



## OPEN ACCESS

## EDITED BY

Junmei Wang,  
University of Pittsburgh, United States

## REVIEWED BY

Thibault Viennet,  
Aarhus University, Denmark  
Yoonji Lee,  
Chung-Ang University, Republic of Korea

## \*CORRESPONDENCE

Jeanine F. Amacher,  
✉ amachej@wwu.edu

†These authors have contributed equally  
to this work

RECEIVED 27 August 2024

ACCEPTED 14 November 2024

PUBLISHED 03 December 2024

## CITATION

Longshore-Neate F, Ceravolo C, Masuga C,  
Tahti EF, Blount JM, Smith SN and Amacher JF  
(2024) The conformation of the nSrc  
specificity-determining loop in the Src SH3  
domain is modulated by a WX conserved  
sequence motif found in SH3 domains.  
*Front. Mol. Biosci.* 11:1487276.  
doi: 10.3389/fmolb.2024.1487276

## COPYRIGHT

© 2024 Longshore-Neate, Ceravolo, Masuga,  
Tahti, Blount, Smith and Amacher. This is an  
open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with  
these terms.

# The conformation of the nSrc specificity-determining loop in the Src SH3 domain is modulated by a WX conserved sequence motif found in SH3 domains

Frederick Longshore-Neate<sup>†</sup>, Caroline Ceravolo<sup>†</sup>, Cole Masuga, Elise F. Tahti, Jadon M. Blount, Sarah N. Smith and Jeanine F. Amacher\*

Department of Chemistry, Western Washington University, Bellingham, WA, United States

Cellular signaling networks are modulated by multiple protein-protein interaction domains that coordinate extracellular inputs and processes to regulate cellular processes. Several of these domains recognize short linear motifs, or SLiMs, which are often highly conserved and are closely regulated. One such domain, the Src homology 3 (SH3) domain, typically recognizes proline-rich SLiMs and is one of the most abundant SLiM-binding domains in the human proteome. These domains are often described as quite *versatile*, and indeed, SH3 domains can bind ligands in opposite orientations dependent on target sequence. Furthermore, recent work has identified diverse modes of binding for SH3 domains and a wide variety of sequence motifs that are recognized by various domains. Specificity is often attributed to the RT and nSrc loops near the peptide-binding cleft in this domain family, particularly for Class I binding, which is defined as RT and nSrc loop interactions with the N-terminus of the ligand. Here, we used the Src and Abl SH3 domains as a model to further investigate the role of the RT and nSrc loops in SH3 specificity. We created chimeric domains with both the RT and nSrc loop sequences swapped between these SH3 domains, and used fluorescence anisotropy assays to test how relative binding affinities were affected for Src SH3- and Abl SH3-specific ligands. We also used AlphaFold–Multimer to model our SH3:peptide complexes in combination with molecular dynamics simulations. We identified a position that contributes to the nSrc loop conformation in Src SH3, the amino acid immediately following a highly conserved Trp that creates a hydrophobic pocket critical for SH3 ligand recognition. We defined this as the WX motif, where X = Trp for Src and Cys for Abl. A broad importance of this position for modulating nSrc loop conformation in SH3 domains is suggested by analyses of previously deposited SH3 structures, multiple sequence alignment of SH3 domains in the human proteome, and our biochemical and computational data of mutant Src and Abl SH3 domains. Overall, our work uses experimental approaches and structural modeling to better understand specificity determinants in SH3 domains.

## KEYWORDS

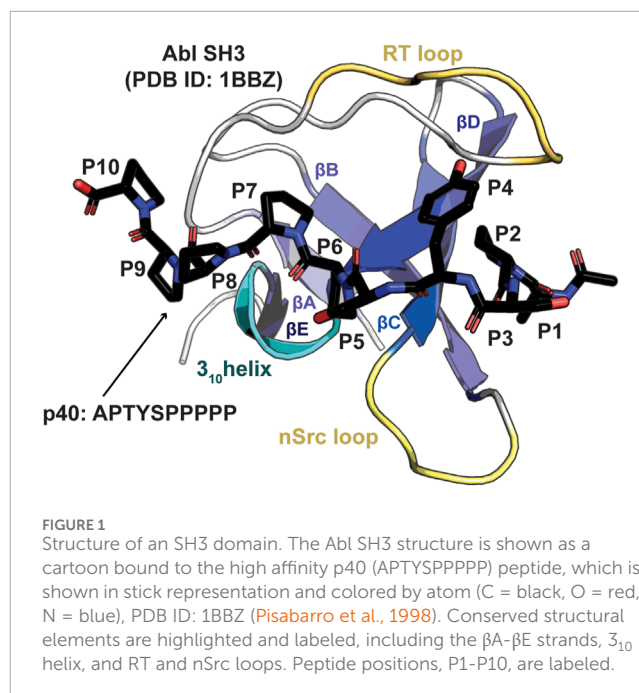
SH3 domains, protein-protein interactions, protein-peptide interactions, short linear motifs, specificity, signal transduction

## Introduction

Short linear motif (SLiM) or peptide binding is a critical component of signal transduction pathways, often with several SLiM-binding domains in multiple proteins modulating the activity and regulation of the signaling cascade (Pawson and Nash, 2003). One such domain, the Src-homology 3 (SH3) domain, was first described in 1988 as a domain located on the proto-oncogene Src tyrosine kinase ((Mayer et al., 1988; Stahl et al., 1988)). It is now recognized that SH3 domains exist in all kingdoms of life and viruses, and there are over 300 SH3 domains in 200 proteins in the human proteome ((Mehrabipour et al., 2023; Teyra et al., 2017)). Although several noncanonical exceptions exist, SH3 domains are generally characterized as binding to proline-rich sequences, specifically containing a PXXP motif (where X = any amino acid), with affinities in the low-to-mid micromolar range [(Teyra et al., 2017; Pisabarro and Serrano, 1996)]. Many SH3 target sequences adopt a type II polyproline helix (PPII) structure, presenting a hydrophobic surface to which the SH3 domain recognizes and binds [(Lim et al., 1994; Mayer, 2001)].

SH3 domains share a general conserved structure, despite displaying a relatively large degree of plasticity in ligand binding. SH3 domains are typically about 60 amino acids in length, with a  $\beta$ -barrel fold, consisting of approximately 5  $\beta$ -strands and a  $3_{10}$  helix [(Mehrabipour et al., 2023; Kaneko et al., 2008a)] (Figure 1). Specificity in SH3 domain binding is determined by the nSrc, RT, and  $\beta 4$ - $\alpha 3_{10}$  loops (Mehrabipour et al., 2023). Interestingly, the same SH3 domain can bind ligands in opposite orientations, depending on whether it corresponds to a Class I/“plus” (consensus sequence: RXLPXP, where X = any amino acid) or Class II/“minus” (XPPLPXR) target sequence [(Lim et al., 1994), (Feng et al., 1994; Kurochkina and Guha, 2013; Saksela and Permi, 2012)]. In the Class I orientation, N-terminal residues of the ligand interact directly with the RT and nSrc loops; in the Class II orientation, these regions interact with C-terminal residues of the ligand [(Lim et al., 1994; Feng et al., 1994)]. Thus, the same SH3 domain can bind ligands N- to C-terminal or *vice versa* depending on the target sequence.

The *Src module*, consisting of an SH3 domain, SH2 domain, and tyrosine kinase, is shared amongst several families of cytoplasmic tyrosine kinases (Shah et al., 2018). This includes Abl, a Src-related tyrosine kinase, which regulates actin (Colicelli, 2010). The kinase domains of Src and Abl are 46% identical, but these proteins are known to be differentially regulated [(Seeliger et al., 2007; Sicheri et al., 1997; Eck et al., 1994; Nagar et al., 2003)]. Furthermore, therapeutics that target the BCR-Abl fusion protein, which is the underlying cause of disease in approximately 95% of chronic myelogenous leukemia (CML) cases, e.g., imatinib, bind and inhibit Abl kinase, but do not bind Src kinase (Seeliger et al., 2007). There are also specificity differences in the SH3 domains of these proteins, which are 43% identical over 58 residues. A previous study of SH3 domain specificity investigated an initial ligand, derived from the 3BP1 protein (sequence: RAPTMPPLPP), which bound the Abl and Fyn (a Src family kinase), the SH3 domains of Fyn and Src share 77% sequence identity) SH3 domains with similar 30  $\mu$ M affinity [(Pisabarro and Serrano, 1996; Pisabarro et al., 1998)]. This sequence was then used as a template to design a selective and high affinity Abl SH3 ligand, termed p40 [(Pisabarro and Serrano, 1996; Pisabarro et al., 1998)]. The Class I p40



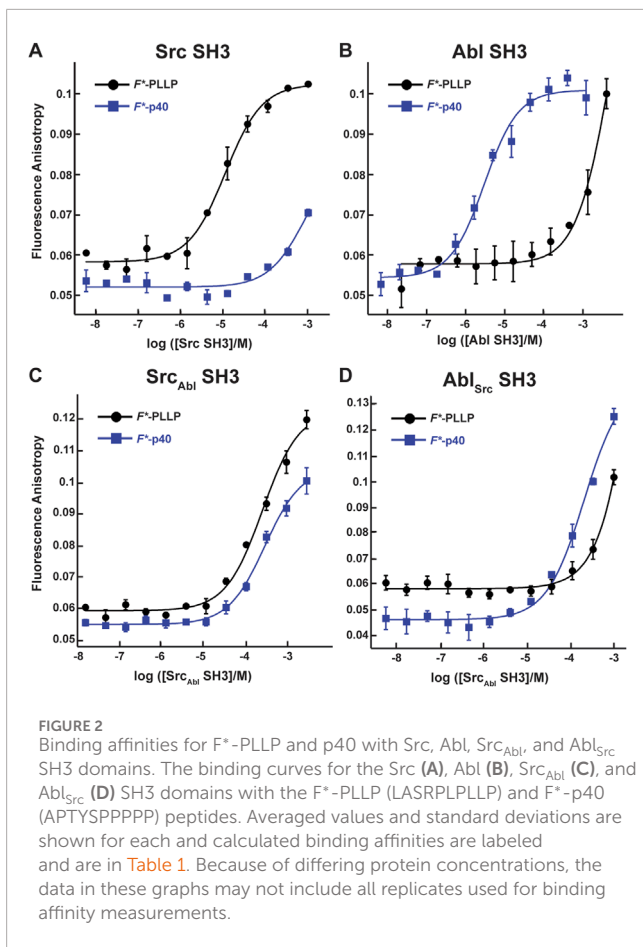
ligand (APTYSPPPPP) bound the Abl and Fyn SH3 domains with 0.4  $\mu$ M and 470  $\mu$ M affinity, respectively [(Pisabarro and Serrano, 1996; Pisabarro et al., 1998)].

Here, we aimed to better understand this result by expanding this investigation to Abl and Src SH3 specificity using chimeric proteins, Abl<sub>Src</sub> and Src<sub>Abl</sub>, which swap both the RT and nSrc loop sequences from the other domain. We calculated binding affinities of these chimeric domains with the Class I p40 ligand, as well as a Class I Src-specific target sequence, LASRPLPLL, termed PLLP (Sparks et al., 1996). We found that while the WT Src and Abl SH3 domains are specific for their ligands, the Src<sub>Abl</sub> chimeric protein revealed relatively weak but similar binding to both peptide sequences. To visualize the tested SH3:ligand interactions, we used AlphaFold–Multimer to model the complexes and analyze potential binding interfaces, in combination with molecular dynamics simulations for Src SH3 variants. We also tested single mutations of a WW (for Src SH3) and WC (for Abl SH3) sequence motif at the C-terminus of the nSrc loop in the wild-type and chimeric proteins, showing that these can modulate ligand affinity several fold. These WX positions in Src SH3, the amino acids W121 and W122, have previously been shown to be important for establishing a hydrophobic binding specificity pocket and modulating lipid binding [(Sipeki et al., 2021; Le Roux et al., 2019)]. Here, we argue that the W122 position, or X of the WX motif, additionally determines the conformation (and potentially the flexibility as well) of the nSrc loop. Taken together, our studies provide further insight into the role of these specificity-determining loops for the Src and Abl SH3 domains.

## Results

### Differing specificities of Src and Abl SH3 domains

To investigate the effect of RT and nSrc loop variation on the Src and Abl SH3 domains, we first wanted to confirm that these domains



**TABLE 1** Relative binding affinities using fluorescence anisotropy for Src- and Abl-derived SH3 domains.

	$K_d$ ( $\mu$ M)	
	F*-PLL	F*-p40
Src SH3	9.7 ± 0.9	>1,000
Src <sub>Abl</sub> SH3	140 ± 20	130 ± 40
W122C Src SH3	810 ± 160	>1,000
W122C Src <sub>Abl</sub> SH3	>1,000	450 ± 190
Abl SH3	>1,000	3.1 ± 0.4
Abl <sub>Src</sub> SH3	>1,000	210 ± 30
C100W Abl SH3	>1,000	55 ± 20

sequence, i.e., our peptides also include N-terminal fluorescein and Ahx linkers, and the use of Fyn SH3 as opposed to Src SH3, whose SH3 domains are 77% identical by sequence, as stated previously. Despite variation in absolute values, our results also indicated a difference of approximately three orders of magnitude for Abl SH3 binding to F\*-p40 as compared to Src/Fyn SH3, indicating internal consistency with the previous results. Taken together, our results confirmed differing ligand specificities for the Abl and Src SH3 domains.

### Chimeric SH3 domains confer differences in ligand specificities

Previous reports indicated that the RT and nSrc loops of SH3 domains drive ligand specificity (Mehrabipour et al., 2023). Therefore, we wanted to engineer chimeric proteins to test if these loops alone could mediate specificity switching. In our chimeras, we swapped both the RT (sequence: <sup>96</sup>ESRTET<sup>101</sup>) and nSrc (<sup>117</sup>TEGD<sup>120</sup>) loops of Src with the RT (<sup>73</sup>VASGDN<sup>78</sup>) and nSrc (<sup>95</sup>HNGE<sup>98</sup>) loop sequences of Abl, and vice versa in each SH3 domain. We termed the resulting proteins, Src<sub>Abl</sub> and Abl<sub>Src</sub>, where the subscript indicates the sequence identity of both the RT and nSrc loops. These proteins were also recombinantly expressed and purified as SUMO fusions and we determined binding affinities using fluorescence anisotropy experiments with our F\*-PLL and F\*-p40 peptides, as described in the Materials and Methods (Figure 2; Table 1).

In our Src<sub>Abl</sub> SH3 chimera, binding affinity for the Src-specific F\*-PLL peptide was reduced ~14-fold, to  $K_D = 140 \pm 20 \mu$ M, as compared to the wild-type Src SH3 domain (Figure 2A; Table 1). Interestingly, Src<sub>Abl</sub> SH3 bound the Abl-specific F\*-p40 peptide with similar affinity,  $K_D = 130 \pm 40 \mu$ M, indicating that specificity was indeed altered by this substitution (Figure 2C; Table 1). In contrast, while binding to the Abl-specific F\*-p40 peptide was reduced by > 60-fold for Abl<sub>Src</sub> SH3,  $K_D = 210 \pm 30 \mu$ M, this chimeric protein continued to show undetectable binding to the Src-specific F\*-PLL peptide,  $K_D > 1,000 \mu$ M (Figure 2D; Table 1).

contain differing ligand specificities. We recombinantly expressed and purified the Src and Abl SH3 domains as SUMO-tagged fusion proteins, as described in the Materials and Methods. All sequences used in this work are in the Supporting Information. We then used fluorescence anisotropy experiments, as described previously and in the Materials and Methods, to test the binding affinities of each domain with two fluorescein (F\*)-tagged peptides, one corresponding to the p40 sequence, F\*-Ahx-APTYSPPPPP, termed F\*-p40, and to a Src SH3 identified ligand, F\*-Ahx-LASRPLPLL, termed F\*-PLL (Figure 2; Table 1) [(Pisabarro and Serrano, 1996; Pisabarro et al., 1998; Sparks et al., 1996)]. For clarity, we will refer to peptide positions numerically from the N-terminus and include one-letter amino acid codes for identification, e.g., for p40, A1-P2-T3-etc.

Consistent with previous work, our results indicated that while Src SH3 bound F\*-PLL with relatively high affinity,  $K_D = 9.7 \pm 0.9 \mu$ M, it did not bind F\*-p40, here defined as  $K_D > 1,000 \mu$ M (Figure 2A; Table 1). In contrast, Abl SH3 bound F\*-p40 with  $K_D = 3.1 \pm 0.4 \mu$ M, and F\*-PLL with  $K_D > 1,000 \mu$ M (Figure 2B; Table 1). Notably, our Abl SH3 affinity for F\*-p40 differs by an order of magnitude as compared to the previously published value,  $0.4 \pm 0.1 \mu$ M (for reference, in this paper Fyn SH3 bound with  $K_D = 472 \pm 55 \mu$ M) [(Pisabarro and Serrano, 1996; Pisabarro et al., 1998)]; however, we attribute these differences to the experimental methods used, variations in

## Structural analyses of SH3-ligand models

To better understand our biochemical assay results, we generated structural models of our SH3-ligand complexes using AlphaFold–Multimer on the ColabFold server (Kaneko et al., 2008b; Donaldson et al., 2002). We created models of wild-type Src and Abl SH3 domains with each peptide, PLLP and p40, as well as our chimeric Src<sub>Abl</sub> and Abl<sub>Src</sub> SH3 proteins with each peptide (Figures 3A, B; Table 2). Because of relatively poor confidence measures, we were unable to analyze models of Abl<sub>Src</sub> bound to either peptide (Table 2). Overall, the 5 output models generated for each complex were internally consistent, with alignments of <0.18 Å in all cases (Figures 3A, B). However, there were notable exceptions, which included steric clash based on the PyMOL library for several of the peptides with SH3 domains. Therefore, our analyses will be based on models without steric clash.

We observed RT and nSrc loop differences between the Abl and Src SH3 domains, however, our Src<sub>Abl</sub> chimeric protein loops are consistent with the Src parent/wild-type protein (Figure 3C). The PXXP motifs at the C-terminus of each peptide (PLLP and PPPP, respectively, and referred to as positions P7–P10, as described above and based on amino acid identity) interact largely with the βA–βB loop and side-chains in the <sub>310</sub> helix in a consistent manner between the two peptide sequences and other SH3 domains (Pisabarro et al., 1998). All PxxP-interacting residues are conserved between Abl and Src SH3 (Figure 3D). At the N-terminal end of the peptide, we observed differing peptide conformations for the first 6 amino acids for each peptide sequence (Figures 3E, F), with an upward translation of p40 binding as compared to PLLP, including for the Src<sub>Abl</sub> chimera (Figure 3C).

As others have previously identified as well, these translation differences in peptide binding are likely due to an electrostatic interaction between the R4 Arg of the PLLP sequence and D102 Src SH3, using full-length Src numbering (Figure 4A). This interaction is a well-studied characteristic of Src SH3 specificity (Saksela and Permi, 2012; Shah et al., 2018; Musacchio et al., 1994), and appears to result in the lowered translation of the peptide in the binding cleft (Figure 4B). The D102 amino acid immediately follows the RT loop. In Abl, this position corresponds to T79. Because of the R4 residue, an upward translation of PLLP in the Src SH3 binding cleft would result in steric clash with the RT loop, as the main chain atoms of the peptide are translated ~2 Å between the two peptide conformations (Figure 4B). Our modeling results also revealed that the RT and nSrc loops of Src<sub>Abl</sub> are in a Src-like conformation, despite containing sequences from Abl SH3 (Figure 4C).

## The WX sequence motif at the base of the nSrc loop influences nSrc loop orientation

Although the peptides bind in differing orientations with respect to the peptide-binding cleft, we wanted to further investigate the relative conformation of the RT and nSrc loops in Src versus Src<sub>Abl</sub>, which include the sequences derived from Src or Abl, respectively (Figure 4C). Structural analyses of our AlphaFold-generated models suggested that an intra-SH3 interaction on either side of the nSrc loop may modulate its loop conformation, independent of the sequence identity of the loop itself. Specifically, an aromatic

Trp (W122) C-terminal to the nSrc loop is counterbalanced by an aromatic Tyr (Y93) N-terminal to the nSrc loop in Abl SH3 (Figure 5A). Because these positions are outside the nSrc loop, we reasoned this may be why the Abl sequence in Src<sub>Abl</sub> maintains a Src-like nSrc loop conformation in our models (Figure 4C). As we do not see specific W122 interactions with the side chain atoms of N116, we predicted that the W122 in Src SH3 (analogous position is C100 in Abl SH3) may be primarily responsible for this observation.

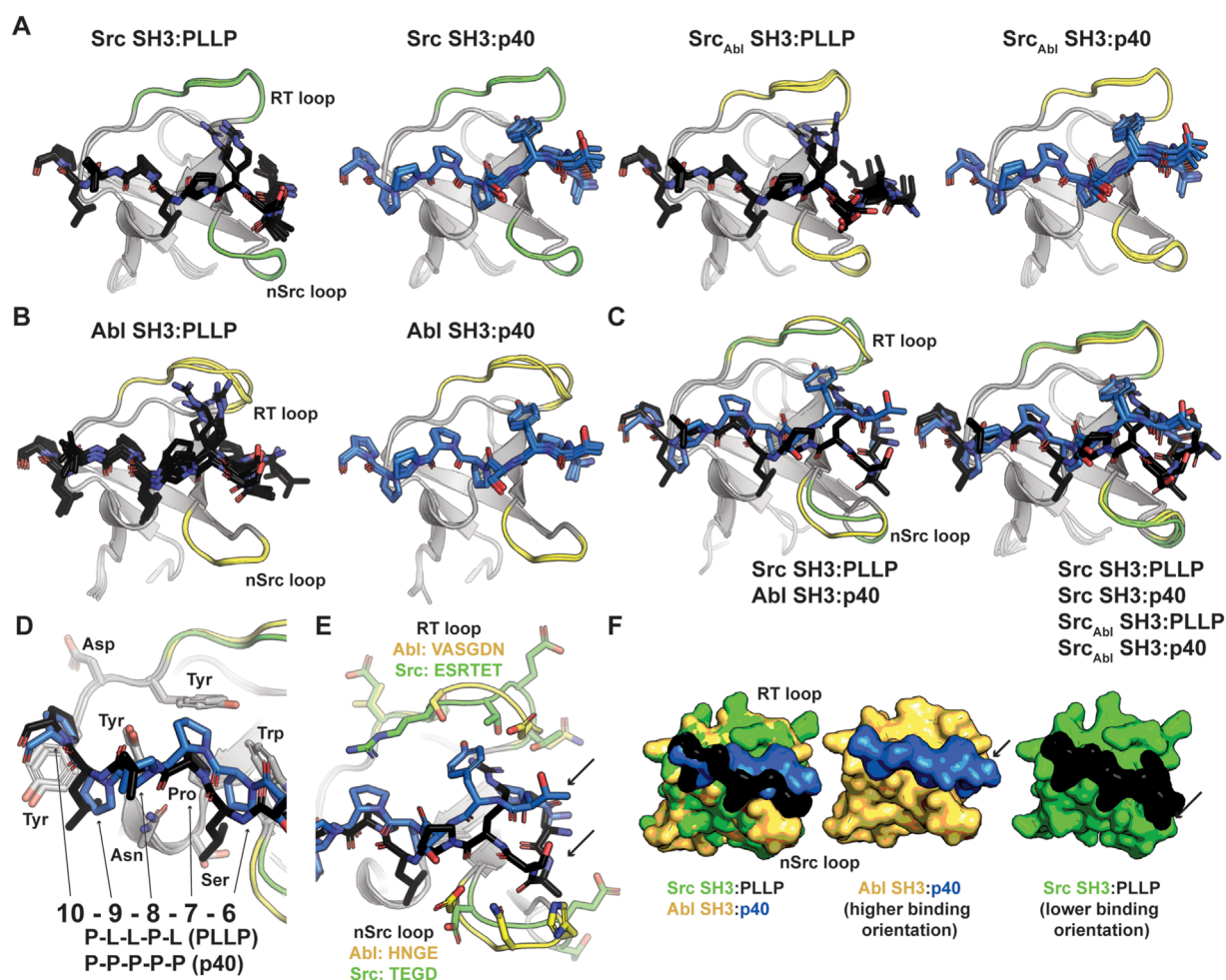
To test this hypothesis, we recombinantly expressed and purified a W122C mutation in Src SH3 and Src<sub>Abl</sub> SH3, as described in the Materials and Methods. While W122C Src SH3 showed little to no binding for either peptide, the W122C Src<sub>Abl</sub> SH3 domain was able to bind F\*–p40 weakly,  $K_D = 450 \pm 190 \mu\text{M}$  (Figure 5B; Table 1). We also recombinantly expressed and purified C100W Abl SH3 and C100W Abl<sub>Src</sub>; however, the C100W Abl<sub>Src</sub> SH3 protein was unstable in our purification. We saw that C100W Abl SH3 revealed undetectable binding to F\*–PLLP, defined as >1,000 μM in our assay. Binding of C100W Abl SH3 to F\*–p40 was weaker by ~18-fold as compared to the wild-type protein,  $K_D = 55 \pm 20 \mu\text{M}$ , but was still relatively high for the SH3 domain interactions tested (Figure 5B; Table 1).

Protein structure prediction models generated using AlphaFold–Multimer show differences for nSrc loop conformation in the W122C Src SH3 and C100W Abl SH3 mutations. Based on our models, the W122C mutation in Src SH3 allows the nSrc loop to adopt a more “Abl-like” conformation, although it is still mostly “Src-like” (Figure 5C). In contrast, we see no difference in the nSrc loop for a C100W Abl SH3 model as compared to the wild-type protein (Figure 5D). This is consistent with our binding affinity results.

## Molecular dynamics simulations investigate the relative flexibility in the nSrc loop upon mutation of the WX motif

Because we were unable to generate an Abl SH3 variant that showed binding to the Src target, PLLP, and based on our C100W Abl SH3 results, we decided to focus on nSrc loop conformational differences in Src SH3 for additional study. To investigate the relative stability of our AlphaFold models and further investigate the WX motif, we ran 1 μs molecular dynamics (MD) simulations of several of our Src SH3-peptide complexes, including (with relative binding affinities, as reported in Table 1, in parentheses): Src–PLLP (9.3 μM), Src–p40 (>1,000 μM), Src<sub>Abl</sub>–PLLP (140 μM), Src<sub>Abl</sub>–p40 (130 μM), W122C Src–PLLP (810 μM), and W122C Src<sub>Abl</sub>–p40 (450 μM) (Figure 6A; Supplementary Figure S1).

The Src SH3 domain backbone remained relatively stable during the length of all simulations, as assessed by calculated root mean square deviation (RMSD) values, compared to the initial structure (Supplementary Figure S1A). Alignment of 200 states from each simulation, including a state every 5 ns, revealed that in general, the PLLP peptide is more stable in the pocket than p40 (Figure 6A). Because all but one of these interactions are of relatively low affinity (>100 μM), it is challenging to use these MD simulations (or the AlphaFold models) to make predictions about why one peptide will bind over another. For example, we included the Src SH3:p40 complex as a control (affinity >1,000 μM or undetectable in our



**FIGURE 3** AlphaFold–Multimer (ColabFold) models of Src, Abl, and SrcAbl SH3 domains with F\* -PLL and p40 ligands. **(A, B)** AlphaFold–Multimer results for Src and SrcAbl **(A)** or Abl **(B)** SH3 domains with the PLLP (black sticks, colored by atom) or p40 (blue sticks, colored by atom) ligands. In all cases, the results for each complex align with RMSD values of <1.75 Å, with many even closer [e.g., Abl SH3:PLL aligns with RMSD <1.3 Å]. SH3 domains are shown in cartoon representation with the RT and nSrc loops colored based on the origin sequence, green for Src SH3 and yellow for Abl SH3. The PLLP and p40 peptides are shown in stick representation and colored by atom (C = black (PLL) or blue (p40)). **(C)** Although the Src and Abl RT and nSrc loops are in distinct conformations (left figure), alignment of the SrcAbl chimeric SH3 domain with the wild-type Src SH3 domain indicates that the loop conformations largely match the wild-type scaffold. Structures are rendered and colored as in **(A, B)**. **(D)** SH3 residues in Src and Abl that interact with the 6–10 ligand positions are identical. SH3 domains are in cartoon representation with side chain sticks for interacting residues (C = gray). Peptide positions are labeled. **(E)** The p40 peptide binds in a “higher” orientation and PLLP in a “lower” orientation. In the left figure, the SH3 domains are shown in cartoon, with side chain atoms of the RT and nSrc loops shown as sticks (C = green (Src) or yellow (Abl)). The peptides are rendered as in **(A–D)**. **(F)** Surface representation of **(E)**. Here, arrows indicate the differing peptide orientations for PLLP (black, with Src in green) and p40 (blue, with Abl in yellow). In all stick representation images, O = blue, N = red.

assay) (Table 1); however, our MD simulation does show binding, although the 1-6 positions of the peptide are relatively unstable in the pocket (Figure 6A). This variability is also observed in the W122C Src<sub>Abl</sub>-p40 simulation (Figure 6A). These factors aside, we do conclude that PLLP is more stable in the binding pocket because of the R4 Arg-D102 interaction previously described (Figure 4A).

Ultimately, our MD simulations do allow us to assess relative flexibility in the nSrc loop upon mutation. Notably, we see the largest root mean square fluctuations (RMSF), as compared to the average, in Cα atoms for specific RT and nSrc loop residues in all simulations (Supplementary Figure S1B). In the RT loop, the largest

RMSF values were seen for R98 in the Src-PLL and W122C Src-PLL simulations (Supplementary Figure S1B). In the nSrc loop, the largest RMSF values were seen for E118 (or E118N in Src<sub>Abl</sub>) in Src<sub>Abl</sub>-PLL, Src<sub>Abl</sub>-p40, W122C Src<sub>Abl</sub>-p40, and W122C Src-PLL. Comparing Src-PLL and W122C Src-PLL states throughout our simulations confirms this result (Figure 6B). Measurements between the loops farthest from each other in our Src-PLL simulation revealed a distance of 3.3 Å. For W122C Src-PLL, this distance was 10.6 Å (7.9 Å without an outlier loop) (Figure 6B). Taken together, this suggests that the single W122C mutation may increase flexibility in the nSrc loop of the Src SH3 domain.

**TABLE 2** Confidence measures from AlphaFold–Multimer predictions (for the rank\_001 model plus a standard deviation for all 5 output models in parentheses). Italics indicate that these predictions are either not robust based on the confidence measures, or that only certain models (e.g., rank\_001) can be interpreted; therefore, these models were excluded from analyses.

	pLDDT	pTM	ipTM
Src SH3:PLLP	94.3 (0.95)	0.838 (0.014)	0.775 (0.024)
Src SH3:p40	91.8 (1.08)	0.801 (0.009)	0.749 (0.018)
Src <sub>Abl</sub> :PLLP	92.7 (0.92)	0.818 (0.014)	0.728 (0.028)
Src <sub>Abl</sub> :p40	91.8 (1.08)	0.801 (0.009)	0.749 (0.018)
W122C Src SH3:PLLP	90.6 (1.59)	0.782 (0.020)	0.715 (0.049)
W122C Src SH3:p40	89.1 (1.59)	0.754 (0.012)	0.697 (0.034)
W122C Src <sub>Abl</sub> SH3:p40	88.3 (1.14)	0.752 (0.008)	0.706 (0.032)
Abl SH3:PLLP	91.8 (1.72)	0.806 (0.017)	0.669 (0.049)
Abl SH3:p40	94.9 (1.15)	0.841 (0.013)	0.847 (0.018)
Abl <sub>Src</sub> SH3:PLLP	89.4 (1.81)	0.779 (0.016)	<i>0.523 (0.108)</i>
Abl <sub>Src</sub> SH3:p40	95.5 (2.26)	0.839 (0.023)	<i>0.839 (0.126)</i>
C100W Abl SH3:p40	96.0 (0.93)	0.853 (0.011)	0.859 (0.016)

## Investigating a loop orientation-determining residue in the WX sequence motif

We next wanted to use existing SH3 structures to assess if this position may broadly affect loop orientation. We analyzed 39 previously solved SH3 structures, most of which had peptides bound in either the Class I (4 structures) or Class II (35 structures) orientation (Kaneko et al., 2008b). We aligned these structures and colored the nSrc loop residues based on the WX motif immediately following the nSrc loop, including blue(s) for X = aromatic and yellow/orange for X = Cys or Leu (Figure 7A). Here, the X residues are the W122 Src SH3 and C100 Abl SH3 positions previously discussed. Although there are multiple structures for several of the SH3 domains, we have representative structures included for SH3 domains from the following proteins: WC (Abl), WL (CMS N-terminal SH3, or CMS-N, p40<sup>phox</sup>, p67<sup>phox</sup>), WY (CSK, Grb2-N, IB1), and WW ( $\beta$ -PIX, CIN85-N, Cortactin, Fyn, GADS C-terminal SH3 or GADS-C, Hck, Itk, p47<sup>phox</sup>, PLC- $\gamma$ 1, SLA1, Src, STAM2) [(Pisabarro and Serrano, 1996; Pisabarro et al., 1998; Pisabarro and Serrano, 1996; Pisabarro et al., 1998; Nagar et al., 2003; Pisabarro et al., 1998; Donaldson et al., 2002; Musacchio et al., 1994; Nagar et al., 2006; Moncalián et al., 2006; Jozic et al., 2005; Kami et al., 2002; Massenet et al., 2005; Andreotti et al., 1997; Deng et al., 2005; Liu et al., 2003; Lewitzky et al., 2004; Kaneko et al., 2003; Mott et al., 2005; Janz et al., 2007; Wu et al., 1995; Hoelz et al., 2006; Arold et al., 1997; Lee et al., 1996; Wittekind et al., 1997; Harkiolaki et al., 2003; Dimasi, 2007; He et al., 2007; Hashimoto et al., 2006; Ogura et al., 2006; Xu et al., 1997;

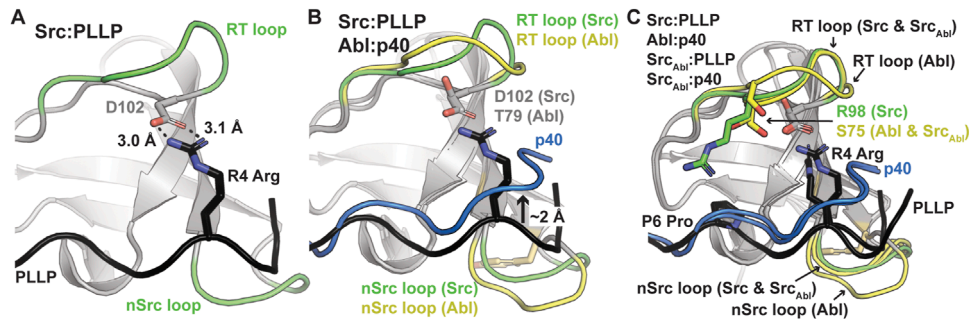
Cowan-Jacob et al., 2005; Witucki et al., 2002; Ghose et al., 2001; Vidal et al., 1999; Kristensen et al., 2006; Xu et al., 1999)].

Our analyses revealed that SH3 domains with WW or WY motifs following the nSrc loop contain more “Src-like” conformations whereas those with WC or WL motifs appear more “Abl-like.” Of the structures analyzed, the WY motif-containing SH3 domains show the most variability (Figure 7A). Notably, these results are based on static structures available in the Protein Data Bank (PDB). It would be interesting to run MD simulations or conduct NMR experiments on a variety of WX motif-containing SH3 domains to test the influence of the WX loop on nSrc loop flexibility, as observed in our MD simulations with Src SH3 versus W122C Src SH3 (Figure 6B). Overall, our results suggest that this position may play an important role in the conformation (and potentially the relative flexibility) of the nSrc loop. In combination with RT loop residues, this directly determines how ligands could interact with the peptide-binding cleft and previously identified specificity pockets for SH3 domains.

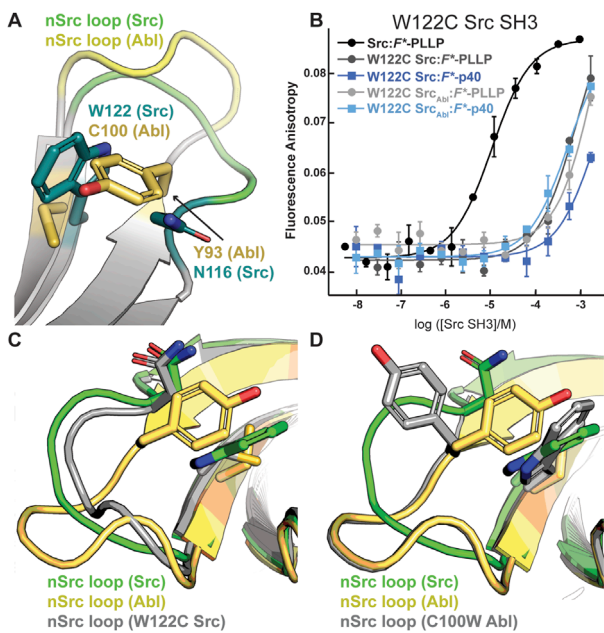
Finally, we analyzed the sequences of all 344 annotated SH3 domains in the human proteome, using the UniProt database [(Magrane and Consortium, 2011; UniProt Consortium, 2021; UniProt Consortium, 2019)]. We extracted the SH3 domain sequences, performed a multiple sequence alignment using Clustal Omega, and visualized the results using Jalview (Supplementary Figure S2) [(Sievers et al., 2011; Clamp et al., 2004)]. There were several atypical SH3 domains based on secondary structural elements according to manual visualization of available AlphaFold structures [(Varadi et al., 2022; Jumper et al., 2021)]; therefore, we chose to exclude any sequences that did not align properly from further analyses. In total, we analyzed the WX sequence motifs for 248 human SH3 domains (Figures 7B, C), and found that indeed, a WW sequence is the most observed, with 138 occurrences (55.6% of total sequences). Position by position, we see a Trp in the first position of the sequence motif in 91% of SH3 domains analyzed, and 59% in the second position (Figure 7C). While there were four alternative residues in the first position (including Leu, Phe, and Tyr), there are an additional thirteen alternatives in the second position (Ala, Cys, Phe, Gly, His, Ile, Leu, Met, Gln, Ser, Thr, Val, Tyr) (Figure 7C). Overall, this result confirms variability in the X, or second, position of the WX sequence motif defined here, within the human SH3 domains.

## Discussion

Despite their relatively small size, SH3 domains display remarkable plasticity in ligand binding and the family has been described as *versatile* and *diverse* [(Mehrabipour et al., 2023; Saksela and Permi, 2012; Kaneko et al., 2008b; Kaneko et al., 2011)]. This includes the ability to bind ligands in opposite orientations, depending on the sequences of the short linear motifs (SLiMs) recognized. Specificity-determining components of SH3 have been described, and the RT and nSrc loops are known to play an important role [(Saksela and Permi, 2012; Kaneko et al., 2008b; Kaneko et al., 2011)]. A conserved Trp residue immediately following the nSrc loop, and at the start of the  $\beta$ C strand was also identified as forming a hydrophobic pocket that is critical for ligand binding (Kaneko et al., 2011). Here, we extend discussion of this Trp residue to include its



**FIGURE 4**  
Ligand interactions with the RT and nSrc loops of Src and Abl SH3 domains. In all, SH3 domains and ligands are shown in cartoon representation, with loops colored by original sequence (green for Src and yellow for Abl), even in the chimeric proteins, as labeled. The ligands are shown as black for PLLP and blue for p40. Relevant side chains are shown as sticks and colored by atom (C = as above, N = blue, O = red). **(A)** The P4 Arg in the ligand interacts with D102 in Src SH3. Distances shown as black dashed lines and labeled. **(B)** The p40 ligand bound to Abl SH3 is translated ~2 Å in the ligand binding pocket. **(C)** Although the RT loop of Src<sub>Abl</sub> SH3 is in a similar conformation as wild-type Src SH3, R98 (Src) is replaced by S75 (Abl), allowing the p40 peptide to bind Src<sub>Abl</sub> in the elevated peptide-binding conformation. The nSrc loop of Src<sub>Abl</sub> is also in a Src-like conformation.



**FIGURE 5**  
The WX motif following the nSrc loop determines its conformation. All structures, generated by AlphaFold-Multimer, are shown in cartoon representation and colored as labeled. Side chain sticks are shown for relevant positions, and colored by atom (C = yellow (Abl), green (Src), or gray (W122C Src or C100W Abl), O = red, N = blue). **(A)** The X position of the WX motif in Src (X = W) and Abl (X = C) makes intra-SH3 interactions with N116 (Src) or Y93 (Abl). **(B)** Binding affinity data by fluorescence anisotropy experiments in triplicate confirms an importance for the W122 residue in the Src SH3 domain (Table 1). All averaged curves were normalized to an identical starting value to better illustrate  $K_d$  shifts. **(C, D)** The nSrc loop in the W122C Src model is shifted with respect to wild-type Src SH3 **(C)**; whereas the nSrc loop in the C100W Abl SH3 model maintains an “Abl-like” conformation **(D)**.

C-terminal neighbor, and define the WX motif, where the identity of the X residue is also important for modulating the conformation (and potentially flexibility) of the nSrc loop.

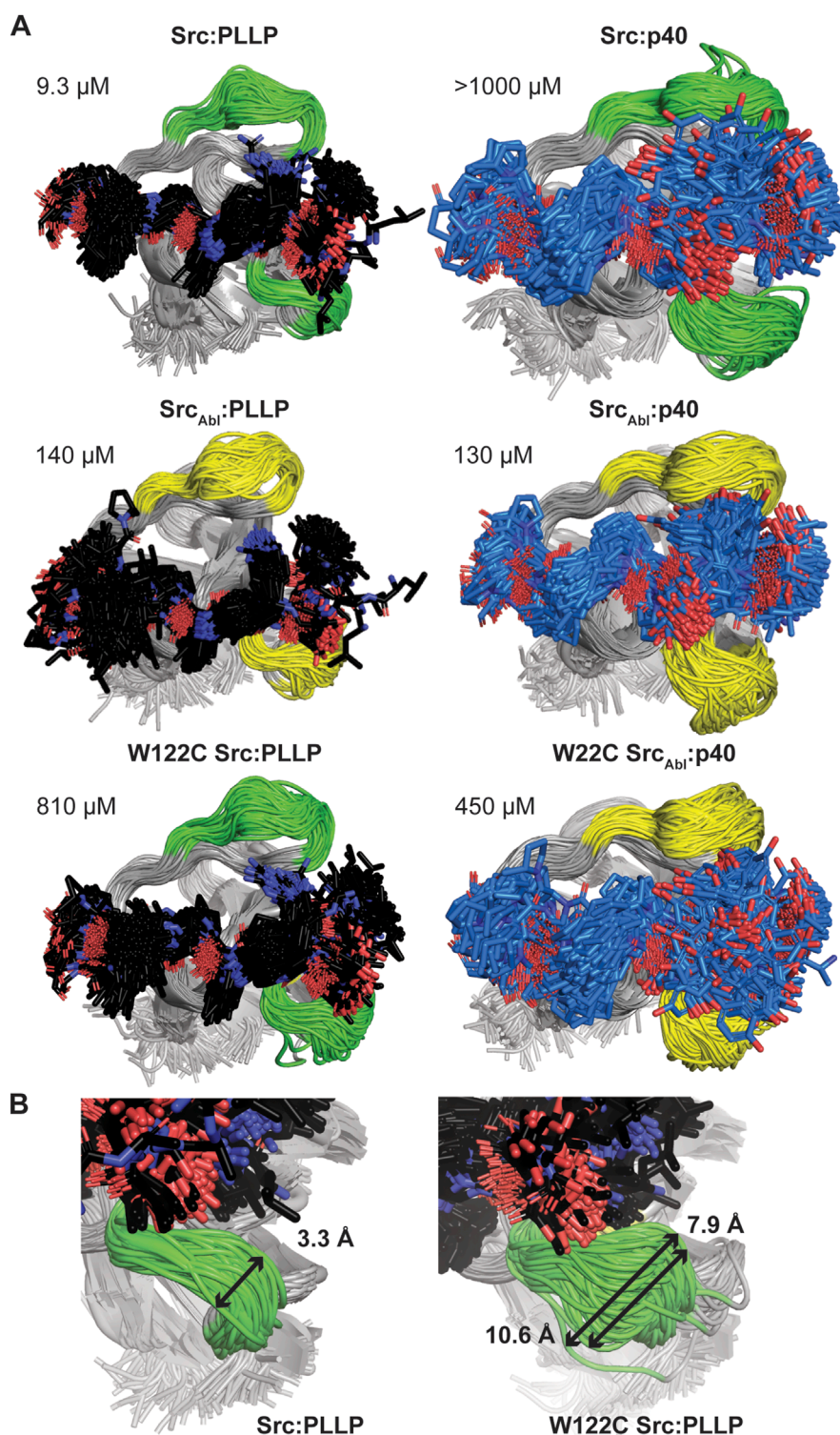
Our biochemical and computational analyses suggest that the WX motif may determine the conformation of the nSrc loop of Src SH3, even in the presence of peptides that bind in distinct conformations (e.g., PLLP versus p40). When we broadened this analysis to include 39 SH3 domain structures, we also saw clear patterns in the positions of the nSrc loops (Figure 7), supporting our biochemical and modeling data. Multiple sequence analyses of the WX sequence motifs of 253 human SH3 domains confirmed that the X position is variable, with a total of fourteen amino acids in these sequences.

While this study was relatively limited in scope, our work identifies an important residue for loop positioning (and potential flexibility), and therefore SH3 specificity. It would be interesting to further test these observations using additional SH3 domains and a variety of ligand sequences. Future work could also include broad mutational characterization and analyses of the X position in the WX sequence motif identified, as well as high throughput studies of the specificity profiles of these differing SH3 domains. Critically, experimental analyses of the effect of the WX motif on relative flexibility of the nSrc loop is needed to confirm our computational data. Considering their widespread and critical roles in the cell, a better understanding of the molecular determinants of SH3 specificity can advance knowledge of complex signaling pathways towards the improvement of human health.

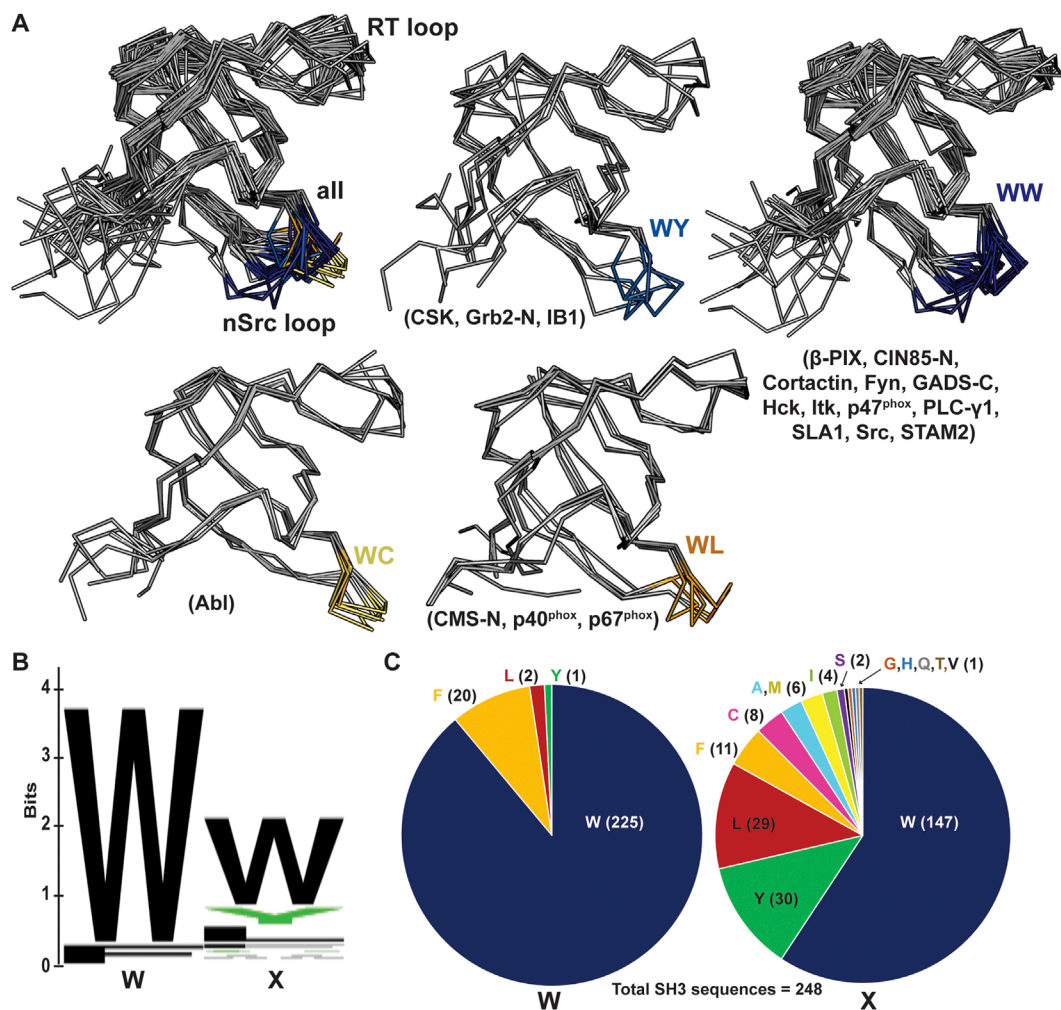
## Materials and methods

### Protein expression and purification

Sequences corresponding to the Abl and Src SH3 (residues 84-145, UniProt ID SRC\_HUMAN) domains were used for the wild-type SH3 domains. All proteins were expressed as His<sub>6</sub>-SUMO fusion proteins using the pET28a (+) plasmid (Genscript). Expression and purification protocols were similar to those used previously for other SLiM-binding domains [(Tahti et al., 2023; Gao et al., 2020; Wofford et al., 2021)]. All sequences used are in the Supporting Information.



**FIGURE 6**  
Molecular dynamics simulations of Src SH3-peptide complexes. All MD simulations were run for 1  $\mu\text{s}$ . **(A)** For each simulation, 200 states ( $\Delta t = 5 \text{ ns}$ ) are shown aligned by SH3 domain, which are shown in cartoon representation and labeled. The RT and nSrc loops are colored by sequence identity (green = amino acids from Src, yellow = amino acids from Abl). The bound peptides are shown in stick representation and colored by atom (C = black for PLLP, marine for p40, O = red, N = blue). Calculated binding affinities (Table 1) are displayed. **(B)** The relative positions sampled by the nSrc loop in the Src:PLLP (left) and W122C Src:PLLP (right) simulations are shown, with distances measured (indicated by a black arrow and as labeled). Structures are rendered as in **(A)**.



**FIGURE 7** nSrc loop conformations and sequences of multiple SH3 domain structures. **(A)** All structures are shown in ribbon representation, with the nSrc loop residues colored based on the “WX” motif immediately following the loop (WW = dark blue, WY = blue, WC = yellow, WL = orange). The SH3 domains that are represented are indicated in parentheses; the majority of these structures include ligand, most in the Class II orientation. PDB ID codes used include: 1BBZ, 1JU5, 1AB0, 1OPL, 2FO0, 1OPK, 2J6F, 2J6O, 1K4U, 1W70, 1AWJ, 1YWO, 1RLP, 1H3H, 1UTI, 1UJ0, 1ZSG, 2P4R, 1CKA, 2DF6, 1AVZ, 1EFN, 1OEB, 2D0N, 2JT4, 2D1X, 2AK5, 2BZ8, 1WLP, 2HCK, 1AD5, 1FMK, 1Y57, 2SRC, 1KSW, 1JEG, 1AZE, 4GBQ, 2FPD (10, 16, 18, 19, 24–54). **(B, C)** The same data shown as a **(B)** WebLogo of the WX motifs for 253 SH3 domain sequences from UniProt, and **(C)** pie charts of the sequence composition per position. Of the 248 sequences analyzed, 91% contain a Trp in the first position and 59% in the second position.

Briefly, plasmids were transformed into BL21 DE3 chemically competent *Escherichia coli* cells. Selected colonies were grown in Terrific Broth (TB) at 37°C with shaking at 210 rpm. Once an optical density (OD) of 0.6–0.8 at  $\lambda = 600$  nm was reached, protein expression was induced by the addition of 0.15 mM isopropyl- $\beta$ -D-1-thiogalactopyranoside (IPTG) for 16–18 h at 18°C. Cells were harvested via centrifugation at 3,000xg for 10 min, followed by resuspension in lysis buffer (0.05 M Tris pH 7.5, 0.2 M NaCl, 0.01 M MgCl<sub>2</sub>, 0.01 M CaCl<sub>2</sub>, 0.05 M imidazole pH 7.5, 20% (w/v) glycerol, 0.25 mM TCEP; 20  $\mu$ g/mL DNase and Complete EDTA-free protease inhibitor tablets (1 tablet/50 mL lysis buffer) were also added). Cells were lysed using sonication at 4°C and whole cell lysate was clarified with centrifugation at 17,500 rpm at 4°C for 30 min. The filtered supernatant was applied to a NiNTA HisTrap

5 mL column (Cytiva) with wash buffer [0.025 M Tris pH 7.5, 0.25 M NaCl, 10% (w/v) glycerol, 0.025 M imidazole pH 7.5, 0.25 mM tris(2-carboxyethyl)phosphine (TCEP)]. Protein was eluted into fractions using elution buffer [0.025 M Tris-HCl pH 7.5, 0.05 M NaCl, 10% (w/v) glycerol, 0.4 M imidazole pH 7.5, 0.25 mM TCEP]. Pooled fractions were dialyzed in the presence of ULP-1 SUMO protease in dialysis buffer [0.025 M Tris pH 7.5, 0.15 M NaCl, 10% (w/v) glycerol, 0.5 mM TCEP] overnight, and then run over another NiNTA His-Trap 5 mL column to separate the SH3 domain from the His<sub>6</sub>-SUMO tag, using running buffer plus 0.025 M imidazole pH 7.5. The flow-through was concentrated using Amicon Ultra 3K Centrifugal Filters and further purified using a Superdex S75 16/600 column, with running buffer. SDS-PAGE was used to assess protein purity and the absorbance at  $\lambda = 280$  nm with the calculated

extinction coefficient (s), by ExPasy ProtParam, was used to calculate protein concentration (Wilkins et al., 1999).

Calculated extinction coefficients used were as follows: SUMO-Src-SH3 = 18,450 M<sup>-1</sup> cm<sup>-1</sup> (cleaved, 16,960 M<sup>-1</sup> cm<sup>-1</sup>); SUMO-Abl-SH3 = 16,960 M<sup>-1</sup> cm<sup>-1</sup> (cleaved, 15,470 M<sup>-1</sup> cm<sup>-1</sup>); SUMO-Src<sub>Abl</sub>-SH3 = 18,450 M<sup>-1</sup> cm<sup>-1</sup> (cleaved, 16,960 M<sup>-1</sup> cm<sup>-1</sup>); SUMO-Abl<sub>Src</sub>-SH3 = 16,960 M<sup>-1</sup> cm<sup>-1</sup> (cleaved, 15,470 M<sup>-1</sup> cm<sup>-1</sup>); W122C SUMO-Src-SH3 = 12,950 M<sup>-1</sup> cm<sup>-1</sup> (cleaved, 16,460 M<sup>-1</sup> cm<sup>-1</sup>); C100W SUMO-Abl-SH3 = 22,460 M<sup>-1</sup> cm<sup>-1</sup> (cleaved, 20,970 M<sup>-1</sup> cm<sup>-1</sup>); W122C SUMO-Src<sub>Abl</sub>-SH3 = 12,950 M<sup>-1</sup> cm<sup>-1</sup> (cleaved, 11,460 M<sup>-1</sup> cm<sup>-1</sup>).

## Fluorescence anisotropy assays

The peptides used for fluorescence anisotropy assays were F\*-PLLP (sequence: FITC-Ahx-LASRPLPLLP), the Src SH3-specific binder, and F\*-p40 (FITC-Ahx-APTYSPPPPP), the Abl SH3-specific binder (Biomatik) (6, 20). Fluorescence anisotropy assays were performed using a BioTek Synergy H1 plate reader, as previously described [(Tahti et al., 2023; Gao et al., 2020; Wofford et al., 2021; Valgardson et al., 2019)]. Solution conditions included running buffer plus 0.1 mg/mL bovine serine albumin (BSA) and 0.5 mM Thesit, with 30 nM fluorescent peptide. Determined K<sub>D</sub> values were the average calculated from triplicate experiments (Cushing et al., 2008). K<sub>d</sub> values were determined as previously, and as the midpoint between the anisotropy resulting from the free reporter peptide and SH3-peptide complex [(Cushing et al., 2008; Vouilleme et al., 2010; Amacher et al., 2013; Amacher et al., 2014)].

## Molecular dynamics simulations

A chosen AlphaFold3 output model (usually, the ranked\_001 model, but an additional model if steric clash was observed) was capped with an N-terminal acetyl and C-terminal N-methyl group, respectively. MD simulations were performed using GROMACS 2022.4 (Bauer et al., 2022). Each system was centered within a simulation box, maintaining a solute-box distance of 1.0 nm, and solvated using the TIP3P water model. The system was then given the solvent configuration spc216, representing a pre-equilibrated configuration of 216 water molecules based on the Simple Point Charge (SPC) (Berendsen et al., 1981). To mimic physiological conditions, ions were added to achieve a final concentration of 0.15 M, with charge neutrality maintained by a balanced addition of sodium and chloride ions. Simulations employed the AMBER99SB-ILDN force field (Wang et al., 2004). Energy minimization was performed using the steepest descent algorithm, with convergence defined by a maximum force of <1,000 kJ/mol/nm. Long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) algorithm (Essmann et al., 1995), while short range electrostatic and Van der Waals cutoffs were set at 1.0 nm. A preliminary equilibration was conducted over 100 ps using an NVT file with parameters set for output control, electrostatics, bond parameters, neighbor searching, and temperature coupling.

Pressure coupling was disabled. Temperature coupling employed a modified Berendsen thermostat, maintaining a reference temperature of 300 K across two coupling groups. The Verlet cutoff scheme was used for neighbor searching with a grid-cell based method. An additional equilibration phase was performed using the isotropic Parrinello-Rahman barostat, targeting a reference temperature of 1 bar (Parrinello and Rahman, 1981). Production runs were conducted for 5 ns, with a timestep (dt) of 2fs. Each production run was then extended by an additional 995 ns. Final simulations were then analyzed and visualized in PyMOL (Schrödinger software).

## Programs used for structural modeling and analyses

AlphaFold-Multimer (in the CoLabFold notebook) was used to model ligand-bound SH3 domains, using the default settings [(Evans et al., 2021; Mirdita et al., 2022)]. Confidence measures are included in Table 2, for reference, pLDDT is a confidence score (>90 = modeled with high accuracy), pTM is an integrated measure (>0.5 = overall predicted fold for complex is similar to true structure), and ipTM correlates to the accuracy of predicted relative positions between multimers (>0.8 = confident high-quality predictions, 0.6-0.8 = “grey zone,” <0.6 likely represent a failed prediction). Binding curves were visualized using KaleidaGraph (Synergy software). Structural figures were rendered using PyMOL. We used UniProt to download the human SH3ome, following by Clustal Omega to generate a multiple sequence alignment (default settings), Jalview to view the results, and WebLogo and Excel to analyze and visualize the data (Supplementary Figure S2) [(Magrane and Consortium, 2011; UniProt Consortium, 2021; UniProt Consortium, 2019; Sievers et al., 2011; Clamp et al., 2004; Crooks et al., 2004)]. In our initial curation of the human SH3 multiple sequence alignment, we used the AlphaFold protein database to manually visualize WX sequence motifs that did not properly align [(Varadi et al., 2022; Jumper et al., 2021)].

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Author contributions

FL-N: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Validation, Visualization, Writing—original draft, Writing—review and editing. CC: Formal Analysis, Funding acquisition, Investigation, Writing—review and editing. CM: Investigation, Writing—review and editing. ET: Investigation, Writing—review and editing. JB: Investigation, Writing—review and editing. SS: Investigation, Writing—review and editing. JA: Conceptualization, Data curation, Formal

Analysis, Funding acquisition, Project administration, Resources, Supervision, Visualization, Writing—original draft, Writing—review and editing.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by NSF CAREER CHE-2044958 and Cottrell Scholar (Research Corporation for Science Advancement) awards to JA, as well as Seagen-funded summer research fellowships at Western Washington University to FL-N and CC, and a Cottrell Postbac fellowship to JB.

## Acknowledgments

The authors would like to thank all members of the Amacher lab for thoughtful discussions and general research support. We would also like to thank Dr. Lionel (Lee) Brooks, who was instrumental in providing computational support towards downloading the human SH3 domain sequences.

## References

- Amacher, J. F., Cushing, P. R., Bahl, C. D., Beck, T., and Madden, D. R. (2013). Stereochemical determinants of C-terminal specificity in PDZ peptide-binding domains: a novel contribution of the carboxylate-binding loop. *J. Biol. Chem.* 288, 5114–5126. doi:10.1074/jbc.M112.401588
- Amacher, J. F., Cushing, P. R., Brooks, L., Boisguerin, P., and Madden, D. R. (2014). Stereochemical preferences modulate affinity and selectivity among five PDZ domains that bind CFTR: comparative structural and sequence analyses. *Structure* 22, 82–93. doi:10.1016/j.str.2013.09.019
- Andreotti, A. H., Bunnell, S. C., Feng, S., Berg, L. J., and Schreiber, S. L. (1997). Regulatory intramolecular association in a tyrosine kinase of the Tec family. *Nature* 385, 93–97. doi:10.1038/385093a0
- Arold, S., Franken, P., Strub, M. P., Hoh, F., Benichou, S., Benarous, R., et al. (1997). The crystal structure of HIV-1 Nef protein bound to the Fyn kinase SH3 domain suggests a role for this complex in altered T cell receptor signaling. *Structure* 5, 1361–1372. doi:10.1016/S0969-2126(97)00286-4
- Bauer, P., Hess, B., and Lindahl, E. (2022). *GROMACS 2022.4 Source code*. Zenodo: KTH - Royal Institute of Technology. doi:10.5281/zenodo.7323393
- Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., and Hermans, J. (1981). “Interaction models for water in relation to protein hydration,” in *Intermolecular forces of The Jerusalem symposia on quantum chemistry and biochemistry*. Editor B. Pullman (Dordrecht: Springer Netherlands), 14, 331–342. doi:10.1007/978-94-015-7658-1\_21
- Clamp, M., Cuff, J., Searle, S. M., and Barton, G. J. (2004). The Jalview Java alignment editor. *Bioinformatics* 20, 426–427. doi:10.1093/bioinformatics/btg430
- Colicelli, J. (2010). ABL tyrosine kinases: evolution of function, regulation, and specificity. *Sci. Signal.* 3, re6. doi:10.1126/scisignal.3139re6
- Cowan-Jacob, S. W., Fendrich, G., Manley, P. W., Jahnke, W., Fabbro, D., Liebetanz, J., et al. (2005). The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure* 13, 861–871. doi:10.1016/j.str.2005.03.012
- Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi:10.1101/gr.849004
- Cushing, P. R., Fellows, A., Villone, D., Boisguerin, P., and Madden, D. R. (2008). The relative binding affinities of PDZ partners for CFTR: a biochemical basis for efficient endocytic recycling. *Biochemistry* 47, 10084–10098. doi:10.1021/bi8003928
- Deng, L., Velikovskiy, C. A., Swaminathan, C. P., Cho, S., and Mariuzza, R. A. (2005). Structural basis for recognition of the T cell adaptor protein SLP-76 by the SH3 domain of phospholipase Cgamma1. *J. Mol. Biol.* 352, 1–10. doi:10.1016/j.jmb.2005.06.072
- Dimasi, N. (2007). Crystal structure of the C-terminal SH3 domain of the adaptor protein GADS in complex with SLP-76 motif peptide reveals a unique SH3-SH3 interaction. *Int. J. Biochem. Cell Biol.* 39, 109–123. doi:10.1016/j.biocel.2006.07.003
- Donaldson, L. W., Gish, G., Pawson, T., Kay, L. E., and Forman-Kay, J. D. (2002). Structure of a regulatory complex involving the Abl SH3 domain, the Crk SH2 domain, and a Crk-derived phosphopeptide. *Proc. Natl. Acad. Sci. USA* 99, 14053–14058. doi:10.1073/pnas.212518799
- Eck, M. J., Atwell, S. K., Shoelson, S. E., and Harrison, S. C. (1994). Structure of the regulatory domains of the Src-family tyrosine kinase Lck. *Nature* 368, 764–769. doi:10.1038/368764a0
- Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *J. Chem. Phys.* 103, 8577–8593. doi:10.1063/1.470117
- Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., et al. (2021). Protein complex prediction with AlphaFold-Multimer. *BioRxiv*. doi:10.1101/2021.10.04.463034
- Feng, S., Chen, J. K., Yu, H., Simon, J. A., and Schreiber, S. L. (1994). Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science* 266, 1241–1247. doi:10.1126/science.7526465
- Gao, M., Mackley, I. G. P., Mesbahi-Vasey, S., Bamonte, H. A., Struyvenberg, S. A., Landolt, L., et al. (2020). Structural characterization and computational analysis of PDZ domains in *Monosiga brevicollis*. *Protein Sci.* 29, 2226–2244. doi:10.1002/pro.3947
- Ghose, R., Shekhtman, A., Goger, M. J., Ji, H., and Cowburn, D. (2001). A novel, specific interaction involving the Csk SH3 domain and its natural ligand. *Nat. Struct. Biol.* 8, 998–1004. doi:10.1038/nsb1101-998
- Harkioliaki, M., Lewitzky, M., Gilbert, R. J. C., Jones, E. Y., Bourette, R. P., Mouchiroud, G., et al. (2003). Structural basis for SH3 domain-mediated high-affinity binding between *Mona/Gads* and SLP-76. *EMBO J.* 22, 2571–2582. doi:10.1093/emboj/cdg258
- Hashimoto, S., Hirose, M., Hashimoto, A., Morishige, M., Yamada, A., Hosaka, H., et al. (2006). Targeting AMAP1 and cortactin binding bearing an atypical src homology 3/proline interface for prevention of breast cancer invasion and metastasis. *Proc. Natl. Acad. Sci. USA* 103, 7036–7041. doi:10.1073/pnas.0509166103
- He, Y., Hicke, L., and Radhakrishnan, I. (2007). Structural basis for ubiquitin recognition by SH3 domains. *J. Mol. Biol.* 373, 190–196. doi:10.1016/j.jmb.2007.07.074
- Hoelz, A., Janz, J. M., Lawrie, S. D., Corwin, B., Lee, A., and Sakmar, T. P. (2006). Crystal structure of the SH3 domain of betaPIX in complex with a high affinity peptide from PAK2. *J. Mol. Biol.* 358, 509–522. doi:10.1016/j.jmb.2006.02.027

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2024.1487276/full#supplementary-material>

- Janz, J. M., Sakmar, T. P., and Min, K. C. (2007). A novel interaction between atrophin-interacting protein 4 and beta-p21-activated kinase-interactive exchange factor is mediated by an SH3 domain. *J. Biol. Chem.* 282, 28893–28903. doi:10.1074/jbc.M702678200
- Jozic, D., Cárdenes, N., Deribe, Y. L., Moncalián, G., Hoeller, D., Groemping, Y., et al. (2005). Cbl promotes clustering of endocytic adaptor proteins. *Nat. Struct. Mol. Biol.* 12, 972–979. doi:10.1038/nsmb1000
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. doi:10.1038/s41586-021-03819-2
- Kami, K., Takeya, R., Sumimoto, H., and Kohda, D. (2002). Diverse recognition of non-PxxP peptide ligands by the SH3 domains from p67(phox), Grb2 and Pex13p. *EMBO J.* 21, 4268–4276. doi:10.1093/emboj/cdf428
- Kaneko, T., Kumasaka, T., Ganbe, T., Sato, T., Miyazawa, K., Kitamura, N., et al. (2003). Structural insight into modest binding of a non-PXXP ligand to the signal transducing adaptor molecule-2 Src homology 3 domain. *J. Biol. Chem.* 278, 48162–48168. doi:10.1074/jbc.M306677200
- Kaneko, T., Li, L., and Li, S. S. C. (2008a). The SH3 domain—a family of versatile peptide- and protein-recognition module. *Front. Biosci.* Volume, 4938–4952. doi:10.2741/3053
- Kaneko, T., Li, L., and Li, S. S. C. (2008b). The SH3 domain—a family of versatile peptide- and protein-recognition module. *Front. Biosci.* 13, 4938–4952. doi:10.2741/3053
- Kaneko, T., Sidhu, S. S., and Li, S. S. C. (2011). Evolving specificity from variability for protein interaction domains. *Trends biochem. Sci.* 36, 183–190. doi:10.1016/j.tibs.2010.12.001
- Kristensen, O., Guenat, S., Dar, I., Allaman-Pillet, N., Abderrahmani, A., Ferdaoussi, M., et al. (2006). A unique set of SH3-SH3 interactions controls I $\beta$ 1 homodimerization. *EMBO J.* 25, 785–797. doi:10.1038/sj.emboj.7600982
- Kurochkina, N., and Guha, U. (2013). SH3 domains: modules of protein-protein interactions. *Biophys. Rev.* 5, 29–39. doi:10.1007/s12551-012-0081-z
- Lee, C. H., Saksela, K., Mirza, U. A., Chait, B. T., and Kuriyan, J. (1996). Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. *Cell* 85, 931–942. doi:10.1016/s0092-8674(00)81276-3
- Le Roux, A.-L., Mohammad, I. L., Mateos, B., Arbesú, M., Gairi, M., Khan, F. A., et al. (2019). A myristoyl-binding site in the SH3 domain modulates c-src membrane anchoring. *iScience* 12, 194–203. doi:10.1016/j.isci.2019.01.010
- Lewitzky, M., Harkiolaki, M., Domart, M.-C., Jones, E. Y., and Feller, S. M. (2004). Mona/Gads SH3C binding to hematopoietic progenitor kinase 1 (HPK1) combines an atypical SH3 binding motif, R/KXXX, with a classical PXXP motif embedded in a polyproline type II (PPII) helix. *J. Biol. Chem.* 279, 28724–28732. doi:10.1074/jbc.M402745200
- Lim, W. A., Richards, F. M., and Fox, R. O. (1994). Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* 372, 375–379. doi:10.1038/372375a0
- Liu, Q., Berry, D., Nash, P., Pawson, T., McGlade, C. J., and Li, S. S. C. (2003). Structural basis for specific binding of the Gads SH3 domain to an RxxK motif-containing SLP-76 peptide: a novel mode of peptide recognition. *Mol. Cell.* 11, 471–481. doi:10.1016/s1097-2765(03)00046-7
- Magrane, M., and Consortium, U. (2011). UniProt Knowledgebase: a hub of integrated protein data. *Database (Oxford)* 2011, bar009. doi:10.1093/database/bar009
- Massenet, C., Chenavas, S., Cohen-Addad, C., Dagher, M. C., Brandolin, G., Pebay-Peyroula, E., et al. (2005). Effects of p47phox C terminus phosphorylations on binding interactions with p40phox and p67phox. Structural and functional comparison of p40phox and p67phox SH3 domains. *J. Biol. Chem.* 280, 13752–13761. doi:10.1074/jbc.M412897200
- Mayer, B. J. (2001). SH3 domains: complexity in moderation. *J. Cell Sci.* 114, 1253–1263. doi:10.1242/jcs.114.7.1253
- Mayer, B. J., Hamaguchi, M., and Hanafusa, H. (1988). A novel viral oncogene with structural similarity to phospholipase C. *Nature* 332, 272–275. doi:10.1038/332272a0
- Mehrabipour, M., Jasemi, N. S. K., Dvorsky, R., and Ahmadian, M. R. (2023). A systematic compilation of human SH3 domains: a versatile superfamily in cellular signaling. *Cells* 12, 2054. doi:10.3390/cells12162054
- Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nat. Methods.* 19, 679–682. doi:10.1038/s41592-022-01488-1
- Moncalián, G., Cárdenes, N., Deribe, Y. L., Spínola-Amilibia, M., Dikic, I., and Bravo, J. (2006). Atypical polyproline recognition by the CMS N-terminal Src homology 3 domain. *J. Biol. Chem.* 281, 38845–38853. doi:10.1074/jbc.M606411200
- Mott, H. R., Nietlispach, D., Evetts, K. A., and Owen, D. (2005). Structural analysis of the SH3 domain of beta-PIX and its interaction with alpha-p21 activated kinase (PAK). *Biochemistry* 44, 10977–10983. doi:10.1021/bi050374a
- Musacchio, A., Saraste, M., and Wilmanns, M. (1994). High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides. *Nat. Struct. Biol.* 1, 546–551. doi:10.1038/nsb0894-546
- Nagar, B., Hantschel, O., Seeliger, M., Davies, J. M., Weis, W. I., Superti-Furga, G., et al. (2006). Organization of the SH3-SH2 unit in active and inactive forms of the c-Abl tyrosine kinase. *Mol. Cell.* 21, 787–798. doi:10.1016/j.molcel.2006.01.035
- Nagar, B., Hantschel, O., Young, M. A., Scheffzek, K., Veach, D., Bornmann, W., et al. (2003). Structural basis for the autoinhibition of c-Abl tyrosine kinase. *Cell* 112, 859–871. doi:10.1016/s0092-8674(03)00194-6
- Ogura, K., Nobuhisa, I., Yuzawa, S., Takeya, R., Torikai, S., Saikawa, K., et al. (2006). NMR solution structure of the tandem Src homology 3 domains of p47phox complexed with a p22phox-derived proline-rich peptide. *J. Biol. Chem.* 281, 3660–3668. doi:10.1074/jbc.M505193200
- Parrinello, M., and Rahman, A. (1981). Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* 52, 7182–7190. doi:10.1063/1.328693
- Pawson, T., and Nash, P. (2003). Assembly of cell regulatory systems through protein interaction domains. *Science* 300, 445–452. doi:10.1126/science.1083653
- Pisabarro, M. T., and Serrano, L. (1996). Rational design of specific high-affinity peptide ligands for the Abl-SH3 domain. *Biochemistry* 35, 10634–10640. doi:10.1021/bi960203t
- Pisabarro, M. T., Serrano, L., and Wilmanns, M. (1998). Crystal structure of the Abl-SH3 domain complexed with a designed high-affinity peptide ligand: implications for SH3-ligand interactions. *J. Mol. Biol.* 281, 513–521. doi:10.1006/jmbi.1998.1932
- Saksela, K., and Permi, P. (2012). SH3 domain ligand binding: what's the consensus and where's the specificity? *FEBS Lett.* 586, 2609–2614. doi:10.1016/j.febslet.2012.04.042
- Seeliger, M. A., Nagar, B., Frank, F., Cao, X., Henderson, M. N., and Kuriyan, J. (2007). c-Src binds to the cancer drug imatinib with an inactive Abl/c-Kit conformation and a distributed thermodynamic penalty. *Structure* 15, 299–311. doi:10.1016/j.str.2007.01.015
- Shah, N. H., Amacher, J. F., Nocka, L. M., and Kuriyan, J. (2018). The Src module: an ancient scaffold in the evolution of cytoplasmic tyrosine kinases. *Crit. Rev. Biochem. Mol. Biol.* 53, 535–563. doi:10.1080/10409238.2018.1495173
- Sicheri, F., Moarefi, I., and Kuriyan, J. (1997). Crystal structure of the Src family tyrosine kinase Hck. *Nature* 385, 602–609. doi:10.1038/385602a0
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539. doi:10.1038/msb.2011.75
- Sipeki, S., Koprivanac, K., Takács, T., Kurilla, A., László, L., Vas, V., et al. (2021). Novel roles of SH2 and SH3 domains in lipid binding. *Cells* 10, 1191. doi:10.3390/cells10051191
- Sparks, A. B., Rider, J. E., Hoffman, N. G., Fowlkes, D. M., Quillam, L. A., and Kay, B. K. (1996). Distinct ligand preferences of src homology 3 domains from src, yes, abl, cortactin, p53bp2, PLCgamma, crk, and Grb2. *Proc. Natl. Acad. Sci. USA.* 93, 1540–1544. doi:10.1073/pnas.93.4.1540
- Stahl, M. L., Ferenz, C. R., Kelleher, K. L., Kriz, R. W., and Knopf, J. L. (1988). Sequence similarity of phospholipase C with the non-catalytic region of src. *Nature* 332, 269–272. doi:10.1038/332269a0
- Tahti, E. F., Blount, J. M., Jackson, S. N., Gao, M., Gill, N. P., Smith, S. N., et al. (2023). Additive energetic contributions of multiple peptide positions determine the relative promiscuity of viral and human sequences for PDZ domain targets. *Protein Sci.* 32, e4611. doi:10.1002/pro.4611
- Teyra, J., Huang, H., Jain, S., Guan, X., Dong, A., Liu, Y., et al. (2017). Comprehensive analysis of the human SH3 domain family reveals a wide variety of non-canonical specificities. *Structure* 25, 1598–1610. doi:10.1016/j.str.2017.07.017
- UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 47, D506–D515. doi:10.1093/nar/gky1049
- UniProt Consortium (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* 49, D480–D489. doi:10.1093/nar/gkaa1100
- Valgardson, J., Cosbey, R., Houser, P., Rupp, M., Van Bronkhorst, R., Lee, M., et al. (2019). MotifAnalyzer-PDZ: a computational program to investigate the evolution of PDZ-binding target specificity. *Protein Sci.* 28, 2127–2143. doi:10.1002/pro.3741
- Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., et al. (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 50, D439–D444. doi:10.1093/nar/gkab1061
- Vidal, M., Goudreau, N., Cornille, F., Cussac, D., Gincel, E., and Garbay, C. (1999). Molecular and cellular analysis of Grb2 SH3 domain mutants: interaction with Sos and dynamin. *J. Mol. Biol.* 290, 717–730. doi:10.1006/jmbi.1999.2899
- Vouilleme, L., Cushing, P. R., Volkmer, R., Madden, D. R., and Boisguenier, P. (2010). Engineering peptide inhibitors to overcome PDZ binding promiscuity. *Angew. Chem. Int. Ed.* 49, 9912–9916. doi:10.1002/anie.201005575

- Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and testing of a general amber force field. *J. Comput. Chem.* 25, 1157–1174. doi:10.1002/jcc.20035
- Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., et al. (1999). Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* 112, 531–552. doi:10.1385/1-59259-584-7:531
- Wittekind, M., Mapelli, C., Lee, V., Goldfarb, V., Friedrichs, M. S., Meyers, C. A., et al. (1997). Solution structure of the Grb2 N-terminal SH3 domain complexed with a ten-residue peptide derived from SOS: direct refinement against NOEs, J-couplings and <sup>1</sup>H and <sup>13</sup>C chemical shifts. *J. Mol. Biol.* 267, 933–952. doi:10.1006/jmbi.1996.0886
- Witucki, L. A., Huang, X., Shah, K., Liu, Y., Kyin, S., Eck, M. J., et al. (2002). Mutant tyrosine kinases with unnatural nucleotide specificity retain the structure and phospho-acceptor specificity of the wild-type enzyme. *Chem. Biol.* 9, 25–33. doi:10.1016/s1074-5521(02)00091-1
- Wofford, H. A., Myers-Dean, J., Vogel, B. A., Alamo, K. A. E., Longshore-Neate, F. A., Jagodzinski, F., et al. (2021). Domain analysis and motif matcher (dammm): a program to predict selectivity determinants in monosiga brevicollis pdz domains using human pdz data. *Molecules* 26, 6034. doi:10.3390/molecules26196034
- Wu, X., Knudsen, B., Feller, S. M., Zheng, J., Sali, A., Cowburn, D., et al. (1995). Structural basis for the specific interaction of lysine-containing proline-rich peptides with the N-terminal SH3 domain of c-Crk. *Structure* 3, 215–226. doi:10.1016/s0969-2126(01)00151-4
- Xu, W., Doshi, A., Lei, M., Eck, M. J., and Harrison, S. C. (1999). Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol. Cell.* 3, 629–638. doi:10.1016/s1097-2765(00)80356-1
- Xu, W., Harrison, S. C., and Eck, M. J. (1997). Three-dimensional structure of the tyrosine kinase c-Src. *Nature* 385, 595–602. doi:10.1038/385595a0