View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by DigitalCommons@URI

THE UNIVERSITY OF RHODE ISLAND

University of Rhode Island
DigitalCommons@URI

Cell and Molecular Biology Faculty Publications

Cell and Molecular Biology

2018

Structural versatility that serves the function of the HRD motif in the catalytic loop of protein tyrosine kinase, Src

Yixin Cui
University of Rhode Island

Gongqin Sun
University of Rhode Island, gongqinsun@uri.edu

Follow this and additional works at: https://digitalcommons.uri.edu/cmb_facpubs

The University of Rhode Island Faculty have made this article openly available.
Please let us know how Open Access to this research benefits you.

This is a pre-publication author manuscript of the final, published article.

Terms of Use
This article is made available under the terms and conditions applicable towards Open Access Policy Articles, as set forth in our Terms of Use.

Citation/Publisher Attribution
Cui, Y. and Sun, G. (2019), Structural versatility that serves the function of the HRD motif in the catalytic loop of protein tyrosine kinase, Src. Protein Science. doi: 10.1002/pro.3554 Available at: https://doi.org/10.1002/pro.3554

This Article is brought to you for free and open access by the Cell and Molecular Biology at DigitalCommons@URI. It has been accepted for inclusion in Cell and Molecular Biology Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.

Protein Science

Structural versatility that serves the function of the HRD motif in the catalytic loop of protein tyrosine kinase, Src

Journal:	*Protein Science*
Manuscript ID	PRO-18-0176.R1
Wiley - Manuscript type:	Full-Length Papers
Date Submitted by the Author:	n/a
Complete List of Authors:	Cui, Yixin; University of Rhode Island, Cell and Molecular Biology Sun, Gongqin; University of Rhode Island, Cell and Molecular Biology
Keywords:	Src protein tyrosine kinase, catalytic loop, HRD motif, bacterial colony screening

SCHOLARONE™
Manuscripts

Structural versatility that serves the function of the HRD motif in the catalytic loop of protein tyrosine kinase, Src

Yixin Cui and Gongqin Sun*

Department of Cell and Molecular Biology, Center for Biotechnology and Life Science, University of Rhode Island, 120 Flagg Road, Kingston, RI 02881

***Correspondence should be addressed to**:

Gongqin Sun

Department of Cell and Molecular Biology, Center for Biotechnology and Life Science, University of Rhode Island, 120 Flagg Road, Kingston, RI 02881

Email: gsun@uri.edu,

Telephone: 1-401-874-5937

Running title: Structural flexibility in the HRD motif of PTK catalytic loop

List of manuscript pages, tables and figures:

Total manuscript pages (including references): 26

Total number of figures: 7

[1]Abbreviations:

HRD motif, His-Arg-Asp residues in the catalytic loop; PTK, protein tyrosine kinase; PTP, protein tyrosine phosphatase.

Abstract: Site-directed mutagenesis is a traditional approach for structure-function analysis of protein tyrosine kinases, and it requires the generation, expression, purification, and analysis of each mutant enzyme. In this study, we report a versatile high throughput bacterial screening system that can identify functional kinase mutants by immunological detection of tyrosine phosphorylation. Two key features of this screening system are noteworthy. First, instead of blotting bacterial colonies directly from Agar plates to nitrocellulose membrane, the colonies were cultured in 96-well plates, and then spotted in duplicate onto the membrane with appropriate controls. This made the screening much more reliable compared to direct colony blotting transfer. A second feature is the parallel use of a protein tyrosine phosphatase (PTP)-expressing host and a non-PTP-expressing host. Because high activity Src mutants are toxic to the host, the PTP system allowed the identification of Src mutants with high activity, while the non-PTP system identified Src mutants with low activity. This approach was applied to Src mutant libraries randomized in the highly conserved HRD motif in the catalytic loop, and revealed that structurally diverse residues can replace the His and Arg residues, while the Asp residue is irreplaceable for catalytic activity.

Keywords: Src protein tyrosine kinase, catalytic loop, HRD motif, bacterial colony screening.

Impact statement: Protein tyrosine kinases are a large family of regulatory enzymes associated with cancer development. Structural variations among protein tyrosine kinases determine their function and regulation in the cells. Identifying structure-function relationships is essential for understanding their catalytic and regulatory mechanisms and for drug development. This study presents a versatile approach for identifying functional mutants of protein tyrosine kinase, in the effort to establish protein tyrosine kinase structure-function relationships.

1. Introduction

Protein tyrosine kinases (PTKs)[1] are important signaling proteins and anticancer drug targets[1, 2]. A full understanding of their biological roles in normal and pathological conditions, and rational inhibitor design require an understanding of the structural basis of PTK catalysis, regulation, and substrate specificity. The catalytic domains of PTKs are generally similar, sharing many residues essential for structural integrity, catalysis, and regulation [3, 4]. Elucidating the roles of evolutionarily conserved as well as variable residues in PTK catalysis, regulation or structural integrity requires extensive research using multidisciplinary tools, such as biochemistry, structural biology and biophysics.

One useful tool for this research is site-directed mutagenesis and heterologous expression. A residue in a kinase can be replaced by any other standard amino acid residue and the mutant enzyme can often be readily expressed in an appropriate bacterial host. Upon purification of a mutant enzyme, it can be characterized to determine the effects of the mutation. The challenge with this approach is that it is time-consuming because each mutant needs to be expressed, purified and characterized. Considering that each residue could require up to 19 mutants for a full analysis, the challenge grows exponentially with the addition of each residue to analyze. Furthermore, when a mutant has no catalytic activity, the effect of the mutation becomes difficult to characterize. It would therefore be desirable to develop a convenient and high throughput screening system in bacteria to identify active PTK variants from a library of mutants.

There are several challenges in developing this approach. First, while some PTKs can be readily expressed in bacteria, others are toxic to the bacterial host [5, 6]. This toxicity is due to the phosphorylation of bacterial proteins by the heterologous PTK, but can be overcome by the

co-expression of a protein tyrosine phosphatase (PTP) [6-9]. Second, expression of PTKs in a bacterial host is often plagued by other generic issues, such as differences in codon usage between higher organisms and bacteria, protein folding and solubility [10]. Solutions to each of these problems have been developed. Plasmids that express tRNAs rarely used in bacteria can help alleviate the codon usage problem [11, 12], while co-expression of chaperones, such as GroES/EL can help heterologous PTK fold more efficiently [13-16].

In this study, we developed a bacterially based rapid high throughput screening method that incorporates each of these elements. We validate this system by using it to study the functionality of the HRD motif in the protein tyrosine kinase Src. The HRD motif is the first three residues in the highly conserved catalytic loop of protein kinases [17]. In Src, The HRD motif is three residues 384-386. Asp386 of this motif is the catalytic base, and is considered indispensable for catalysis [18-22]. His384 and Arg385 have been thought to be key regulatory residues as their interactions vary in different activity states of several kinases [20, 23-28]; however their exact roles in PTK regulation have not been fully defined. This study confirms that the Asp386 residue is indeed essential for catalysis as previously demonstrated, while the His384 and Arg385 can be replaced by a wide variety of residues without abolishing kinase activity.

2. Results

2.1. Bacterial proteins are phosphorylated on tyrosine residues when PTKs are heterologously expressed in E coli. For screening of active PTKs or mutants in bacteria, it would be desirable if the kinases phosphorylate bacterial proteins that can be detected by Western blotting using anti-phosphotyrosine antibodies. It will then spare the need to express a specific substrate for the kinase in the bacterial cells. To confirm that eukaryotic PTKs phosphorylate

bacterial proteins, four protein tyrosine kinases, Abl, Csk, FGFR1, and Lck were each cloned into pGEX-4T-1 and the plasmids were introduced into *E. coli* DH5α (hereafter called *E. coli*). Fig. 1A shows a Western blot detecting protein tyrosine phosphorylation after the bacteria were cultured in LB medium at 37^0C for 24 hours. *E. coli* cells containing the empty expression vector did not produce any detectable protein tyrosine phosphorylation (Fig 1A, lane 1). Expression of each PTK resulted in significant tyrosine phosphorylation of proteins of various sizes. Even with FGFR1, which resulted in relatively low levels of protein phosphorylation, it was still readily detectable.

2.2. Active Src expression in E. coli requires the co-expression of a protein tyrosine phosphatase. Fully active Src is toxic to *E. coli* and cannot be expressed unless it is co-expressed with a PTP [6-9]. While the co-expression of a PTP helps reduce the toxicity associated with PTK expression, the PTP might reduce the phosphorylation of bacterial proteins to a level that is undetectable by Western blotting. That would preclude phosphotyrosine-based screening. We therefore examined whether phosphotyrosine can still be detected in *E. coli* when active Src is co-expressed with a PTP. The wtSrc, kinase-defective Src (kdSrc, containing a Lys295Met mutation) [13, 14] and truncated Src (SΔD-Src: His384 was replaced with Ser, and Arg385 replaced with a stop codon) in the pRSET-a plasmid were introduced into *E. coli* BL21(DE3)RIL harboring the pREP4groESL and pCDF-PTP1B plasmids. The truncated Src mutant, SΔD-Src, was included here because it will be used as a parental plasmid for generating Src mutant libraries in later experiments. By mutating the HR residues in the HRD motif to SΔ we achieved two purposes: the mutation completely eliminated Src kinase activity and generated a restriction site (AGTACT for Sca I) on the plasmid. Using SΔD-Src as a parental plasmid for

library generation eliminates parental plasmids as false positives in the library screening. Incorporating a restriction site allows convenient mutation verification.

Individual colonies were grown for 24 h at 37^0C in LB broth. Soluble fractions from the cultured cells were analyzed by Western blotting using monoclonal antibodies specific for phosphotyrosine or phosphotyrosine 416 of Src, which is the site of Src autophosphorylation [29]. The active wtSrc phosphorylated bacterial proteins and Src itself on Tyr416, whereas neither kdSrc nor SΔD-Src resulted in detectable phosphorylation (Fig. 1B). The presence of the PTP significantly reduced the number of proteins phosphorylated on tyrosine compared to those by other kinases without the PTP (Fig. 1A and 1B). Even in the presence of the PTP, however, the residual phosphotyrosine level was still sufficient for Western blot detection when wtSrc was expressed.

Co-expression of a PTP poses a dilemma for phosphotyrosine-based screening to identify active Src mutants. While the expression of wtSrc and Src mutants with high activity requires a PTP, co-expression of a PTP may reduce the phosphotyrosine signal to a point as to prevent identification of mutants with low activity. As previously demonstrated, low activity Src mutants can be expressed in bacteria without the help of a PTP [6]. With the opposing requirements by low activity and high activity Src mutants, we introduced each Src mutant library into *E. coli* co-expressing the PTP 1B (the PTP system) and *E. coli* not co-expressing PTP 1B (the non-PTP system) for dot blot screening. We anticipated that the more active mutants would likely be selected in the PTP system and the less active mutants would likely be selected in the non-PTP system.

2.3. Dot blot screening. We explored the dot blotting format to make the screening more high

throughput. There are two general ways of transferring colonies onto a nitrocellulose membrane for screening: blotting colonies from an Agar plate onto the membrane directly, or culturing the colonies in liquid and dot blotting the cultures on to the membrane. While the former is more direct and convenient, it did not work well in our hands. Both the colony size and the blotting transfer efficiency from the plate to the membrane varied from colony to colony, resulting in low reproducibility. The second approach worked much better and is outlined in Fig. 2B. It involves inoculating colonies for overnight individual liquid cultures in 96-well plates, and spotting a similar number of cells on the membrane. Positive controls (cells expressing an active kinase) and negative controls (cells that do not express an active kinase) can be spotted on the membrane, and each can be spotted in duplicate. Using this approach, the detection signals from the positive and negative controls consistently reflected the kinase activity expressed.

We tested this screening strategy on Src by determining the residues that can functionally replace the HRD motif. The HRD motif is the first three residues (His384-Arg385-Asp386 in Src) in the catalytic loop. Asp386 in Src (or its equivalent in other kinases) is the catalytic base [22, 30] in catalyzing the phosphoryl transfer reaction. As such, this Asp residue is thought to be essential for catalysis, but it is not clear if any other residue can functionally replace it. His384 and Arg385 are thought to be important regulatory residues [28, 31], even though the specific roles they play in regulation are not clear. It has been shown that His384 in Src can be replaced by several residues and still maintain some level of kinase activity [6]. Mutation of Arg385 to Ala results in low kinase activity in v-Src [32], while an Arg to Cys mutation has little effect on the kinase activity of Drosophila Src64 [33]. Considering the ambiguous requirements for these residues, we thought a thorough evaluation of these positions may be informative, and identifying functional HRD mutants of Src would be a good test for the screening system.

To apply the high-throughput screening to the HRD motif in Src, we first generated a truncated form of Src, SΔD-Src, in which Ser replaced His384 and a stop codon replaced Arg385. With this truncation in the HRD motif, SΔD-Src loses approximately half of the catalytic domain, and all its catalytic activity. The mutant libraries are generated by mutating the codons for SΔD into appropriate codons required for each mutant library, as explained in the following sections. The reason we used SΔD-Src, instead of wtSrc, as the parental plasmid is that because SΔD-Src is catalytically truncated and thus not active, any unmutated parental plasmid will not result in a false positive colony during screening. Otherwise, the false positives from unmutated original plasmids would outnumber the active mutants from the mutation library.

2.4. Identification of residues that can replace His384 in the HRD motif. A His384 mutation library was created by mutating the SΔ codons in SΔD-Src to the codons for XRD, where X is the randomized residue, while R restored the wt Arg residue at the 385 position. The randomized plasmids were introduced into both the non-PTP system and the PTP system, and the colonies were subjected to the screen outlined in Fig. 2B. Representative screening results from the PTP system and non-PTP system, respectively, are shown in Fig. 3. As shown in the blots, the positive controls (wtSrc) consistently produced positive signals while the negative controls (kdSrc and SΔD-Src) consistently produced undetectable signals. When a box containing a colony from the mutant library produced two visible signals, the mutant was considered a positive. The colonies were then cultured and the plasmids from the positive colonies were isolated and sequenced to identify the mutations.

Approximately 630 random colonies were screened in the PTP system (Fig. 3A), and five positives were identified. All five mutants had histidine at the 384 position, with both CAT and

CAC codons used. This result indicated that only wild type Src with a His at this position generated sufficient tyrosine phosphorylation signal to be selected as a positive in the PTP system. In the non-PTP system, ~20 mutants on average were positives for each 96-slot screening (Fig. 3B). Fifty random positives were selected and their plasmids were sequenced. Nine different mutants, His384Gln, His384Tyr, His384Thr, His384Ser, His384Phe, His384Leu, His384Val, His385Ile and His384Gly, were identified. Wild type Src was not identified from the non-PTP system, presumably because wtSrc is toxic to *E. coli* without a PTP.

To confirm the catalytic activity of the His384 mutants, all nine mutant enzymes were expressed and purified. The relative kinase activity of these mutants was determined using polyE_4Y as a substrate (Fig. 4). All mutants maintained active kinase activity and His385Gln had the highest catalytic activity at ~14% of that of the wild type. The functional substitutions for His384 were either hydrophobic residues, Phe (10.2%), Leu (1.1%), Val (0.6%) I (0.6%) and Gly (0.2%), or polar but not charged residues, Gln (14.2%), Tyr (6.2%), Thr (6.0%) and Ser (2.6%). This result supports the proposal that His384 is involved in hydrophobic interactions in stabilizing the active conformation [25], and also suggests that His384 may interact with other residues through hydrogen bonds in some conformations.

2.5. Identification of residues that can replace Arg385 in the HRD motif. We next applied this strategy to Arg385 by replacing SΔ codons in SΔD-Src with codons for HX, where H restores the His384 and X is randomized residue at the 385 position. Representative screening results using the PTP and non-PTP systems are shown in Fig. 5. In both expression systems, a large number of positive colonies were identified. Approximately 50 functional mutants were selected and sequenced from each screening system. While each system produced its own unique

variants, some mutants were found in both systems. Arg385Gln and Arg385Tyr were selected only from the non-PTP system, while Arg385Lys, Arg385Val, Arg385Asn, Arg385Leu, Arg385Cys and wtSrc were only found in the PTP system. Arg385Ala, Arg385Gly, Arg385Ser, and Arg385Thr were selected from both systems. These results suggested that a variety of Arg385 mutants maintained significant catalytic activity.

Selected mutants from both systems were purified and the catalytic activity was determined (Fig. 6). As expected, many mutants displayed significant kinase activity, with some approaching that of wtSrc. It is remarkable that Val, Asn and Cys, with side chains very different from each other and from that of Arg, can all replace Arg385 and maintain most of the kinase activity. No negatively charged residue was identified at this position from either system.

2.6. The PTP system and non-PTP system can each accommodate Src mutants with kinase activities within a defined range. An examination of the results of all of the above screening suggests that the PTP system and the non-PTP system can each accommodate Src mutants with a broad range of kinase activities. The 9 mutants selected from the PTP system had kinase activities ranging from ~17.2 (Arg385Ser) to 88.3% (Arg385Asn). Because wtSrc was also selected from the PTP system, this result indicated that the range of kinase activity that can be selected by the PTP system is ~17% to 100% of the wtSrc activity. Mutants with activities below 17% were likely unable to produce a positive signal due to the PTP dephosphorylation. Among the 13 mutants identified from the non-PTP system, including 9 His384 mutants and 4 Arg385 mutants, Arg385Thr had the highest activity at 30.8%, while His384Gly had the lowest activity at 0.2%, indicating that the selective range for the non-PTP system was 0.2% to 31% of the wild type activity. The upper limit is likely determined by the threshold of kinase toxicity that can be

tolerated by the bacterial host cells.

2.7. Identification of residues that can replace Asp386 in the HRD motif. The same approach was applied to identify residues that can replace Asp386 in the HRD motif. In generating the Asp386 mutants, the codons for SΔD in SΔD-Src were replaced with the codons for HRX. The HR codons restored the wildtype His and Arg residues at 384 and 385 positions, while the X represented the randomized codon for the 386 position. The library was screened in the PTP and non-PTP systems, and typical results are shown in Fig. 7. A total of 630 colonies were screened in each screening system. Only the PTP system produced positive colonies, and sequencing indicated that all the positive colonies contained wtSrc, with Asp386 being coded by either GAT or GAC codons. With 630 colonies screened in each system, the probability of any of the 64 codons not being represented among these colonies in each of the systems is about 0.3%. This result indicated that replacing Asp386 by any other residue could not retain even 0.2% of wtSrc kinase activity, the lower limit of the non-PTP system. This result supports the assignment of Asp386 as an indispensable residue for Src catalysis [20].

3. Discussion

We report a bacterially based high throughput screening method for identifying functional PTK mutants. This method incorporates all the tools that have been developed to help protein kinase expression in bacteria, including expressing rarely used tRNAs in bacteria, co-expressing a chaperone to help the target PTK fold, and co-expression of a protein tyrosine phosphatase to help overcome toxicity associated with the expression of some PTKs. While the co-expression of a PTP helps with the expression of a toxic PTK, its inclusion in the screening

system results in a bias that favors high activity mutants and prevents the identification of low activity mutants. Using parallel PTP and non-PTP systems removes the bias, and enables the identification of mutants with all levels of activity. While such a system may not be necessary for all PTKs, it is particularly useful for a toxic PTK, such as Src. Using several small Src mutant libraries randomizing the highly conserved HRD motif, the benefits of the parallel PTP and non-PTP systems became obvious, as each system selected Src mutants with kinase activities that fall within a broad range.

Another important feature of the screening system is how the colonies from Agar plates are transferred to a nitrocellulose membrane. Instead of transferring the colonies by directly blotting the colonies from the plates to the membrane, each colony was cultured in 96 well plates, and comparable number of cells were then dot-blotted onto the filter paper. This eliminated the variation in colony size, uneven colony blotting, and also enabled duplication of colonies and incorporation of positive and negative controls in the screening. Despite the extra step that may limit the number of colonies that can be conveniently screened, this approach is superior for most applications. A previous report described colony screening to identify substrates for a given kinase [34]. In this earlier study, no PTP and chaperones were used. While this method served the purposes of the previous study, the incorporation of the PTP and chaperone is necessary to eliminate the bias caused by toxic PTKs. A comparison between the results from the PTP and non-PTP systems indicates that the non-PTP system has an inherent bias that would have precluded the identification of the high activity Src mutants.

The screening method was successfully used to identify mutants in the HRD motif of Src protein tyrosine kinase. The HRD motif is a highly conserved structural feature in PTKs, and among protein kinases in general. It is well established that the Asp386 in the HRD motif is the

catalytic base [20, 22]. Consistent with this designation is the extreme conservation at this position in the kinome. Among 447 active protein kinases encoded by the human genome, only one contains a non-Asp residue, a Ser in IRAK3, at this position[4]. In contrast, among 50 human pseudokinases, 33 contain a residue other than Asp at the equivalent position. Therefore, missing the Asp residue is considered a hallmark of pseudokinases, even though some pseudokinases were discovered to possess significant kinase catalytic activity [35-38]. Whether Asp can be replaced by any other residue has not been fully tested in any kinase. The screening of an Asp386 mutant library indicated that no other residue was able to replace Asp386 to allow significant kinase activity in Src (> 0.2% of wild type).

The roles His384 and Arg385 play in the kinase function are less understood. There are considerable variations at these positions even though there is a general conservation in the kinome. Among 447 active kinases encoded by the human genome, 52 (12%) contain a non-His residue at the position equivalent to His384, including Tyr (44/52), Leu (5/52), Cys (1/52), Phe (1/52), and Ile (1/5Most of these residues (except Cys) were also identified in our His384 mutant library (Fig. 4). Loss of His at this position is not a defining feature of pseudokinases, as 80% of pseudokinases retain a His at this position (compared to 88% among active kinases). Furthermore, most of the same residues used by active kinases are also used by pseudokinases. Approximately 83% of the 447 active human protein kinases and only 40% pseudokinase contain an Arg residue at the position equivalent to Arg385, suggesting that the Arg residue at this position is important for the kinase function. Among the 75 protein kinases using a non-Arg residue at this position, the top six substitutions are Leu (19/75), Gly (15/75), Cys (9/75), Ala (7/75), Thr (5/75) and Lys (5/75). All of these residues could indeed replace Arg to afford Src kinase activity in the Arg385 library (Fig. 6). Among the 50 pseudokinases, 20 have an Arg, 18

have a Gly, and a small number used Lys (4), Cys (2) among others. Because the similar sets of residues are used by the active kinases and peudokinases, the residue at this position does not define an active or inactive kinase. It is interesting to note that negatively charged residues, Asp or Glu, do not appear at the positions equivalent to H384 or R385 in any active or inactive human kinases, and were not selected in the H384 and R385 libraries. A question remains unanswered as to why His and Arg are conserved through evolution although they are not required in most kinases. It is plausible that conservation is based on certain non-catalytic functions unique to some kinases, such as regulation. Protein kinases in different families often have unique regulatory mechanisms, yet all the mechanisms are based on a common conformational elasticity embedded in the kinase catalytic domain [39-43]. Further studies will be needed to determine whether the HRD motif is involved in controlling conformational changes underlying kinase regulation.

4. Materials and Methods

4.1. Reagents and chemicals. Consumables, culture media, and media components were purchased from Fisher Scientific. The QuikChange mutagenesis kit was purchased from Stratagene. Specific antibodies were purchased from Cell Signaling Technology. DNA primers were synthesized by Integrated DNA Technologies. The chromatographic resin, Ni-NTA agarose, was purchased from Qiagen. Bradford reagent and pre-cast SDS-PAGE gels were purchased from Bio-Rad. All the other chemicals were purchased from Sigma.

4.2. Construction of PTK expression plasmids. Wild type human Abl catalytic domain [44], full-length Csk [14], FGFR1 catalytic domain [45], and full-length Lck [5] were cloned into

pGEX-4T-1 plasmid using *BamHI* and *EcoRI* restriction sites as described previously and expressed in *Escherichia coli* DH5α. All constructs were verified by sequence analysis. Plasmids were introduced into the specific expression systems by electroporation.

4.3. Molecular construction of Src and Src mutants and their expression in a PTP system and a non-PTP system. Wild type Src was cloned into the pRSETA plasmid using *Bam HI* and *HindIII* restriction sites as described previously [6, 13, 14]. Kinase-defective Src (kdSrc, containing a Lys295Met mutation) [13] and truncated Src (SΔ-Src: His384 was replaced with Ser, and Arg385 replaced with a stop codon) were generated by QuikChange site-directed mutagenesis. Active Src and Src mutants were expressed in a PTP system (BL21(DE3)RIL harboring the pREP4groESL and pCDF-PTP1B plasmid) [13] and a non-PTP system (BL21(DE3)RIL harboring the pREP4groESL) [6] for screening purposes.

4.4. Mutant library construction and active mutant screening. A QuikChange mutagenesis kit was used for the construction of mutant libraries. The plasmid pRSETA SΔ-Src was used as template for the construction of mutant plasmids to avoid wild type Src plasmid contamination in constructed libraries. The primers for the His384 mutant library were 5'-GGATGAACTACGTGXXXCGAGACCTGCGGGCGG and its complimentary primer. These primers changed the His CAC codon to a random sequence (XXX) while keeping all other neighboring residues as wildtype. The primers for the Arg385 mutant library were, 5'-GGATGAACTACGTGCACXXXGACCTGCGGGCGGCC and its complimentary primer. These primers changed the Arg CGA codon to a random sequence (XXX) while keeping all other neighboring residues as wildtype.

The primers for the Asp386 mutant library were 5'-GGATGAACTACGTGCACCGAXXXCTGCGGGCGGCCAAC and its complimentary primer. These primers changed the Asp GAC codon to a random sequence (XXX) while keeping all other neighboring residues as wildtype.

QuikChange was performed as instructed by the manufacturer. The constructed mutant plasmids were transformed into a PTP-system (BL21(DE3)RIL harboring the pREP4groESL and pCDF-PTP1B plasmids) and a non-PTP system (BL21(DE3)RIL harboring the pREP4groESL). Mutant colonies were inoculated into 96-well plates containing 200 µl in LB broth in each well with appropriate antibiotics. The plates were incubated for 24 h at 37^0C shaking at 250 rpm. Then two 0.4 µl samples from each well (representing approximately $8x10^5$ cells) were blotted directly onto a nitrocellulose membrane. The membrane was heated at 95^0C for 5 min to lyse the cells and blocked with milk. This was followed by immunoblotting using a mouse monoclonal antibody against phosphotyrosine (PY100; Cell Signaling Technology) and an anti-mouse secondary antibody conjugated to horseradish peroxidase (HRP) (Stratagene).

4.5. Expression and purification of Src and active Src mutants. As described above, wild type Src and active Src mutants were expressed in *Escherichia coli* BL21(DE3)RIL cells harboring the pREP4groESL plasmid and pCDF-PTP1B by the use of the pRSETA plasmid as described [6]. The plasmid pREP4groESL contains the genes for the GroES/EL chaperone, which helps Src fold correctly [13]. The pCDF-PTP1B directs the expression of PTP1B, which keeps Src and bacterial proteins dephosphorylated and reduces the toxicity of Src expression to bacterial cells [7]. The (His)6-tagged active Src constructs were purified using immobilized Ni-iminodiacetic acid-agarose as described previously [46]. Briefly, individual colonies were inoculated into 500

mL of LB culture medium containing appropriate antibiotics. The cultures were grown to an OD_{600} of 2.5 and diluted with 500 mL of fresh LB containing the antibiotics. The cultures were air-cooled to room temperature, and 0.4 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) was added to induce the production of the fusion protein. Cultures were allowed to induce for 6 h at room temperature with shaking at 250 rpm. Bacterial cultures were harvested by centrifugation at 4,500g and stored at -20 °C before use. Cell pellets were resuspended with ice-cold lysis buffer [50 mM Hepes (pH 8.0), 200 mM NaCl, 10 mM imidazole, and 0.1% Triton-X]. Cells were lysed by French press and clarified by centrifugation at 47,800g for 30 min at 4 °C. The supernatant was added to 0.75 mL of Ni-NTA agarose (Qiagen) and gently mixed by rotation at 4 °C for 30 min. The beads were loaded into a column and then washed with washing buffer (50 mM Hepes, pH 8.0, and 20 mM imidazole). Proteins were eluted by using 100–300 mM imidazole in 50 mM Hepes (pH 8.0). Enzymes were desalted and stored in 50 mM Hepes (pH 8.0) at -20 °C in 40% glycerol. Protein concentration was determined by Bradford reagent standardized with BSA (0.2–1.0 mg/ml). Purity of protein fractions was determined by SDS/PAGE (Bio-Rad) and stained with Coomassie blue.

4.6. Kinase assay. The kinase activity of wild type Src and Src mutants was determined by a luminescent ADP detection assay (ADP-Glo™ Kinase Assay, Promega) using $polyE_4Y$ and ATP as substrates. The phosphorylation reactions were performed in 50 µl volumes at 25°C in the protein kinase assay buffer: 50 mM *N*-(2-hydroxyethyl)piperazine-*N*′-3-propanesulfonic acid (pH 8.0) containing 5% glycerol, 0.005% Triton X-100, and 0.05% 2-mercaptoethanol. The standard assay used 12 mM $MgCl_2$, 0.2 mM ATP, and 1 mg/ml $polyE_4Y$. After a 30 min incubation, 5 µl of the kinase reaction mixture was mixed with 5 µl ADP-Glo™ reagent and incubated for 40 min

to stop the kinase reaction and deplete the unconsumed ATP. Kinase detection reagent (10 μl) was added and the mixture was incubated for another 30 min to convert produced ADP to ATP and introduce luciferase and luciferin to detect converted ATP. Then the luminescence produced was measured by a plate-reading luminometer. Assays were performed in duplicate, and each assay was repeated three times with reproducible results.

5. Acknowledgement

The authors have no conflict of interest in relation to this manuscript.

6. References

1. Blume-Jensen, P. and T. Hunter, *Oncogenic kinase signalling.* Nature, 2001. **411**(6835): p. 355-65.
2. Capdeville, R., et al., *Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug.* Nat Rev Drug Discov, 2002. **1**(7): p. 493-502.
3. Hubbard, S.R. and J.H. Till, *Protein tyrosine kinase structure and function.* Annu Rev Biochem, 2000. **69**: p. 373-98.
4. Manning, G., et al., *The protein kinase complement of the human genome.* Science, 2002. **298**(5600): p. 1912-34.
5. Sun, G. and R.J. Budde, *A modified pGEX expression system that eliminates degradation products and thrombin from the recombinant protein.* Anal Biochem, 1995. **231**(2): p. 458-60.
6. Kemble, D.J., Y.H. Wang, and G. Sun, *Bacterial expression and characterization of catalytic loop mutants of SRC protein tyrosine kinase.* Biochemistry, 2006. **45**(49): p.

14749-54.

7. Seeliger, M.A., et al., *High yield bacterial expression of active c-Abl and c-Src tyrosine kinases.* Protein Sci, 2005. **14**(12): p. 3135-9.

8. Wang, W., et al., *Structural characterization of autoinhibited c-Met kinase produced by coexpression in bacteria with phosphatase.* Proc Natl Acad Sci U S A, 2006. **103**(10): p. 3563-8.

9. Wang, Y.H., et al., *A new strategy to produce active human Src from bacteria for biochemical study of its regulation.* Biochem Biophys Res Commun, 2006. **346**(2): p. 606-11.

10. Rosano, G.L. and E.A. Ceccarelli, *Recombinant protein expression in Escherichia coli: advances and challenges.* Front Microbiol, 2014. **5**: p. 172.

11. Kleber-Janke, T. and W.M. Becker, *Use of modified BL21(DE3) Escherichia coli cells for high-level expression of recombinant peanut allergens affected by poor codon usage.* Protein Expr Purif, 2000. **19**(3): p. 419-24.

12. Rosano, G.L. and E.A. Ceccarelli, *Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted Escherichia coli strain.* Microb Cell Fact, 2009. **8**: p. 41.

13. Wang, D., X.Y. Huang, and P.A. Cole, *Molecular determinants for Csk-catalyzed tyrosine phosphorylation of the Src tail.* Biochemistry, 2001. **40**(7): p. 2004-10.

14. Lee, S., et al., *Determination of the substrate-docking site of protein tyrosine kinase C-terminal Src kinase.* Proc Natl Acad Sci U S A, 2003. **100**(25): p. 14707-12.

15. Amrein, K.E., et al., *Purification and characterization of recombinant human p50csk protein-tyrosine kinase from an Escherichia coli expression system overproducing the*

bacterial chaperones GroES and GroEL. Proc Natl Acad Sci U S A, 1995. **92**(4): p. 1048-52.

16. Caspers, P., M. Stieger, and P. Burn, *Overproduction of bacterial chaperones improves the solubility of recombinant protein tyrosine kinases in Escherichia coli.* Cell Mol Biol (Noisy-le-grand), 1994. **40**(5): p. 635-44.
17. Hanks, S.K., A.M. Quinn, and T. Hunter, *The protein kinase family: conserved features and deduced phylogeny of the catalytic domains.* Science, 1988. **241**(4861): p. 42-52.
18. Gibbs, C.S. and M.J. Zoller, *Rational scanning mutagenesis of a protein kinase identifies functional regions involved in catalysis and substrate interactions.* J Biol Chem, 1991. **266**(14): p. 8923-31.
19. Skamnaki, V.T., et al., *Catalytic mechanism of phosphorylase kinase probed by mutational studies.* Biochemistry, 1999. **38**(44): p. 14718-30.
20. Xu, W., et al., *Crystal structures of c-Src reveal features of its autoinhibitory mechanism.* Mol Cell, 1999. **3**(5): p. 629-38.
21. Ablooglu, A.J., et al., *Probing the catalytic mechanism of the insulin receptor kinase with a tetrafluorotyrosine-containing peptide substrate.* J Biol Chem, 2000. **275**(39): p. 30394-8.
22. Cole, P.A., et al., *The role of the catalytic base in the protein tyrosine kinase Csk.* J Biol Chem, 1995. **270**(38): p. 22105-8.
23. Xu, W., S.C. Harrison, and M.J. Eck, *Three-dimensional structure of the tyrosine kinase c-Src.* Nature, 1997. **385**(6617): p. 595-602.
24. Yamaguchi, H. and W.A. Hendrickson, *Structural basis for activation of human lymphocyte kinase Lck upon tyrosine phosphorylation.* Nature, 1996. **384**(6608): p. 484-

9.

25. Kornev, A.P., et al., *Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism.* Proc Natl Acad Sci U S A, 2006. **103**(47): p. 17783-8.

26. Kornev, A.P., S.S. Taylor, and L.F. Ten Eyck, *A helix scaffold for the assembly of active protein kinases.* Proc Natl Acad Sci U S A, 2008. **105**(38): p. 14377-82.

27. Williams, J.C., et al., *The 2.35 A crystal structure of the inactivated form of chicken Src: a dynamic molecule with multiple regulatory interactions.* J Mol Biol, 1997. **274**(5): p. 757-75.

28. La Sala, G., et al., *HRD Motif as the Central Hub of the Signaling Network for Activation Loop Autophosphorylation in Abl Kinase.* J Chem Theory Comput, 2016. **12**(11): p. 5563-5574.

29. Roskoski, R., Jr., *Src protein-tyrosine kinase structure and regulation.* Biochem Biophys Res Commun, 2004. **324**(4): p. 1155-64.

30. Hines, A.C., et al., *Bisubstrate analog probes for the insulin receptor protein tyrosine kinase: molecular yardsticks for analyzing catalytic mechanism and inhibitor design.* Bioorg Chem, 2005. **33**(4): p. 285-97.

31. Banavali, N.K. and B. Roux, *Flexibility and charge asymmetry in the activation loop of Src tyrosine kinases.* Proteins, 2009. **74**(2): p. 378-89.

32. Senften, M., et al., *Catalytic activity and transformation potential of v-Src require arginine 385 in the substrate binding pocket.* Oncogene, 1995. **10**(1): p. 199-203.

33. Strong, T.C., G. Kaur, and J.H. Thomas, *Mutations in the catalytic loop HRD motif alter the activity and function of Drosophila Src64.* PLoS One, 2011. **6**(11): p. e28100.

34. Lock, P., et al., *A new method for isolating tyrosine kinase substrates used to identify fish, an SH3 and PX domain-containing protein, and Src substrate.* EMBO J, 1998. **17**(15): p. 4346-57.

35. Boudeau, J., et al., *Emerging roles of pseudokinases.* Trends Cell Biol, 2006. **16**(9): p. 443-52.

36. Ungureanu, D., et al., *The pseudokinase domain of JAK2 is a dual-specificity protein kinase that negatively regulates cytokine signaling.* Nat Struct Mol Biol, 2011. **18**(9): p. 971-6.

37. Taylor, S.S., et al., *Pseudokinases from a structural perspective.* Biochem Soc Trans, 2013. **41**(4): p. 981-6.

38. Byrne, D.P., D.M. Foulkes, and P.A. Eyers, *Pseudokinases: update on their functions and evaluation as new drug targets.* Future Med Chem, 2017. **9**(2): p. 245-265.

39. Huse, M. and J. Kuriyan, *The conformational plasticity of protein kinases.* Cell, 2002. **109**(3): p. 275-82.

40. Johnson, L.N., M.E. Noble, and D.J. Owen, *Active and inactive protein kinases: structural basis for regulation.* Cell, 1996. **85**(2): p. 149-58.

41. Taylor, S.S., et al., *PKA: a portrait of protein kinase dynamics.* Biochim Biophys Acta, 2004. **1697**(1-2): p. 259-69.

42. McClendon, C.L., et al., *Dynamic architecture of a protein kinase.* Proc Natl Acad Sci U S A, 2014. **111**(43): p. E4623-31.

43. Saleh, T., P. Rossi, and C.G. Kalodimos, *Atomic view of the energy landscape in the allosteric regulation of Abl kinase.* Nat Struct Mol Biol, 2017.

44. Tiwari, R.K., et al., *Design, Synthesis, and Evaluation of Dasatinib-Amino Acid and*

Dasatinib-Fatty Acid Conjugates as Protein Tyrosine Kinase Inhibitors. ChemMedChem, 2017. **12**(1): p. 86-99.

45. Kemble, D.J. and G. Sun, *Direct and specific inactivation of protein tyrosine kinases in the Src and FGFR families by reversible cysteine oxidation.* Proc Natl Acad Sci U S A, 2009. **106**(13): p. 5070-5.

46. Huang, K., et al., *Identification of N-terminal lobe motifs that determine the kinase activity of the catalytic domains and regulatory strategies of Src and Csk protein tyrosine kinases.* J Mol Biol, 2009. **386**(4): p. 1066-77.

Figure Legends:

Figure 1. Protein tyrosine phosphorylation by PTKs when expressed in *E coli*. (A) Protein phosphorylation by Abl, Csk, FGFR1 and Lck. As described in the Materials and Methods section, *E. coli* cells expressing these kinases were collected from overnight cultures and analyzed by Western Blotting with a monoclonal antibody specific for phosphotyrosine residues. The same amount of soluble cell lysates for each PTK sample were compared. (B) Protein tyrosine phosphorylation by Src when it is co-expressed with a protein tyrosine phosphatase in *E. coli*. The wtSrc and Src mutants were co-expressed with GroESL in the PTP system. The wtSrc, kdSrc, and SΔD-Src were expressed at 37^0C for 24 hours. The same amount of soluble lysates were analyzed by Western Blotting with monoclonal antibodies specific for phosphotyrosine and phosphor-Tyr416, respectively.

Figure 2. Schemes of library construction and proposed screening strategy. A. Construction of SΔD-Src and HRD mutant libraries. B. Colony screening. Individual steps are depicted and labeled, and explained in detail in Materials and Methods.

Figure 3. Representative screening results of His384 mutation library. The His384 mutatant library was constructed as described in Materials and Methods. Overnight cultures (0.4 μl) of control colonies and unknown colonies were spotted on the nitrocellulose membrane in duplicate for the dot blotting screen. (A) A representative screening blot of the His384 library in the PTP system: boxes a1, a2, d1 and d2 have wtSrc-expressing cells as positive controls; boxes b1, b2, e1 and e2 have kdSrc-expressing cells as negative controls; boxes c1, c2, f1 and f2 have SΔD-Src-expressing cells as negative controls. Other boxes have colonies expressing random His384 mutants. (B) A representative screening blot of the His384 library in the non-PTP system. The arrangement of colonies is the same as in Fig. 3A.

Figure 4. Relative catalytic activity of Src His384 mutants. A. Coomassie blue staining of selected purified His384 mutant proteins. B. The catalytic activity of His384 mutants toward an artificial substrate, polyE_4Y was determined. The wtSrc activity (210 min^{-1}) is taken as 100%, and the activity of all other mutants was expressed as a percentage of that of the wtSrc. The kinase activities of each mutant can be found in the Supplementary Information. The wtSrc is represented by His as the wtSrc has His at the 384 position, and all other variants are represented by the residue at this position.

Figure 5. Representative screening results of the Arg385 mutant library. The Arg385 mutant library was constructed as described in the Materials and Methods section. Colony screening was performed as described in the legend to Fig. 3. (A) A representative screening blot of the Arg385 library in the PTP system. Boxes a1 and a2 have wtSrc expressing cells as positive controls; Boxes b1 and b2 have kdSrc-expressing cells as negative controls; Boxes c1 and c2 have SΔ-Src expressing cells as negative controls. The other boxes have random Arg385 mutant cells. (B) A representative screening blot of the Arg385 library in the non-PTP system. Colony arrangement is the same as in Fig. 5A.

Figure 6. Relative catalytic activity of Src Arg385 mutants. A. Coomassie staining of purified Arg385 mutants fractionated on SDS-PAGE. B. The kinase activity of selected Arg385 mutants toward an artificial substrate, polyE_4Y was determined. The assay was performed as described in Materials and Methods and the legend to Fig. 4. Detailed information on the kinase mutants is given in Supplementary Information.

Figure 7. Representative screening results of the Asp386 mutant library. The Asp386 mutant library was constructed as described in Materials and Methods. (A) A representative screening

blot of the Asp386 library in the PTP system. Boxes a1, a2, d1 and d2 have wtSrc-expressing cells as positive controls; Boxes b1, b2, e1 and e2 have kdSrc-expressing cells as negative controls; boxes c1, c2, f1 and f2 have SΔ-Src-expressing cells as negative controls. Other boxes have random Asp386 mutant-expressing colonies. (B) A representative screening blot of Asp386 library in the non-PTP system. The colony arrangement is identical to that in Fig. 7A.

Fig. 1.

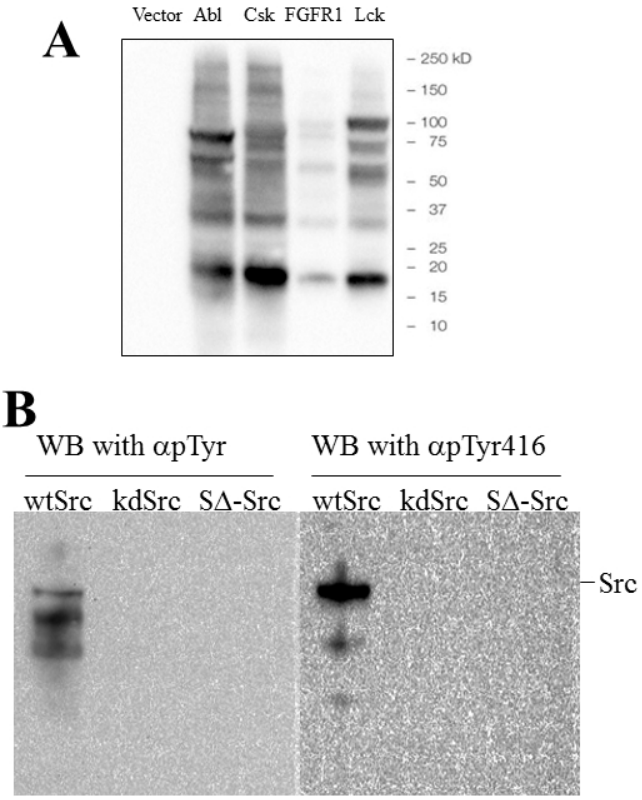
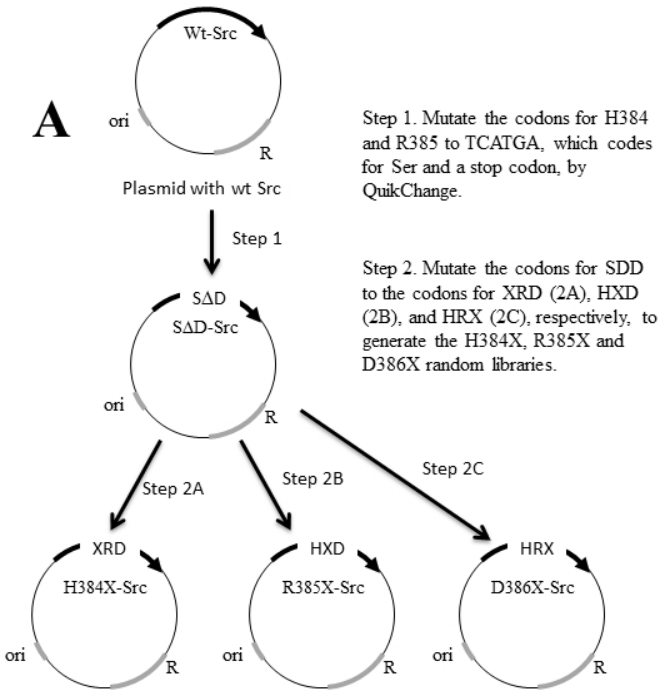


Fig. 1

190x254mm (96 x 96 DPI)

Fig. 2A



Figu. 2A

190x254mm (96 x 96 DPI)

Fig. 2B.

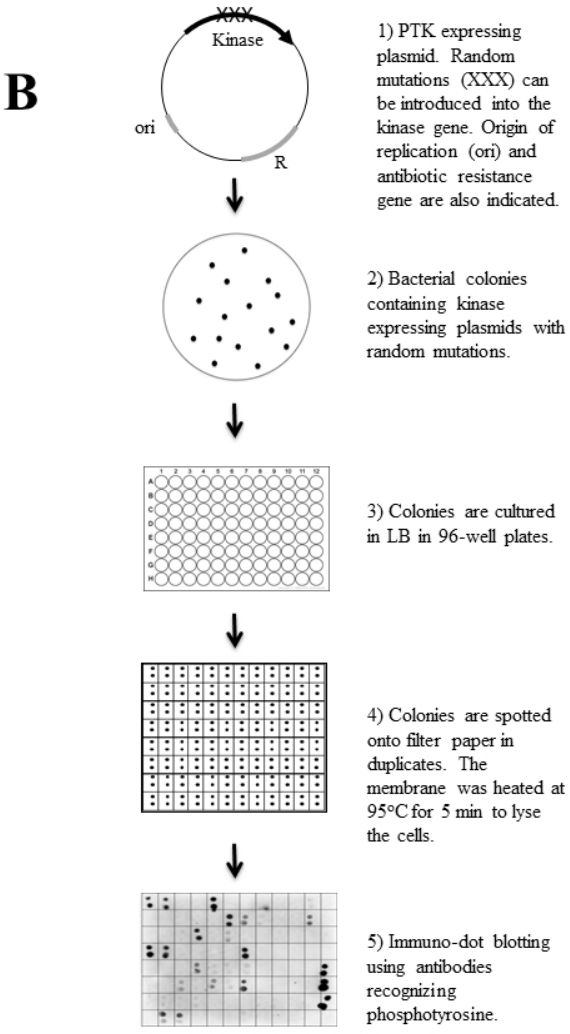


Fig. 2B

190x254mm (96 x 96 DPI)

Fig. 3.

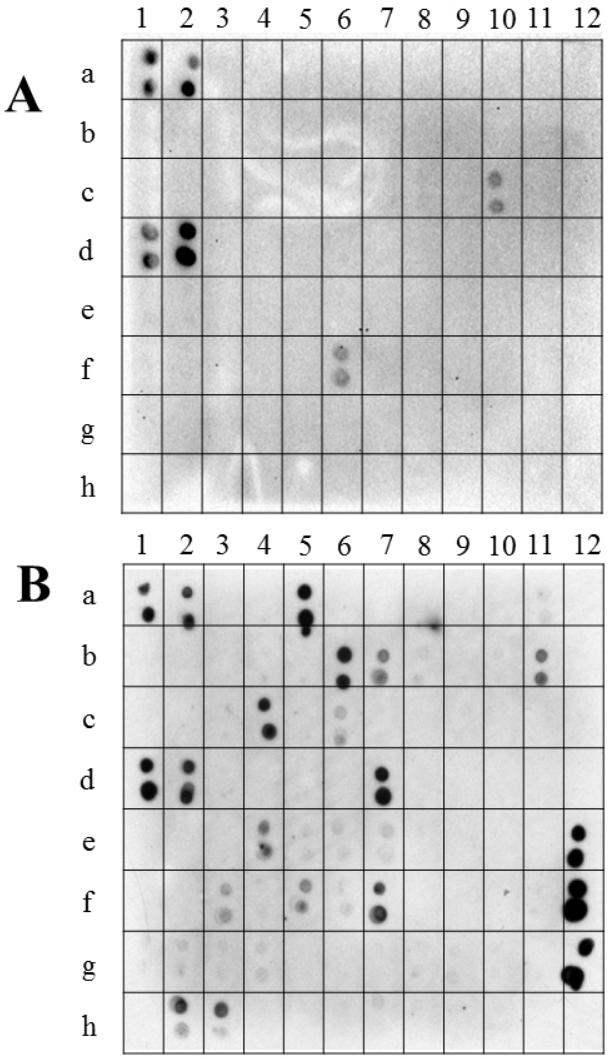


Fig. 3

190x254mm (96 x 96 DPI)

Fig. 4

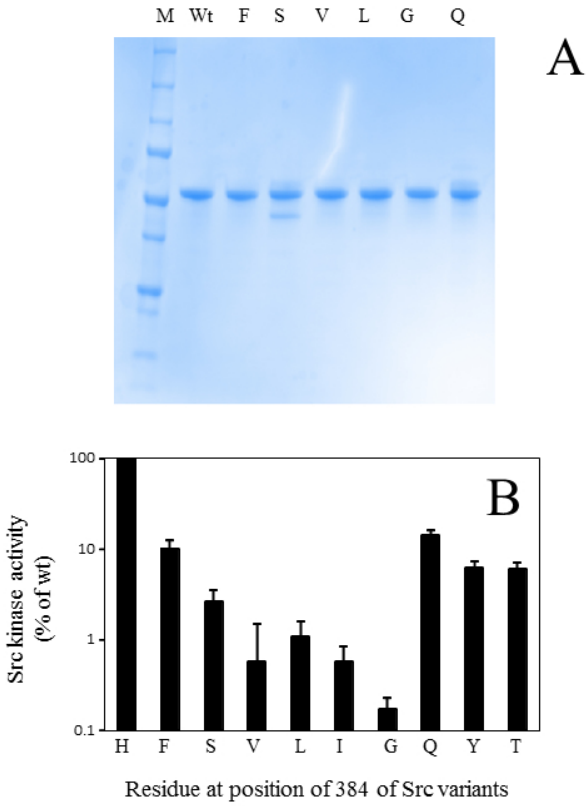


Fig. 4

190x254mm (96 x 96 DPI)

Fig. 5.

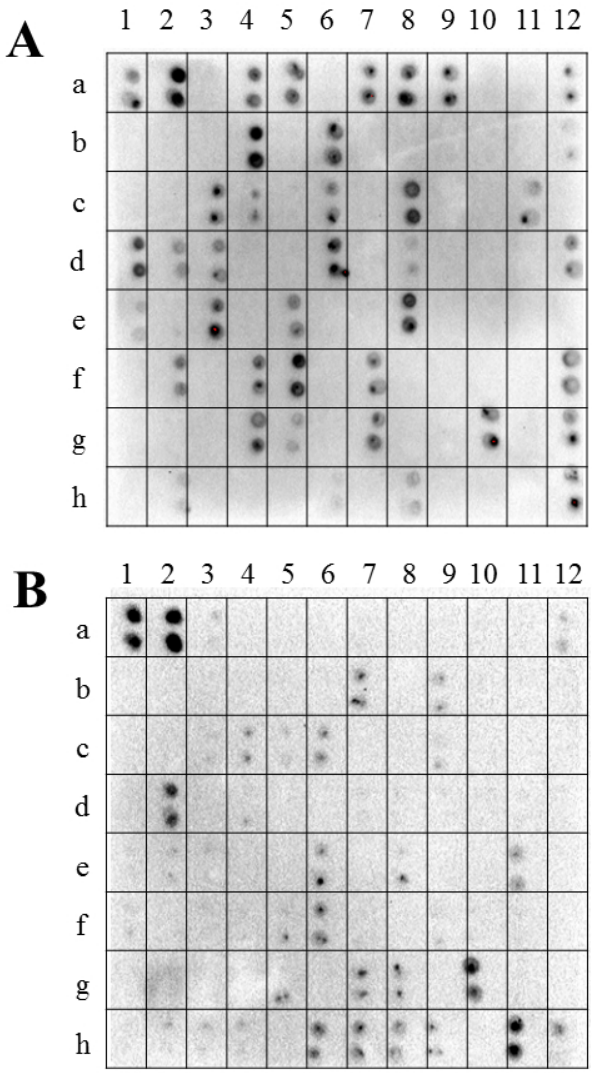


Fig. 5

190x254mm (96 x 96 DPI)

Fig. 6.

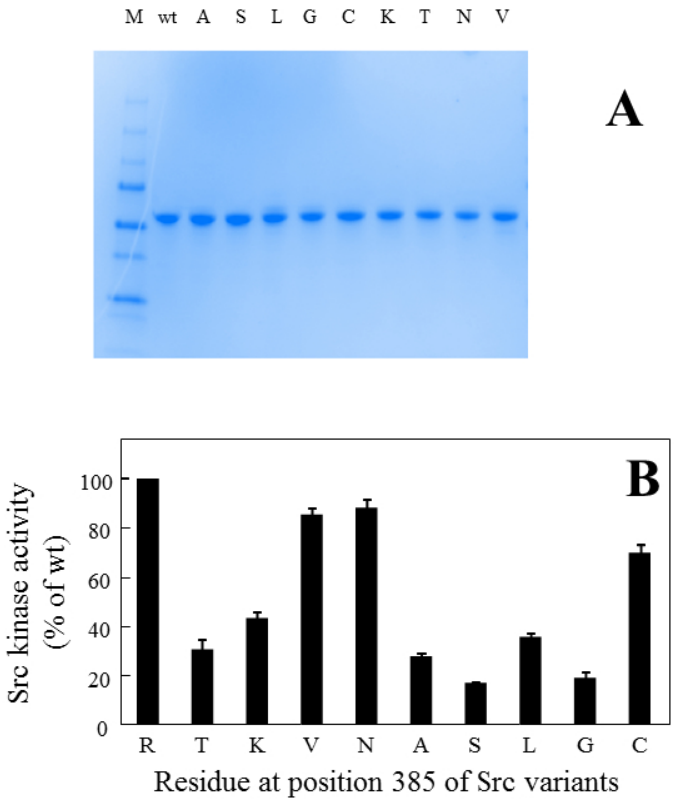


Fig. 6

190x254mm (96 x 96 DPI)

Fig. 7.

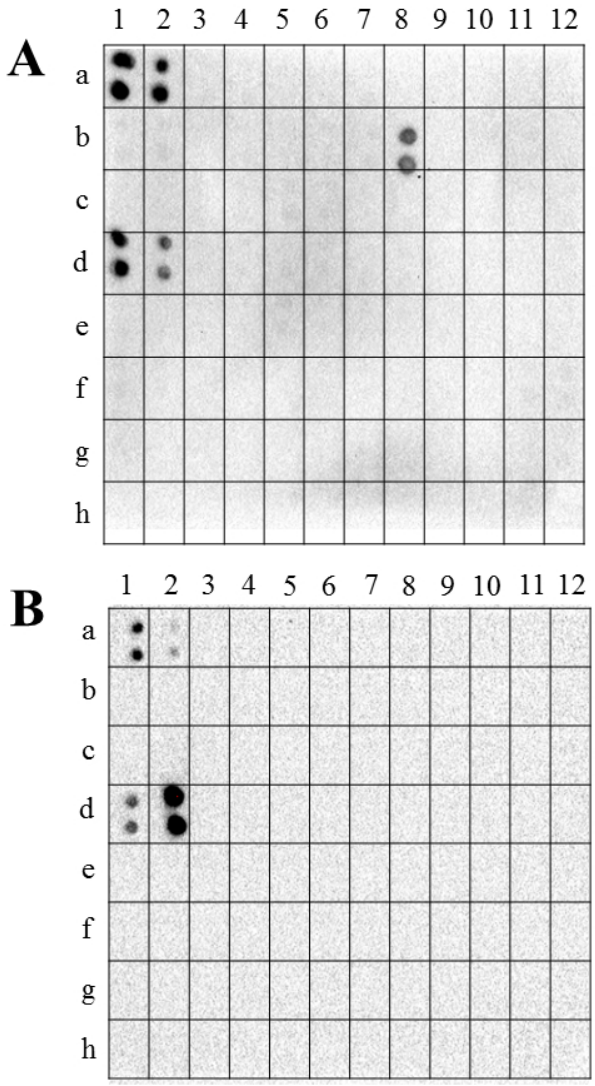


Fig. 7

190x254mm (96 x 96 DPI)

Cui Y. and Sun G. Supplementary Information

Kinase activity determination for His384 mutants. (Complementary to Figure 4)

	wtsrc	R385T	R385K	R385V	R385N	R395A	R385S	R385L	R385G	R385C
Assay 1 (mi	200.25	62.43	100.15	187.16	182.38	55.66	36.57	75.00	40.02	155.54
Assay 2 (mi	207.50	79.77	82.78	181.76	197.85	62.78	36.26	80.12	48.14	150.71
Assay 3 (mi	222.10	51.75	91.51	169.28	176.15	57.59	35.54	70.89	33.18	134.79
mean	209.95	64.65	91.48	179.40	185.46	58.68	36.13	75.34	40.45	147.02
standard er	6.43	8.16	5.01	5.29	6.45	2.13	0.30	2.67	4.32	6.27
Relative act	100.00	30.79	43.56	85.43	88.31	27.94	17.20	35.87	19.26	70.01
standard er	0.00	3.89	2.39	2.52	3.07	1.01	0.14	1.27	2.06	2.98

Kinase activity determination for Arg385 mutants. (Complementary to Figure 6)

	wtsrc	H384F	H384S	H384V	H384L	H384I	H384G	H384Q	H384Y	H384T
Assay 1 (mi	200.25	21.23	9.28	-2.57	1.43	2.34	0.51	34.24	17.08	12.70
Assay 2 (mi	207.50	13.18	4.76	3.53	0.94	0.34	0.13	23.21	8.58	8.09
Assay 3 (mi	222.10	29.63	2.56	2.67	4.43	0.97	0.45	32.49	13.51	16.91
mean	209.95	21.35	5.53	1.21	2.27	1.22	0.36	29.98	13.06	12.57
standard er	6.43	4.75	1.98	1.91	1.09	0.59	0.12	3.42	2.47	2.55
Relative act	100.00	10.17	2.63	0.57	1.08	0.58	0.17	14.28	6.22	5.98
standard er	0.00	2.26	0.94	0.91	0.52	0.28	0.06	1.63	1.17	1.21