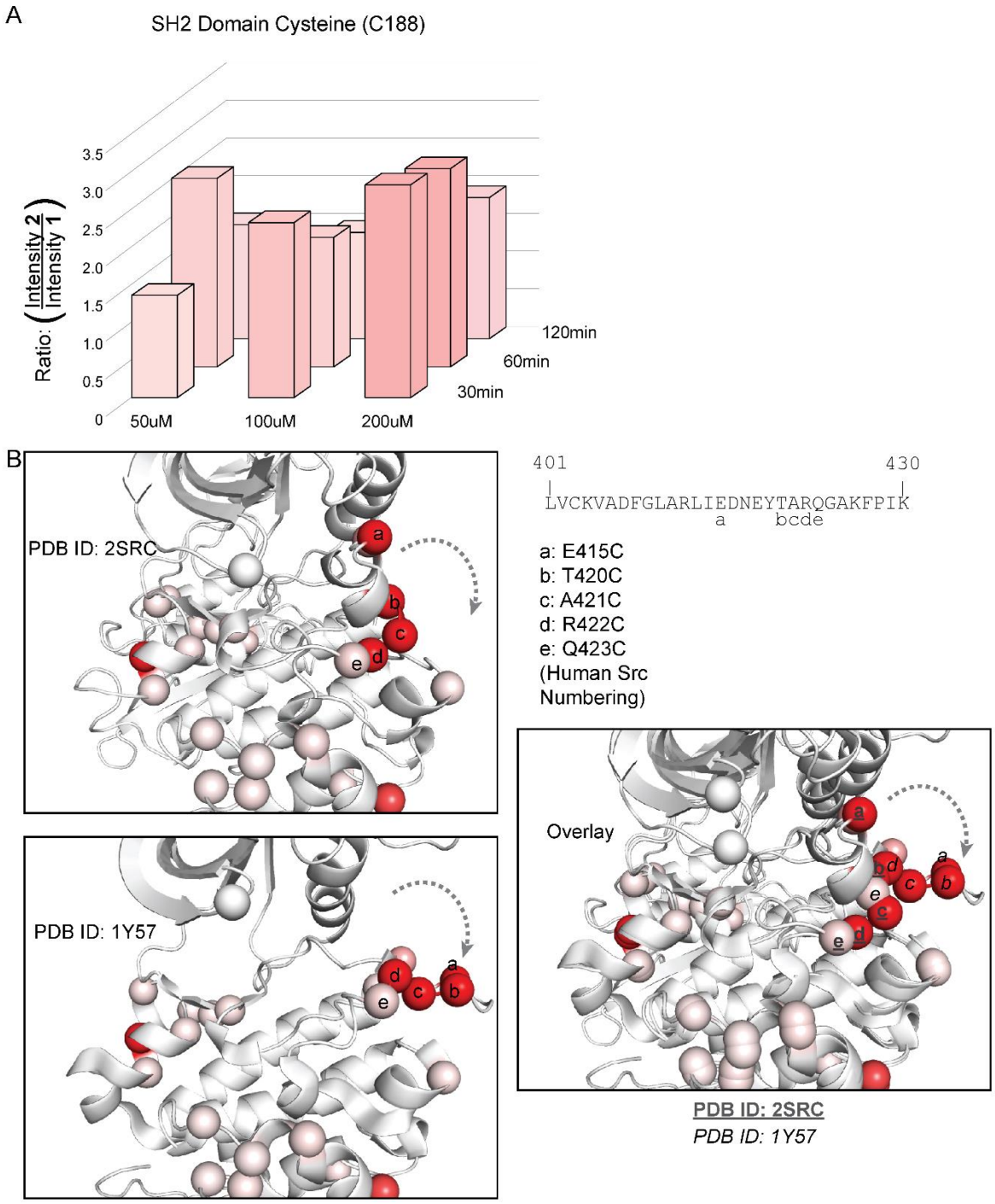


Figure S6. Movement of activation loop

(A) 50uM 60min data mapped to crystal structures PDB ID: 2Src and 1Y57. Indicated residues move out into solution in the open global conformation (PDB ID: 1Y57) relative to the closed global conformation (PDB ID: 2Src). (B) Quantification of WT wild-type cysteine residue C188 which residues on the SH2 domain.

Supplemental Figure 7

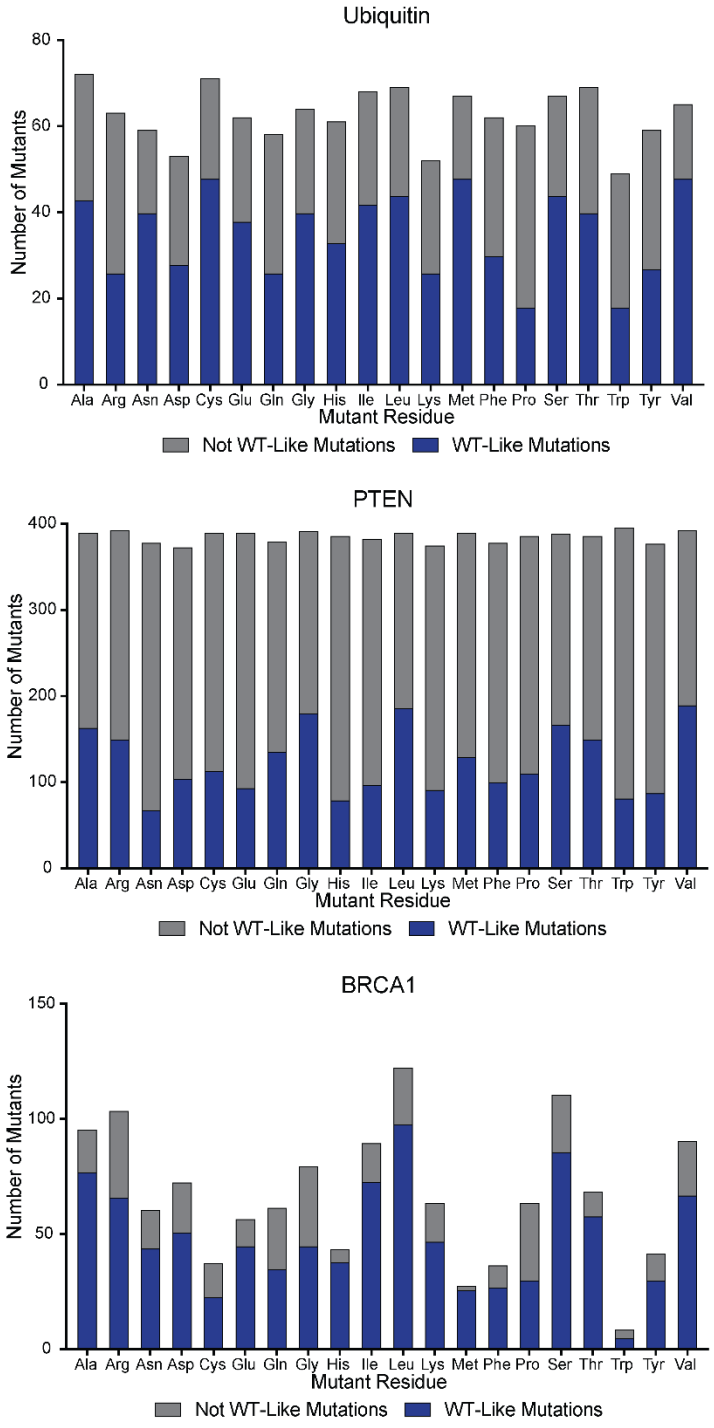


Figure S7. Distribution of all mutations contained in the Deep Mutational Scans analyzed
All mutational effect scores were extracted from the original datasets and binned as WT-like or not WT-like using the definition provided in the original reports (relative to synonymous distribution).

Chapter 2: An activity screen of Lck yields insight into molecular regulation and variants for enhanced T cell-based therapies

Introduction

Lymphocyte-specific protein tyrosine kinase (Lck) is a member of the Src-Family of Kinases (SFKs) that plays a central role in T cell function. Like all SFKs, Lck contains five domains—a kinase domain (KD), SH2 (phosphotyrosine binding), SH3 (polyproline motif binding), unique, and SH4 (membrane binding). Lck's kinase activity is regulated by multiple layers of regulation governed by inter- and intramolecular protein-protein interactions (PPIs), and phosphorylation events resulting in a highly tuned conformational equilibrium defined by two extremes (Figure 1Fig. 1A). When all regulatory domains are disengaged, and Lck is phosphorylated at its activation loop, it is most catalytically active and in the "open" global conformation. On the other hand, when the kinase is dephosphorylated at its activation loop, and phosphorylated at the C-terminal tail, the regulatory domains make physical contact with the kinase domain, perturbing the alignment of the key active site residues, abrogating kinase activity and the kinase is said to be in the "closed" global conformation.

T cells are essential members of the immune system's adaptive response to antigens presented by pathogens and tumors. T cells are rapidly activated by pathogen- or tumor-derived antigenic peptides presented by Major Histocompatibility Complex proteins (pMHC) and avoid autoreactivity with self-pMHCs. Although T cells are not activated by self-pMHC/T Cell Receptor (TCR) interactions, tonic signaling resulting from such interactions is required for the maturation and survival of naïve T cells. TCRs lack enzymatic activity, and extracellular detection of antigens in transduced intracellularly by the tyrosine kinases Lck and Zap70. Upon antigen engagement, Lck is recruited to the pMHC-engaged TCR where it phosphorylates the immunoreceptor tyrosine-based activation motifs (ITAMs) of the CD3ζ chains of the TCR complex. These Lck-mediated phosphorylation events lead to the recruitment and activation of Zap70. Zap70 binds to phosphorylated ITAM motifs using

its SH2 domain, after which Lck further phosphorylates Zap70. Lck then binds to phosphorylated and activated Zap70 using its SH2 domain, and binds to a polyproline-rich motif of the Linker for Activation of T Cells (LAT) using its SH3 domain (Lo, et al 2018). This interaction serves as a molecular bridge linking Zap70 to its substrate, LAT, facilitating efficient TCR signal transduction.

Results

Kinase Activity Measurements for 4,892 Lck Variants

To study the intramolecular regulation of Lck, we performed a deep mutational scan (DMS) (Fowler and Fields, 2014) of Lck's kinase domain—in the context of full-length enzyme—using a commercially synthesized variant library and the S. cerevisiae growth-based experimental workflow previously established for the closely related kinase Src (Ahler and Register, et al. 2019). Yeast provides a eukaryotic cellular environment devoid of extrinsic regulatory factors in which to profile the molecular regulation of tyrosine protein kinases. In this system, yeast growth correlates with kinase activity (Figure 1A). We commercially synthesized a DNA barcoded library of full-length Lck with the kinase domain mutagenized (Figure S1B), transformed this library into yeast, collected samples throughout outgrowth, and used next-generation sequencing to quantify the frequency of each barcode—and in turn, variant—over time (Figure 1B). Using our sequencing data, we calculated activity scores for the 4,892 single amino acid variants observed in our library (Figure 1C) and classified variants as wild-type like (WT-like), activating, or deactivating (Figure 1D and 1E). We generated a sequence-function map using the average activity score per-position for Lck (Figure 1G) which bared some expected trends. For example, most positions lining the ATP binding site are on average deactivating. However, like the DMS previously reported for Src, we observed patches of resides that are on average activating. We hypothesized that since most substitutions at these positions are on

average activating, they must participate in the negative regulation of Lck's kinase activity. Thus, we chose these regions to investigate in greater detail.

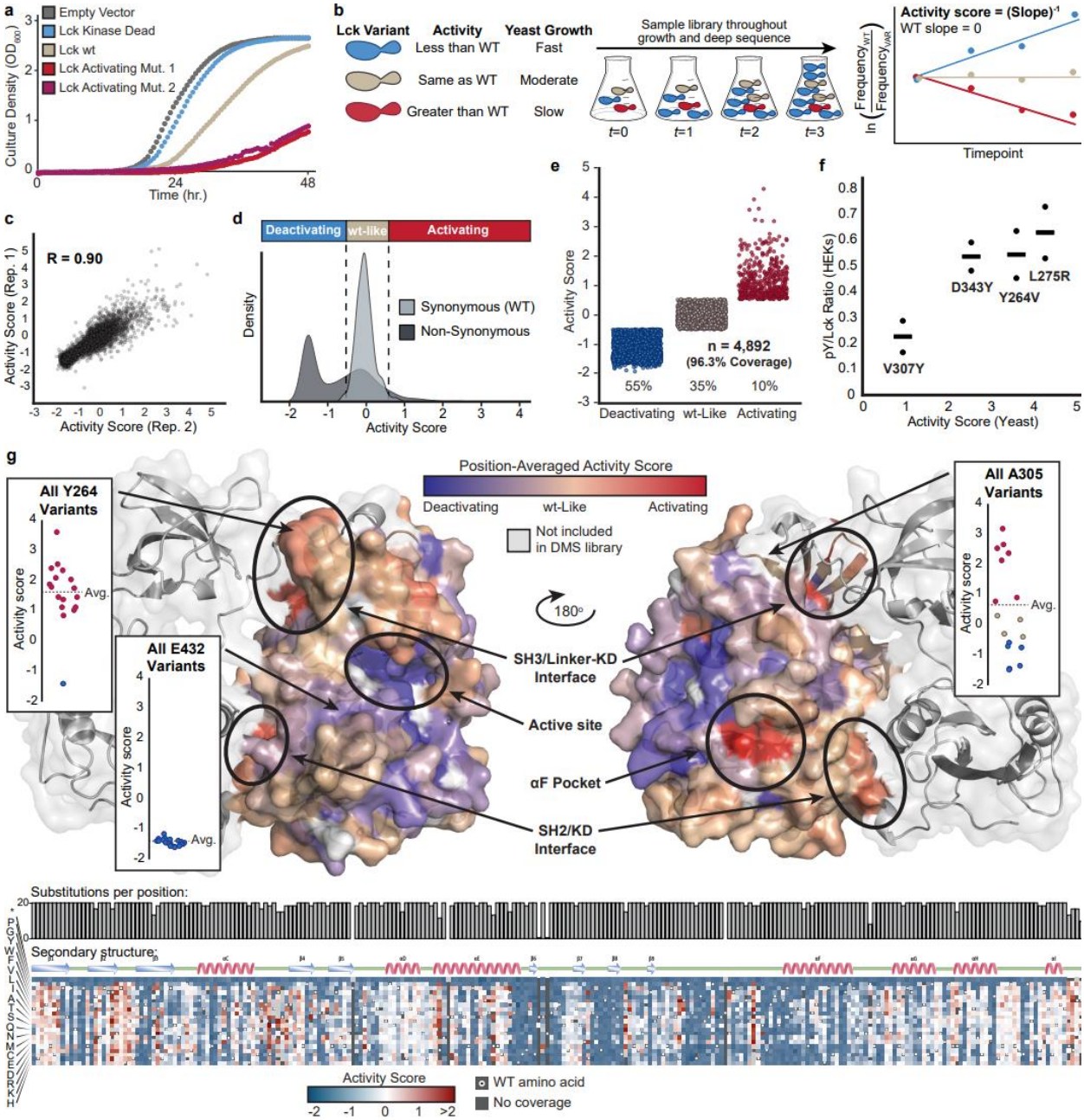


Figure 1. Kinase Activity Measurements of 4,892 Single Amino Acid Variants of Lck
(A) Yeast growth curves for select Lck variants expressed in yeast. Datapoints represent average of three replicates. (B) Schematic for yeast growth-based deep mutational scan of Lck's KD. (C) Scatterplot showing correlation between activity scores calculated from two independent biological replicates of the Lck variant library. (D) Histogram of synonymous and non-synonymous distributions of activity scores. Two standard deviations from average of synonymous distribution used to define Activating and Deactivating variants. (E) Activity scores for variants classified as Deactivating, wt-Like, or Activating. (F) Scatter plot of activity scores calculated in yeast as they correlate with

phosphotyrosine levels in HEK293T cells. (G) Position-averaged activity scores mapped onto the AlphaFold2 predicted structure of full-length Lck (N-terminal SH4 domain removed for simplicity). Bar graph represents mutational completeness at each amino acid position. Secondary structure and functional motif annotations were obtained from the ProKinO database.

Residue-Level Map of Lck's Intramolecular Regulation Reveals a Uniquely Tuned Regulatory Interface Between the SH2-SH3 and Kinase Domain

To investigate Lck's intramolecular regulation, we inferred that residues on the kinase domain (KD) which participate in autoinhibitory protein-protein interactions would exhibit multiple activating mutations. Therefore, we clustered the 34 residues on Lck's KD that had 5 or more activating substitutions using their coordinates in three-dimensional space extracted from the AlphaFold2 predicted structure of Lck (**Figure 2A**). Ten clusters emerged and were analyzed further.

All seven clusters previously identified in Src are represented in our Lck data (**Figure 2A** and **2B**). This includes cluster 9 (red), which defines the αF pocket of Lck, an orphan ligand binding pocket which interacts with the SH4 domain of the SFK to exert autoinhibition (Ahler et al., 2019). Lck's αF pocket was predicted to exist based on hyperactivating mutations observed at conserved residues (Chakraborty, et al 2022 bioRXiv). Two additional activating clusters, 3 and 8, lie along the SH2 and SH3 domain interface with Lck's KD. These clusters overlap with the clusters identified in the Src DMS (**Figure 2B**, Ahler et al., 2019) and represent known binding interfaces between SFK KDs and their regulatory SH2 and SH3 domains. The importance of the SH2-SH3/KD interface with respect to the regulation of SFK activity has long been appreciated, however the specific contributions of individual residues—particularly for Lck which lacks crystallographic data of a the SH3-SH2-KD module—is incomplete. Additionally, the heterogeneity of this interface amongst the SFKs—and the contribution of this heterogeneity to unique modes of autoinhibition and biological function—is virtually unexplored. We noticed that the size of cluster 3 (blue) and the nearby clusters 4 (orange) and

10 (yellow) appear to be more elaborated, or contain completely unique residues, in Lck (**Figure 2A**, **2B**, and **2C**).

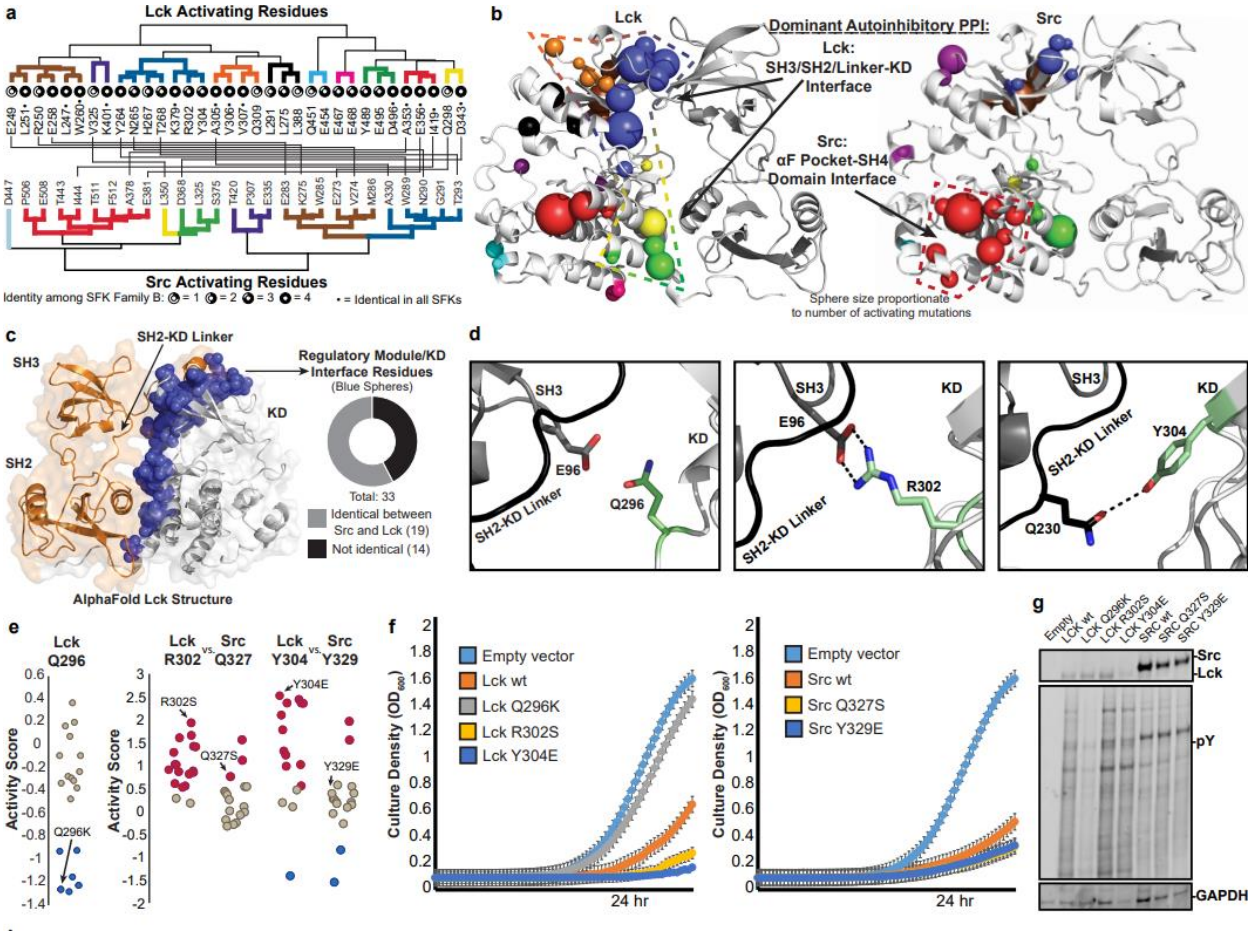


Figure 2. Characterization of Activity-Modulating Variants of Lck

(A) Hierarchical clustering of Lck residues with at least five activating mutations based on centroided atomic coordinates for each amino acid side chain extracted from the AlphaFold2 structure prediction of full-length Lck (AlphaFold: AF-P06239-F1) (top) compared to the hierarchical clustering of Src residues with five or more activating mutations (Ahler et al, 2019). Lines indicate corresponding residues shared between the two kinases. Pie charts represent identity between Src Family B members (Lck, Blk, Hck, Lyn). (B) Hierarchical clustering data projected onto the structures of Lck (AlphaFold: AF-P06239-F1) and Src (PDB: 2SRC). Sphere size represents the number of activating mutations per position, color from **Figure 2A**. (C) SH2-SH3-Linker contact residues on the kinase domain of full-length Lck taken from the AlphaFold2 predicted structure of Lck (AlphaFold: AF-P06239-F1). (D) Structural detail for residues along the SH2-SH3-Linker/KD interface and their predicted side chain interactions. (E) Activity scores for every variant at select residues on the SH2-SH3-Linker/KD interface. Src activity scores extracted from Ahler et al., 2019. (F) Individual yeast growth curves for Lck (left) or Src (right) variants expressed in ABC16 monster yeast. Dots represent average OD_{600} for three replicates, error bars represent SEM. (G) Phosphotyrosine western blot of lysates from yeast expressing individual Lck or Src variants. Representative of three replicates.

To investigate the relative importance of individual residues along the SH3-SH2/KD interface, we first identified three residues with divergent contributions to Lck's autoinhibition (**Figure 2D**). The AlphaFold2 structure of Lck shows residue Q296, located on the KD, directed towards residue E96 which lies on the SH3 domain (15D, left). Inspection of Q296 activity scores reveals that the lysine (and arginine) substitution are both deactivating (**Figure 2E**). We reasoned that perhaps a salt bridge forms between Q296K and E96, strengthening the autoinhibitory interaction between the KD and SH3 domain. Residues R302 and Y304 both exist on the KD and appear to make hydrogen bonding interactions with the SH3, or SH2-KD linker, respectively, based on the AlphaFold2 predicted structure (**Figure 2D** (middle and right, respectively)). Almost every substitution at these positions is activating, further indicating that specific sidechain interactions along this interface are important for maintaining Lck's typical autoinhibition. The most activating substitution at Lck R302 is only slightly activating in Src. Inspection of the crystal structure of autoinhibited Src (PDB ID: 2SRC) reveals no polar contacts between the sidechain of Q327 and any other residue, perhaps indicating that this specific residue does not engage in an autoinhibitory hydrogen bond as it appears to in Lck. Alternatively, Lck and Src are identical at Y304 (Src Y329) and both share an identical residue Q230 (Src Q254) on the SH2-KD linker (**Figure 2D**, right). Interestingly, however, substitutions at this position in Lck are strongly, and more frequently, activating as compared to Src (**Figure 2E**).

To test whether this region is uniquely activating in Lck and not in Src, we first individually expressed each variant of Lck and Src in yeast and measured their growth. As expected, yeast expressing Lck Q296K grew much faster than Lck wt expressing yeast (**Figure 2F**) and overall phosphotyrosine levels were correspondingly lower (**Figure 2G** and **S2B**). Importantly, western blot analysis revealed that Lck Q296K is expressed at about the same level as Lck wt in yeast indicating

that the increased yeast growth rate and lower phosphotyrosine levels are not simply due to protein degradation. Taken together, these results suggest that the lower activity of full-length Lck Q296K relative to Lck wt is due to the strengthening of the autoinhibitory SH2-SH3/KD interface, promoting the closed and catalytically down-regulated conformation of Lck. On the other hand, yeast expressing Lck R302S and Y304E grew much slower than yeast expressing Lck wt. Analysis of yeast growth rates revealed that yeast expressing Lck Y304E grow at a rate almost 80% slower than yeast expressing Lck wt, whereas Src Y329E expressing yeast grow at a rate about 20% slower than Src wt (**Figure S2A**), confirming that this variant is indeed more activating in Lck than the equivalent variant is in Src. One possible explanation for this difference between Lck and Src is that Lck's SH3-SH2/KD interface is of lower affinity compared to Src's, which allows for more rapid and transient activation—and subsequent deactivation—of Lck during T Cell activation.

Identification of Lck Variants Sensitized to a Panel of Six FDA-Approved ATP-Competitive Kinase Inhibitors

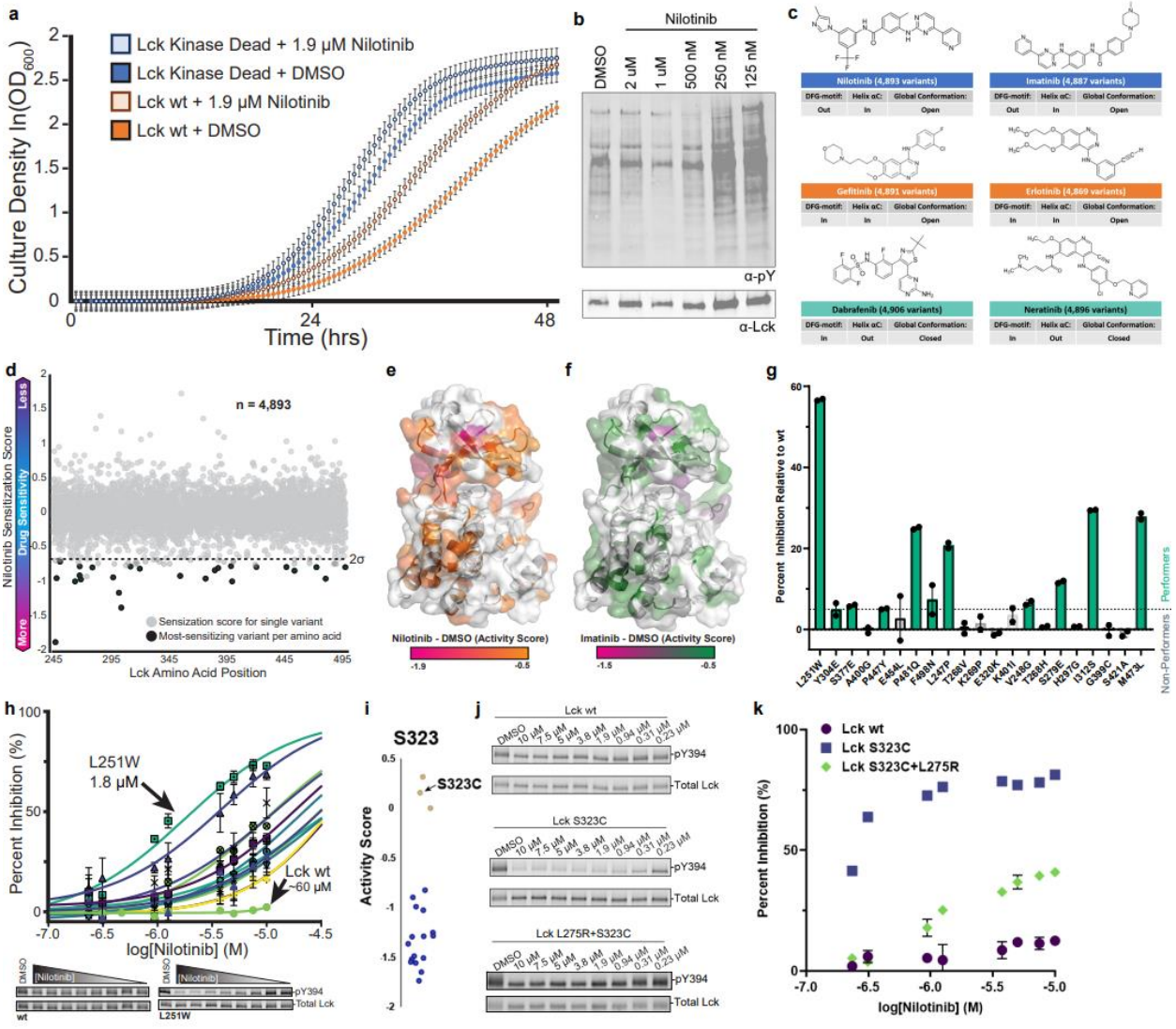
Systematic analysis of co-crystal structures of inhibitor-bound kinase complexes demonstrates that ATP-competitive inhibitors can stabilize large conformational movements of regions lining the ATP-binding pocket (Tong and Seeliger, 2015; Ung, et al., 2018; Zhao, et al., 2014). Specifically, certain pharmacophores stabilize large displacements of the DFG-motif within the activation loop, or the helix αC—which contains a conserved Glu residue that makes a salt bridge with the catalytic Lys residue—from their active conformations in numerous kinases (**Figure S4** from Chapter 1). Moreover, we and others have shown that conformation-selective inhibitors can stabilize large structural changes in regions of multi-domain protein kinases that are distal to the ATP-binding site (Aguis et al., 2019; Ahler et al., 2019; Chakraborty et al., 2019, Fang et al., 2020; Foda et al., 2015; Krishnamurty et al., 2013; Kung and Jura, 2016; Kwarcinski et al, 2016; Leonard et al., 2014; Potter et al., 2020; Register et al., 2014; Tong et al., 2017). In one of the most extensively characterized

examples, Src, conformation-selective inhibitors have been shown to modulate the level of intramolecular SH2 and SH3 domain engagement with the kinase domain by manipulating the same allosteric networks that control kinase activity.

The ATP-competitive kinase inhibitor, dasatinib—originally developed to target breakpoint cluster region-abelson murine leukemia viral oncogene homolog 1 (BCR-ABL) fusion protein, but which also potently inhibits SFKs, including Lck—has been shown to reversibly inhibit T cell receptor signaling and proliferation via ATP-competitive inhibition of Lck (Schade, A. E., 2008, *Blood*). Recently, dasatinib has been advanced into the clinic as a pharmacologic off switch for CAR-T cells (Weber et al, 2019; Mestermann, et al 2019). However, several clinically relevant side effects could present complications for patients chronically treated with high doses of dasatinib (Talpaz, et al 2006). Therefore, to maximize the utility and potential safety of Lck-mediated off switches for T cell-based therapies, we initially identified FDA-approved ATP-competitive kinase inhibitors which met three criteria: those with desirable patient safety profiles for multi-year treatment, those with relatively selective off-target profiles with <10 off targets with IC_{50}s <2 μM, and those that had Lck as an off-target. Initially, this led us to two inhibitors—nilotinib and imatinib—which have been used in the clinic and chronically dosed in patients for more than a decade without significant toxicity. We focused on these two inhibitors initially because both compounds stabilize the DFG-out, αC helix-in active site conformation of their target BCR-Abl, which in SFKs would stabilize the open global conformation. We speculated the stabilization of this conformation would allow an inhibitor-sensitized Lck to bind intramolecular scaffolding partners essential for TCR signaling (Lo, et al 2018) and could enhance Lck's ability to function as an off switch for TCR signaling.

Inhibitor screens in yeast can be difficult due to the rapid efflux of cytosolic inhibitor via ATP-binding cassette (ABC) transporters. However, we have found that we can achieve sufficiently high concentrations of kinase inhibitors using a genetically engineered ABC transporter deletion strain

(ABC16 monsters) (Suzuki et al, 2011). Recently, we used this strain of yeast along with a Src variant library to profile Src's resistance to various ATP-competitive inhibitors (Chakraborty, et al. 2022).



From these screens, we were able to identify general and inhibitor-specific mechanisms of resistance. We reasoned that we could employ a similar workflow to identify mutations that sensitize Lck to a panel of six FDA-approved kinase inhibitors.

Figure 3. Screening of 4,892 Single Amino Acid Lck Variants for Drug Sensitization
(A) Yeast growth curves Lck wt or K273R (kinase dead) expressing ABC16 yeast in the presence of 1.9 μM nilotinib or DMSO. (B) Phosphotyrosine blot for yeast expressing Lck wt grown in the presence of nilotinib or DMSO. (C) Summary of inhibitors used in this study for sensitization screens against the Lck variant library. (D) Sensitivity scores for all variants screened against nilotinib. Dotted line represents two standard deviations from the mean of sensitization scores for synonymous mutations. (E) Top 50 most nilotinib-sensitizing variants mapped onto the crystal structure of Lck

(PDB: 3AD6). (F) Top 50 most imatinib-sensitizing variants mapped onto the crystal structure of Lck (PDB: 3AD6). (G) Percent inhibition calculated from activation loop phosphotyrosine 394 western blots for 21 single amino acid variants. 4 hr treatment with 5 μM nilotinib. Dots represent individual values (n=2). (H) Nilotinib titrations for most nilotinib-sensitized variants. (n=2). (I) All activity scores for position S323. (J) Activation loop phosphotyrosine 394 levels for Lck wt, S323C, and L275R+S323C double mutant following 4 hr treatment with indicated neratinib concentration. Representative of two independent replicates. (K) Graphical representation of neratinib titration data from **Figure 3J** (n=2).

The first step in adapting our yeast growth-based activity assay to identify sensitizing mutations was to ensure that inhibitor treatment dose-dependently rescued yeast growth. We first treated yeast expressing Lck wt or Lck K273R (kinase dead) with or without 1.9 μM nilotinib and measured their growth (**Figure 3A**). We found that drug treatment increased the growth rate of Lck wt expressing yeast, but not Lck K273R expressing yeast, indicating that the effect of inhibitor treatment was specific to inhibition of Lck's kinase activity. We then used western blot analysis of inhibitor-treated yeast to measure phosphotyrosine levels (**Figure 3B** and **S3A**). We identified a concentration of each drug that inhibited about 10% of Lck's kinase activity (data not shown).

Sensitizing Mutations are Distributed Throughout the Kinase Domain and Impart Sensitivity by Stabilizing Active Site Features

We then transformed our Lck variant library into yeast, treated at a single concentration of each inhibitor, and collected samples during growth for analysis. We performed this screen on our panel of inhibitors that stabilize complementary active site conformations. (**Figure 3C** and **S3H**). We first determined activity scores for DMSO and drug treated samples, and then calculated "sensitivity scores" by taking the difference between DMSO and drug treated activity scores representing faster yeast growth relative to wt Lck in the presence of inhibitor (**Figure 3D**). To our surprise, we discovered sensitizing mutations distributed throughout the kinase domain, some more than 20 Å from the ATP-binding site (**Figure 3E** and **3F**) and principal component analysis of our sensitivity

scores revealed that mutations that stabilize certain active site conformations appear to correlate with one another (**Figure S3I**).

Several of the most sensitizing mutations are on the N-terminal lobe of the kinase domain in regions that impart inhibitor resistance in Src (Chakraborty, et al 2022). Using amino acid substitution in Pymol and the crystal structure of Lck's KD bound to imatinib, we were able to speculate why certain mutations are sensitizing for some inhibitors and not others (**Figure S3D** and **S3E**). In particular, the Leu385His mutation appears to only be sensitizing to inhibitors that flip the DFG-motif 180º out of the active site (**Figure S3D** and **S3E**). Given Leu385's proximity to the DFG-motif, and its likely coordinated movement with the DFG-motif, we speculate that the Leu385His mutation could thermodynamically favor the DFG-out conformation, lowering the energy barrier for inhibitor binding and increasing the binding affinity. Similarly, Ser377 is a residue located at the base of the loop that connects the helix αC to the kinase domain. Ser377Glu is sensitizing for inhibitors that stabilize the helix αC-in conformation, with the exception of erlotinib (**Figure S3F** and **S3G**).

To validate our inhibitor sensitivity data, we primarily focused on nilotinib-sensitive mutants. We selected a panel of 21 mutants using the nilotinib sensitivity scores by identifying the most sensitizing variants 2 standard deviations below the average nilotinib sensitivity score (**Figure 3D**, black dots). We first transfected a panel of mutants into HEK293T cells and treated with 5 μM nilotinib for 4 hr. We used western blot analysis to quantify activation loop phospho-tyrosine 394 levels as a proxy for kinase activity. We found that more than half of our chosen panel of mutants showed increased levels of inhibition relative to Lck wt (**Figure 3G** and **S3B**). We further calculated IC_{50}s for these variants in HEK293T cells (**Figure 3H**) and found that even single mutations were enough to increase nilotinib sensitivity by about 30 fold.

We also analyzed the activity scores for Ser323, which is near the active site of Lck. This position is equivalent to the cysteine that is covalently targeted by the FDA-approved drug neratinib.

We found that this substitution was classified as WT-like (**Figure 3I**) and that HEK293T cells expressing this variant of Lck had similar phospho-tyrosine 394 levels (**Figure 3J** (top and middle blots)). When treated with varying concentrations of neratinib, HEK293T cells expressing Lck Ser323Cys showed dose-dependent inhibition of phosphotyrosine 394 levels, while Lck wt expressing cells showed virtually no inhibition. We also combined the Ser323Cys mutation with the Leu275Arg mutation which is the most activating substitution identified in our initial deep mutational scan of Lck. We found that this double mutant did demonstrate higher phsopho-tyrosine 394 levels compared to Lck wt when treated with DMSO (**Figure 3J** (top and bottom)) and also demonstrated dose-dependent inhibition of phospho-tyrosine 394 levels (**Figure 3J** and **3K**).

Supplemental Information for Chapter 2

Supplemental Figure 1

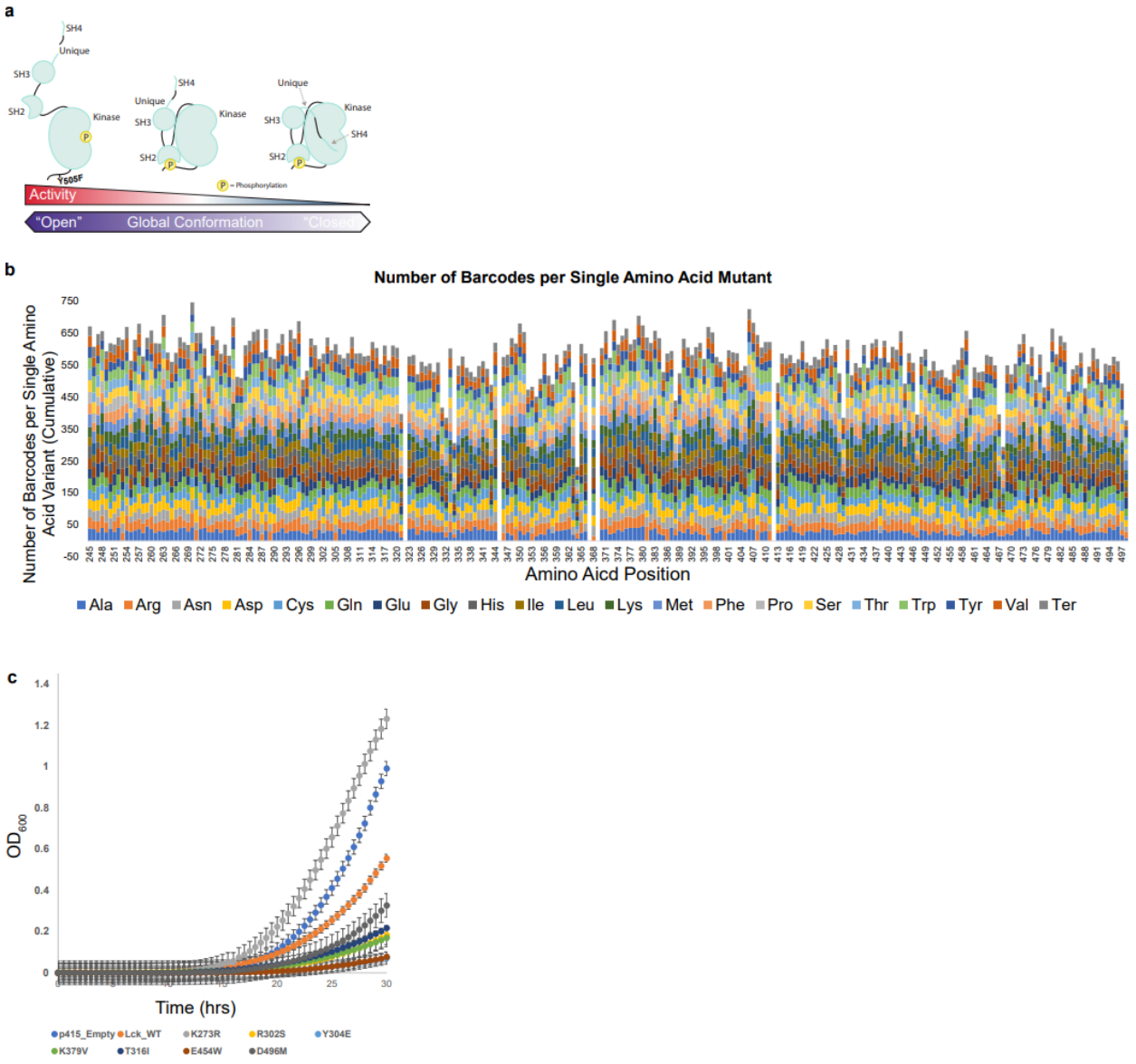


Figure S1. Profiling of Lck's Phosphotransferase Activity in Yeast
(A) Conformational equilibrium of Src Family Kinases and corresponding phosphotransferase activities. (B) Summary of the PacBio sequencing data of the Lck variant library. Stacked bars represent the number of barcodes at each position associated with each amino acid substitution at that position. (C) Individual yeast growth curves for selected Lck variants in ABC16 monster yeast. Dots represent average of three replicates, bars represent SEM.

Supplemental Figure 2

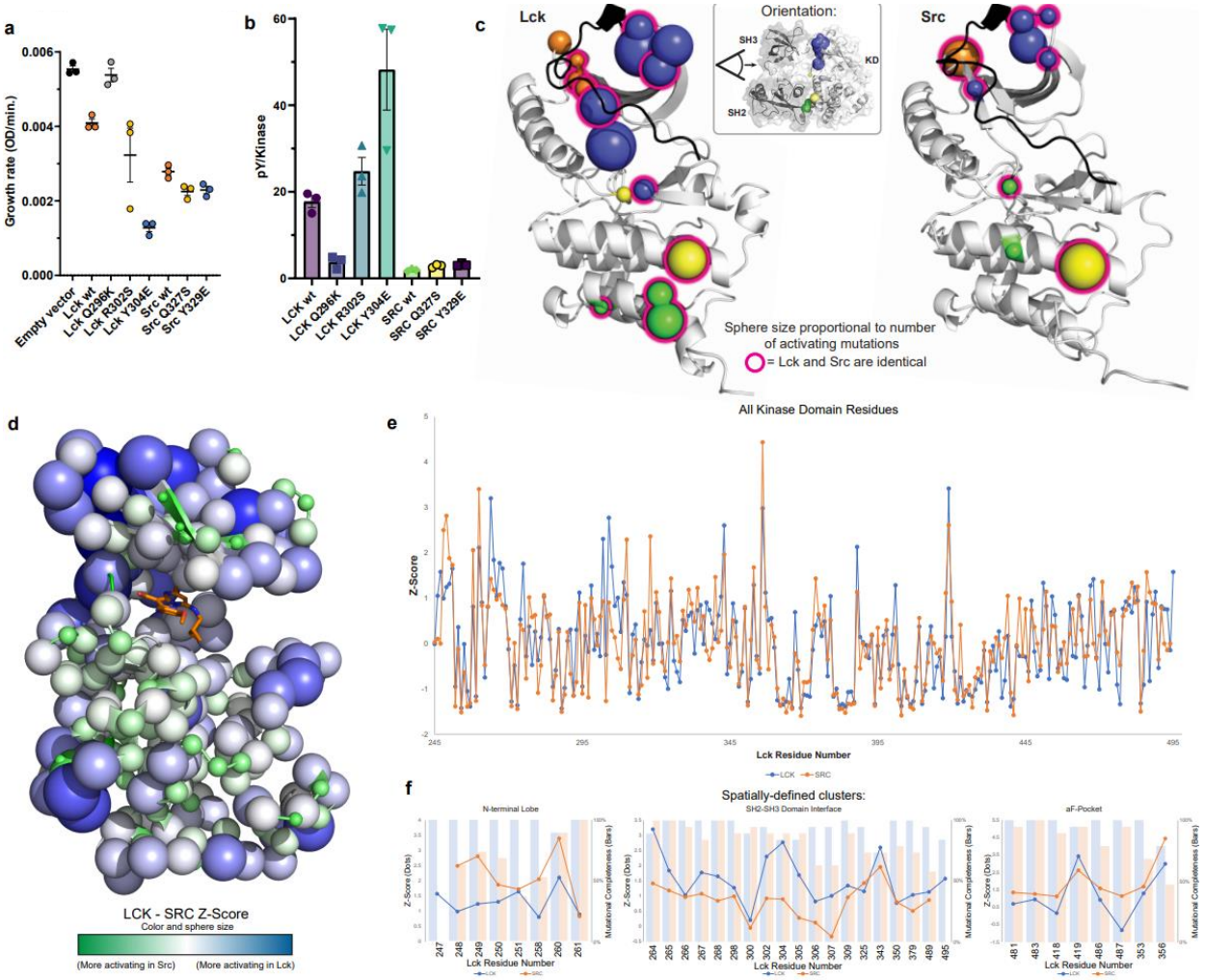


Figure S2. Comparison of Lck and Src Activity Measurements

(A) Quantification of individually assessed yeast growth rates for indicated Lck or Src variants in ABD16 monster yeast. Relates to **Figure 2F**. Dots represent individual replicates (n=3) lines indicate averages, error bars represent SEM. (B) Quantification of phospho-tyrosine levels in yeast expressing indicated Lck or Src variants. Relates to **Figure 2G**. Dots represent individual replicates (n=3). Error bars represent mean. Error bars represent SEM. (C) SH2-SH3-Linker/KD Interface of Src (PDB: 2SRC) and Lck (AlphaFold: AF-P06239-F1). Spheres represent number of activating mutations per position. Color represents classification from hierarchical clustering (**Figure 2A** and **2B)**. (D) Quantitative comparison of the average activity score per position between Lck and Src activity data (Ahler et al. 2019). Z-Scores were calculated for each average activity score within each dataset. The difference between z-scores for each corresponding position on the kinase domain was taken (Lck-Src) and the magnitude of this difference is represented by sphere size and color. Smaller, more green spheres represent residues that are more activating in Src than in Lck. Larger more blue spheres are more activating in Lck than in Src. Left side of figure where blue spheres appear is the SH2-SH3-Linker/KD interface. (E) Z-scores for all kinase domain residues of Src and Lck. (F) Graphical representation of z-scores taken from hierarchical clustering. Higher z-scores represent residues that

were on average more activating. Certain regions are more activating in Src (N-terminal lobe and αF-pocket) whereas the SH2-SH3-Linker/KD interface is more activating in Lck than in Src.

Supplemental Figure 3

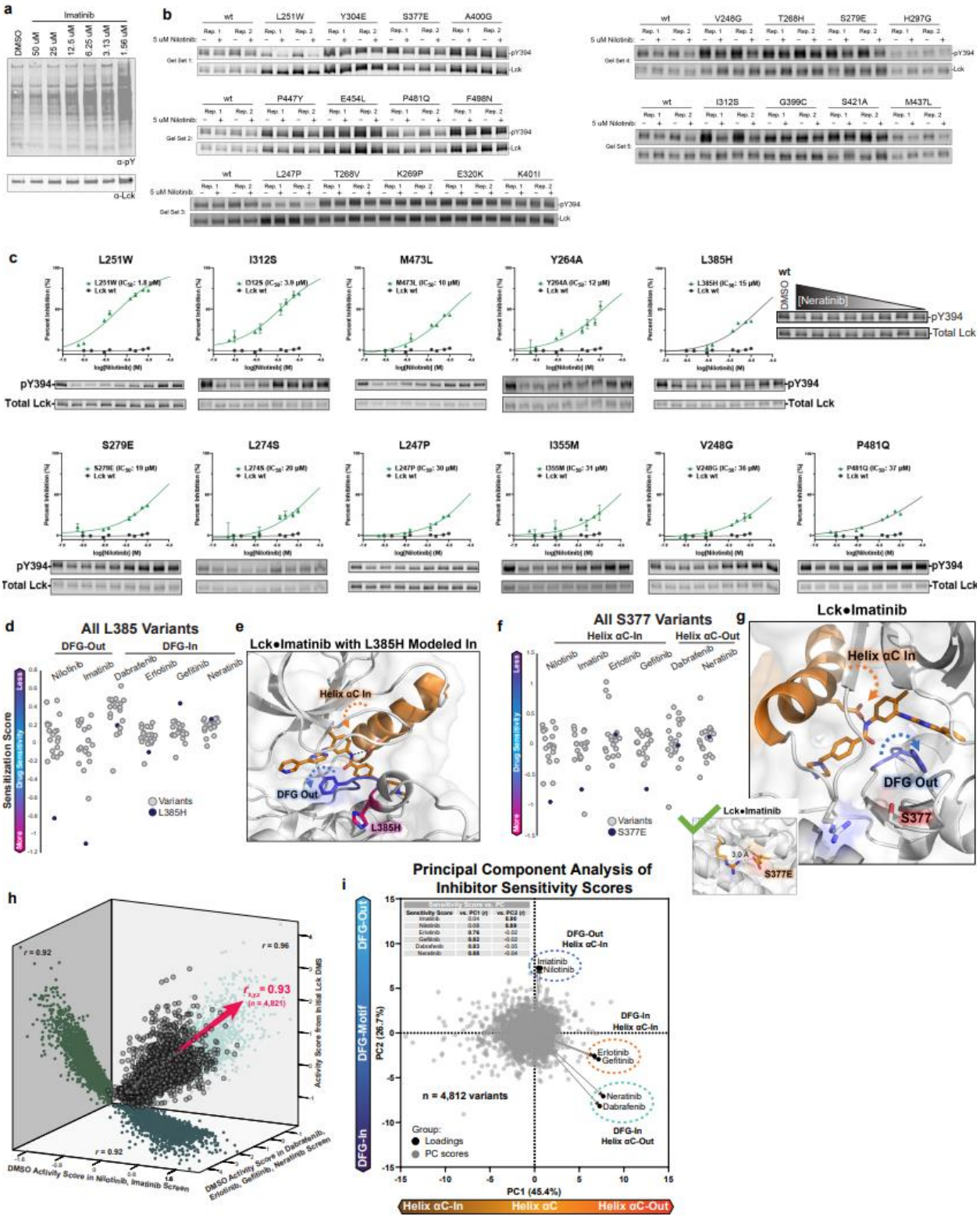


Figure S3. Characterization of Drug-Sensitizing Lck Variants

(A) Western blot of Lck wt expressing ABC16 monster yeast treated with indicated concentrations of imatinib. (B) Western blots of initial nilotinib-sensitized variants individually expressed in HEK293T

cells treated at 5 μM or DMSO for 4 hr (n=2). Relates to **Figure 3G**. (C) Nilotinib titrations for individual nilotinib-sensitized variants that showed increased levels of inhibition relative to Lck wt when treated with a single concentration. Relates to **Figure 3H**. Representative of two replicates. (D) All sensitization scores for residue Leu385. (E) Structural details of Lck's KD bound to imatinib (PDB: 2PL0) with Leu385His modeled using PyMol mutagenesis. (F) All sensitization scores for Ser377. (G) Structural detail of Lck bound to imatinib (PDB: 2PL0) with Ser377Glu modeled using PyMol mutagenesis (bottom). Dotted lines represent polar contacts identified by PyMol following mutagenesis. (H) Three-dimensional correlation coefficient between activity scores calculated for DMSO replicates of sensitization screen and the initial DMS of Lck in the absence of drugs. (I) Principal component analysis of inhibitors sensitization data. Performed using GraphPad Prism8.

Materials and Methods

Yeast growth assay

Yeast codon-optimized full-length Lck and indicated mutants were transformed into *S. cerevisiae* ABC16 monster yeast using LiAc transformation (Geitz and Schiestl, 2007) and plated on C-Leu plates. Three colonies for each Lck variant were used to inoculate 1 mL overnight liquid cultures in C-Leu media with 3% glucose. Strains were back diluted into C-Leu media supplemented with 3% raffinose and were allowed to double at least once. 150 µL of each strain at OD 0.01 in C-Leu media supplemented with 3% galactose were plated per well of a 96 well plate in a BioTek Synergy plate reader under constant orbital shaking at 30 °C. OD_{600} was measured every 30 min over 48 hr.

Yeast western blot protocol

Picked freshly transformed colonies (~1 week old). Grew in 1mL C-Leu 2% glucose overnight at 30 °C with shaking at 250 RPM. Next day, cells were centrifuged at 3,000 g for 5 min in culture tubes, and the media was removed. Cells were resuspended in 5 mL C-Leu 3% raffinose and allowed to double at least once. Cells were pelleted at 3,000 g for 5 min, media was removed, and cells were resuspended in 5 mL C-Leu with 2% galactose. 24 hr later, 100 µL aliquots of cells were centrifuged at 3,000 g for 5 min in microcentrifuge tubes and the media was removed. Cells were stored at -80 °C until further analysis. Cell pellets were thawed on ice and lysed with 200 µL SUMEB Buffer (1% SDS, 8 M urea, 10 mM MOPS, 10 mM EDTA, 0.01% bromophenol blue, 5% BME, 1 mM PMSF, pH 6.8, 1X yeast protease inhibitor (Sigma), 1X phosphatase inhibitor cocktails 2 and 3 (Sigma)). 150 µL glass beads (0.4-0.6 mm) were added to each sample and incubate at ~1,600 rpm for 10 min at room temperature, then an additional 10 min at 65 °C. Lysate was transferred to a new microcentrifuge tube and centrifuged at 17,000 g at 4 °C for 5 min. Samples were then resolved on SDS-PAGE gels (Criterion, BioRad) and transferred to nitrocellulose membranes (TransblotTurbo, BioRad). Western

blots were performed using the indicated primary antibodies and developed using LiCor secondary antibodies.

Lck Variant Library

Lck single amino acid variant library was commercially synthesized by Twist Biosciences (South San Francisco, CA) and subassembled with random 18 nucleotide barcode into the p415 vector backbone. Synonymous codons were used where possible, and stop codons were included. Following subassembly, the pooled, barcoded variant library was transformed into *E. coli* and serial dilutions were used to estimate library size. Pools of diluted library were collected and analyzed by Illumina sequencing for enumerate library size aiming for an average of 30 barcodes per variant (~150,000 barcodes total).

Lck variant library transformation into yeast

Performed following the protocol from Ahler et al 2019. Briefly, a single colony of ABC16 monster yeast was picked and grown in 5 mL YPED overnight at 30 °C with shaking. The saturated culture was diluted into 50 mL YPED at an initial OD_{600} of 0.3 and grown until reaching an OD_{600} of ~2. The culture was centrifuged at 3,000 g for 5 min, the cells were washed with 25 mL sterile H_2O three times and distributed among 10 microcentrifuge tubes following the last wash. Cells were pelleted at 17,000 g for 30 sec and resuspended in transformation mix containing 240 µL 50% PEG, 50 µL 2 mg/mL salmon sperm (Invitrogen), 36 µL 1 M LiAc. Eight of the samples had 500 ng Lck variant library added, one had 250 ng Lck variant library and 250 ng pRS411 plasmid, and one received only H_2O without plasmid. All samples were incubated at 42 °C for 45 min after which cells were pelleted and washed with 1 mL sterile H_2O. The eight library samples were pooled, added to 50 mL C-Leu 2% glucose media and grown with shaking at 30 °C. Serial dilutions of the culture were plated on C-Leu plates to estimate library size. Dilutions of the sample that received the Lck variant library

and the pRS411 plasmid were plated on C-Leu-Met plates to estimate the number of double transformants. After 72 hr outgrowth, aliquots of the library were frozen at -80 °C in 20% glycerol.

Selection and sequencing

The protocol from Ahler et al 2019 was followed with modifications made for drug screens. Briefly, aliquots of the frozen Lck variant library were thawed and grown overnight in C-Leu with 2% glucose. The next morning, the cultures were back diluted to OD_{600} 0.5 in C-Leu with 3% raffinose and allowed to double at least once. Selection was initiated by inoculation at OD_{600} 0.01 in 250 mL C-Leu with 2% galactose. Samples of pre-selected cultures were saved and served as timepoint = 0. Cells grew overnight and timepoints were collected at OD_{600}, representing approximately 4, 6 and 8 doublings. At each timepoint, 4 OD units of cells were pelleted and stored at -80 °C until further analysis. For drug selections, desired drug concentrations were added when cultures were back diluted to OD_{600} 0.01.

Lck variant library growth assay analysis

Illumina reads were demultiplexed using bcl2fastq and ea-utils. Activity scores for Lck variants were calculated using Enrich2 (Rubin et al., 2017). Barcoded Variant SeqLib analysis was used with "weighted least squares" scoring option enabled with "wild type" normalization. Additionally, FASTQ mode parameters included an average quality of 30, and a minimum count of 10. The OD_{600} readings of each timepoint were multiplied by 100 and used as the timepoints. The barcode-variant map from the subassembly was used.

HEK293T drug treatments

500µL HEK293T cells at $0.1x10^6$ cells/mL were plated in a 24 well tissue culture plate. 24 hr later, cells were transfected with Lck variants. Transfection mix was made using 250 ng plasmid DNA, 37 µL serum free DMEM, and 0.75 µL PEI (1 mg/mL). Transfection mixes were incubated at room

temperature for 15 min before being added dropwise to cells. 24 hr later, indicated drug at indicated concentration (or DMSO) was added. Cells were incubated in a humidified incubator at 37 °C with 5% CO_2 for 4 hr. Media was the aspirated, cells were washed once with ice cold PBS, and then cells were lysed directly in 300 µL 1.8X SDS-PAGE loading dye (5X Loading dye: 250 mM Tris, pH 6.8, 45% glycerol, 5% SDS (wt/vol), 0.5 M DTT, 1.86 mM Bromophenol blue). Samples were then collected into microcentrifuge tubes and boiled at 98 °C for 10 min before being resolved via SDS-PAGE (4 µL sample loaded per well) (Criterion gels, BioRad) and transferred to nitrocellulose membranes (TransblotTurbo, BioRad). Westerns were probed with the corresponding primary antibodies, and then developed using LiCor secondary antibodies.

WORKS CITED

Abo M, Li C, Weerapana E. (2018). Isotopically-labeled iodoacetamide-alkyne probes for quantitative cysteine-reactivity profiling. Mol Pharm. 15, 743–749.

Agius MP, Ko KS, Johnson TK, Kwarcinski FE, Phadke S, Lachacz EJ, Soellner MB (2019). Selective proteolysis to study the global conformation and regulatory mechanisms of c-Src kinase. ACS Chem Biol 14 7, 1556–1563.

Ahler E Register AC, Chakraborty S, Fang L, Dieter EM, Sitko KA, Vidadala RSR, Trevillian BM, Golkowski M, Gelman H, Stephany JJ, Rubin AF, Merritt EA, Fowler DM, Maly DJ (2019). A combined approach reveals a regulatory mechanism coupling Src's kinase activity, localization and phosphotransferase-independent functions. Mol. Cell74, 2, 393–408

Akamine P, Madhusudan, Wu J, Xuong NH, Ten Eyck LF, Taylor SS. Dynamic features of cAMPdependent protein kinase revealed by apoenzyme crystal structure. J Mol Biol. 2003 Mar 14;327(1):159-71. PubMed PMID: 12614615.

Azam M, Seeliger MA, Gray N, Kuriyan J, Daley GQ (2008). c-Src kinase domain Thr338Ile mutant in complex with ATPgS Nat. Struct. Mol. Biol 15: 1109–1118

Baker PR, Chalkley RJ (2014). MS-viewer: A web-based spectral viewer for proteomics results. Mol. Cell. Proteomics 13(5) 1392–6.

Ban H, Gavrilyuk J, Barbas CF (2010). Tyrosine bioconjugation through aqueous ene-type reactions: A click-like reaction for tyrosine. JACS 132, 5, 1523–1525

Boggon TJ, Eck MJ (2004). Structure and regulation of Src family kinases. Oncogene 23, 7918–7927

Brigham JL, Perera BG, Maly DJ (2013). A hexylchloride-based catch-and-release system for chemical proteomic applications. ACS Chem. Biol 8(4): 691–9

Butt TR, Edavettal SC, Hall JP, Mattern MR (2005). SUMO fusion technology for difficult-to-express proteins. Protein Expression and Purification 43, 1, 1–9

Chakraborty S, Inukai T, Fang L, Golkowski M, Maly DJ (2019). Targeting dynanmic ATP-binding site features allows discrimination between highly homologous protein kinases. ACS Chem Biol 14 6, 1249–1259

Chakraborty, S., Ahler, E., Simon, J. J., Fang, L., Potter, Z. E., Sitko, K. A., Stephany, J. J., Guttman, M., Fowler, D. M., Maly D. J., Profiling of the drug resistance of thousands of Src tyrosine kinase mutants uncovers a regulatory network that couples autoinhibition to the dynamics of the catalytic domain. bioRxiv 2021.12.05.471322; doi: https://doi.org/10.1101/2021.12.05.471322

Chapman HW, Butler DG, Newell M. The route of liquids administered to calves by esophageal feeder. Can J Vet Res. 1986 Jan;50(1):84-7. PubMed PMID: 3742363

Chea EE, and Jones LM (2018). Modifications generated by fast photochemical oxidation of proteins reflect the native conformations of proteins. Protein Science 27, 1047–1056

Engh RA, Bossemeyer D. Structural aspects of protein kinase control-role of conformational flexibility. Pharmacol Ther. 2002 Feb-Mar;93(2-3):99-111. PubMed PMID: 12191603.

Esposito D, Weile J, Shendure J, Starita LM, Papenfuss AT, Roth FP, Fowler DM, Rubin AF (2019). MaveDB: An open-source platform to distribute and interpret data from multiplexed assays of variant effect. Genome Biol 20, 223.

Fang J, Rand KD, Beuning PJ, Engen JR (2011). False EX1 signatures caused by samples carryover during HX MS analyses. Int. J. Mass Spectrom 302 (1–3) 19–25.

Fang L, Vilas-Boas J, Chakraborty S, Potter ZE, Register A, Seeliger M, Maly DJ (2020). How ATP-Competitive Inhibitors Allosterically Modulate Tyrosine Kinases That Contain a Src-like Regulatory Architecture. ChemRxiv, 10.26434/chemrxiv.12148524.v2

Findlay GM, Daza RM, Martin B, Zhang MD, Leith AP, Gasperini M, Janizek JD, Huang X, Starita LM, Shendure J(2019). Accurate classification of BRCA1 variants with saturation genome editing. Nature 562, 217–222

Foda ZH, Shan Y, Kim ET, Shaw DE, Seeliger MA (2015). A dynamically coupled allosteric network underlies bunding cooperativity in Src kinase. Nat. Comm 6, 5939

Fowler DM, Fields S (2014). Deep mutational scanning: A new style of protein science. Nat. Meth 11, 801–807

Gau BC, Sharp JS, Rempel D, Gross ML (2009). Fast photochemical oxidation of protein footprints faster than protein unfolding. Anal. Chem 81, 16, 6563–6571

Hacker SM, Backus KM, Lazear MR, Forli S, Correia BE, Cravatt BF (2017). Global profiling of lysine reactivity and ligandability in the human proteome. Nat. Chem 9(12):1181–1190

Hambly DM, Gross ML (2005). Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. J. Am. Soc. Mass Spectrom 16, 2057–2063

Hamuro Y (2017). Determination of equine cytochrome c backbone amide hydrogen/deuterium exchange rates by mass spectrometry using a wider time window and isotope envelope. J. Am. Soc. Mass Spectrom 28, 3, 486–497

Hamuro Y and Coales SJ (2018). Optimization of feasibility stage for hydrogen/deuterium exchange mass spectrometry. J. Am. Soc. Mass Spectrom 29(3) 623–629.

Hietpas RT, Jensen JD, Bolon DN (2011). Experimental illumination of a fitness landscape. Proc. Natl. Acad. Sci 108, 19, 7896–901

Hodkinson JP, Jahn TR, Radford SE, and Ashcroft AE (2009). HDX-ESI-MS reveals enhanced conformational dynamics of the amyloidogenic protein β2- microglobulin upon release from the MHC-1. Journal of the American Society for Mass Spectrometry 20, 278–286

Houde D, Arndt J, Domeier W, Berkowitz S, and Engen JR (2009). Characterization of IgG1 Conformation and Conformational Dynamics by Hydrogen/Deuterium Exchange Mass Spectrometry. Analytical Chemistry 81, 2644–2651

Huse M, Kuriyan J. The conformational plasticity of protein kinases. Cell. 2002 May 3;109(3):275-82. Review. PubMed PMID: 12015977.

Jeffrey PD, Russo AA, Polyak K, Gibbs E, Hurwitz J, Massagué J, Pavletich NP. Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. Nature. 1995 Jul 27;376(6538):313-20. PubMed PMID: 7630397.

June, C. H. & Sadelain, M. Chimeric Antigen Receptor Therapy. N Engl J Med 379, 64–73, doi:10.1056/NEJMra1706169 (2018).

Jung JE, Jura N (2016). Structural basis for the non-catalytic functions of protein kinases. Structure. 24, 1, 7–2

Kaur U, Johnson DT, Chea EE, Deredge DJ, Espino JA, Jones LM (2018). Evolution of structural biology through the lens of mass spectrometry. Anal. Chem 91, 142–155

Kornev AP, Haste NM, Taylor SS, Eyck LF. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. Proc Natl Acad Sci U S A. 2006 Nov 21;103(47):17783-8. Epub 2006 Nov 9. PubMed PMID: 17095602

Kornev AP, Taylor SS (2016). Dynamics driven allostery in protein kinases. Trends Biochem. Sci 40(11): 628–647.

Krishnamurty R, Brigham JL, Leonard SE, Ranjitkar P, Larson ET, Dale EJ, Merritt EA, Maly DJ (2013). Active site profiling reveals coupling between domains in Src-family kinases. Nat. Chem. Biol 9 1, 43–50

Kritzer JA, Freyzon Y, Lindquist S (2018). Yeast can accommodate phosphotyrosine: v-Src toxicity in yeast arises from a single disrupted pathway. FEMS Yeast Res. 18, 3

Kuriyan J, Eisenberg D (2007). The origin of protein interactions and allostery in colocalization. Nature12, 450 (7172), 982–990

Kwarcinski FE, Brandvold KR, Phadke S, Beleh OM, Johnson TK, Meagher JL, Seeliger MA, Stuckey JA, Soellner MB (2016). Conformation-selective analogues of dasatinib reveal insight into kinase inhibitor binding and selectivity. ACS Chem Biol 11 5, 1296–1304

Lee JO, Yang H, Georgescu MM, Di Cristofano A, Maehama T, Shi Y, Dixon JE, Pandolfi P, Pavletich NP (1999). Crystal structure of the PTEN tumor suppressor: implications for its phosphoinositide phosphatase activity and membrane association. Cell 99: 323–334

Leonard SE, Register AC, Krishnamurty R, Brighty GJ, Maly DJ (2014). Divergent modulation of Src-family kinase regulatory interactions with ATP-competitive inhibitors. ACS Chem Biol 9 (8), 1894–905

Li S, Yang X, Jia S, Weeks AM, Hornsby M, Peter SL, Nichiporuk RV, Iavarone AT, Wells JA, Toste FD, Chang CJ (2017). Redox-based reagents for chemoselective methionine bioconjugation. Science 355,6325, 597–602

Lo WL, Shah NH, Ahsan N, Horkova V, Stepanek O, Salomon AR, Kuriyan J, Weiss A. Lck promotes Zap70-dependent LAT phosphorylation by bridging Zap70 to LAT. Nat Immunol. 2018 Jul;19(7):733-741. doi: 10.1038/s41590-018-0131-1. Epub 2018 Jun 18. PMID: 29915297; PMCID: PMC6202249.

Majumdar R, Manikwar R, Hickey JM, Arora J, Middaugh CR, Volkin DB, Weis DD (2012). Minimizing carry-over in an online pepsin digestion system used for the H/D exchange mass spectrometric analysis of an IgG1 monoclonal antibody. J. Am. Soc. Mass Spectrom 23(12) 2140–8.

Majzner RG, Mackall CL. Clinical lessons learned from the first leg of the CAR T cell journey. Nat Med. 2019 Sep;25(9):1341-1355. doi: 10.1038/s41591-019-0564-6. Epub 2019 Sep 9. PMID: 31501612.

Manning, D.B. Whyte, R., Martinez, Hunter, T., Sundarsanam, S. The protein kinase complement of the human genome. Science. 2002 Dec 6. Vol. 298, Iss 5600, pp 1912-1934. 10.1126/science.1075762

Matreyek KA, Stariata LM, Stephany JJ, Martin B, Chiasson MA, Gray VE, Kircher M, Dines JN, Hause RJ, Bhatia S, Evans WE, Relling MV, Yang W, Shendure J, Dowler DM (2018). Multiplex assessment of protein variant abundance by massively parallel sequencing. Nat. Genet 50, 874–882

McClendon CL, Kornev AP, Gilson MK, Taylor SS (2014). Dynamic architecture of a protein kinase. Proc. Natl. Acad. Sci. USA 111(43): E4623–E4631.

Mestermann K, Giavridis T, Weber J, Rydzek J, Frenz S, Nerreter T, Mades A, Sadelain M, Einsele H, Hudecek M. The tyrosine kinase inhibitor dasatinib acts as a pharmacologic on/off switch for CAR T cells. Sci Transl Med. 2019 Jul 3;11(499):eaau5907. doi: 10.1126/scitranslmed.aau5907. PMID: 31270272; PMCID: PMC7523030.

NCBI Resource Coordinators, Database resources of the National Center for Biotechnology Information, Nucleic Acids Research, Volume 41, Issue D1, 1 January 2013, Pages D8– D20, https://doi.org/10.1093/nar/gks1189

Pellicena P, Kuriyan J (2006). Protein-protein interactions in the allosteric regulation of protein kinases. Curr. Opin. Struct. Biol 16, 6, 702–709

Potter ZE, Lau HT, Chakraborty S, Fang L, Guttman M, Ong SE, Fowler DM, Maly DJ. Parallel Chemoselective Profiling for Mapping Protein Structure. Cell Chem Biol. 2020 Aug 20;27(8):1084-1096.e4. doi: 10.1016/j.chembiol.2020.06.014. Epub 2020 Jul 9. PMID: 32649906; PMCID: PMC7484201.

Register AC, Leonard SE, Maly DJ (2014). SH2-catalutic domain linker heterogeneity influences allosteric coupling across the SFK family. Biochemistry 53 (44), 6910–23

Roscoe B, Thayer KM, Zeldovish KB, Fushman D, Bolon DN (2013). Analyses of the effects of all ubiquitin point mutants on yeast growth rate. J. Mol. Biol 425(8) 1363–1377.

Rubin AF, Gelman H, Lucas N, Bajjalieh SM, Papenfuss AT, Speed TP, Fowler DM. A statistical framework for analyzing deep mutational scanning data. Genome Biol. 2017 Aug 7;18(1):150. doi: 10.1186/s13059-017-1272-5. Erratum in: Genome Biol. 2018 Feb 7;19(1):17. PMID: 28784151; PMCID: PMC5547491.

Schade AE, Schieven GL, Townsend R, Jankowska AM, Susulic V, Zhang R, Szpurka H, Maciejewski JP. Dasatinib, a small-molecule protein tyrosine kinase inhibitor, inhibits T-cell activation and proliferation. Blood. 2008 Feb 1;111(3):1366-77. doi: 10.1182/blood-2007-04-084814. Epub 2007 Oct 25. PMID: 17962511; PMCID: PMC2214733.

Shiozaki EN, Gu L, Yan N, Shi Y (2004). Structure of the BRCT repeats of BRCA1 bound to a BACH1 phosphopeptide: implications for signaling. Mol. Cell 14: 405–412

Suzuki Y, St Onge RP, Mani R, King OD, Heilbut A, Labunskyy VM, Chen W, Pham L, Zhang LV, Tong AH, Nislow C, Giaever G, Gladyshev VN, Vidal M, Schow P, Lehár J, Roth FP. Knocking out multigene redundancies via cycles of sexual assortment and fluorescence selection. Nat Methods. 2011 Feb;8(2):159-64. doi: 10.1038/nmeth.1550. Epub 2011 Jan 9. PMID: 21217751; PMCID: PMC3076670.

Talpaz M, Shah NP, Kantarjian H, Donato N, Nicoll J, Paquette R, Cortes J, O'Brien S, Nicaise C, Bleickardt E, Blackwood-Chirchir MA, Iyer V, Chen TT, Huang F, Decillis AP, Sawyers CL. Dasatinib in imatinib-resistant Philadelphia chromosome-positive leukemias. N Engl J Med. 2006 Jun 15;354(24):2531-41. doi: 10.1056/NEJMoa055229. PMID: 16775234.

Taylor SS, Keshwani MM, Steichen JM, Kornev AP (2012). Evolution of the eukatyotic protein kinases as dynamic molecular switches. Philos. Trans. R. Soc. Long. B. Biol. Sci 367(1602): 2517–2528.

Taylor SS, Meharena HS, Kornev AP (2019). Evolution of a dynamic molecular switch. IUBMB Life 71, 6

Tong M, Pelton JG, Gill ML, Zhang W, Picart F, Seeliger MA (2017). Survey of solution dynamics in Src kinase reveals allosteric cross talk between the ligand binding and regulatory sites. Nat.. Commun 8 1, 2160.

Tong M, Seeliger MA (2015). Targeting conformation plasticity of protein kinases. ACS Chem. Biol 10, 1, 190–200

Ung PM, Rahman R, Schlessinger A (2018). Redefining the protein kinase conformation space with machine learning. Cell Chem. Biol 25, 7, 916–924

Vahihi S and Konermann L (2016). Probing the time scale of FPOP: Radical reactions over tens of milliseconds. J. Am. Soc. Mass Spectrom 27, 7, 1156–1164

Vijay-Kumar S, Bugg CE, Cook WJ (1987). Structure of ubiquitin refined at 1.8 A resolution. J. Mol. Biol 194: 531–544

Walters BT, Ricciuti A, Mayne L, Englander SW (2012). Minimizing back exhange in the hydrogen exchange-mass spectrometry experiment. J. Am. Soc. Mass Spectrom 23(12) 2132–9.

Wang L, Pan H, Smith DL (2002). Hydrogen Exchange-Mass Spectrometry: Optimization of Digestion Conditions. Mol. Cell. Proteomics 1(2) 132–8

Weber EW, Lynn RC, Sotillo E, Lattin J, Xu P, Mackall CL. Pharmacologic control of CAR-T cell function using dasatinib. Blood Adv. 2019 Mar 12;3(5):711-717. doi: 10.1182/bloodadvances.2018028720. PMID: 30814055; PMCID: PMC6418502.

Weerapana E#, Wang C#, Simon GM, Richter F, Khare S, Dillon BDM, Bachovchin DA, Mowen K, Baker D, Cravatt BF. (2010). Quantitative reactivity profiling predicts functional cysteines in proteomes. Nature 468, 790–795. #These authors contributed equally to this work.

Xu W, Doshi A, Lei M, Eck MJ, Harrison SC (1999). Crystal structures of c-Src reveal features of its autoinhibitory mechanism. Mol. Cell 3: 629–638

Zhang H, Wen J, Huang R, Blankenship RE, Gross ML (2012). Mass spectrometry-based carboxyl footprinting of proteins: method evaluation. Int. J. Mass Spectrom 312: 78–86

Zhang Z, Zhang A, Xiao G (2012). Improved protein hydrogen/deuterium exchange mass spectrometry platform with fully automated data processing. Anal. Chem 84(11) 4942–9.

Zhao Z, Wu H, Wang L, Liu Y, Knapp S, Liu Q, Gray NS (2014). Exploration of type II binding mode: A privileged approach for kinase inhibitor focused drug discovery? ACS Chem. Biol 9, 6, 1230–41