

Functional Classification and Interaction Selectivity Landscape of the Human SH3 Domain Superfamily

Citation for published version (APA):

Kazemein Jasemi, N. S., Mehrabipour, M., Magdalena Estirado, E., Brunsveld, L., Dvorsky, R., & Ahmadian, M. R. (2024). Functional Classification and Interaction Selectivity Landscape of the Human SH3 Domain Superfamily. *Cells*, 13(2), Article 195. <https://doi.org/10.3390/cells13020195>

Document license:

CC BY

DOI:

[10.3390/cells13020195](https://doi.org/10.3390/cells13020195)

Document status and date:

Published: 01/01/2024

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy





If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Article

Functional Classification and Interaction Selectivity Landscape of the Human SH3 Domain Superfamily

Neda S. Kazemein Jasemi ^{1,†}, Mehrnaz Mehrabipour ^{1,†}, Eva Magdalena Estirado ², Luc Brunsveld ², Radovan Dvorsky ¹ and Mohammad R. Ahmadian ^{1,*}

¹ Institute of Biochemistry and Molecular Biology II, Medical Faculty and University Hospital Düsseldorf, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; neda.jasemi@hhu.de (N.S.K.J.); mehrnaz.mehrabipour@hhu.de (M.M.); radovan.dvorsky@gmail.com (R.D.)

² Laboratory of Chemical Biology, Department of Biomedical Engineering and Institute for Complex Molecular Systems (ICMS), Eindhoven University of Technology, P.O. Box 513, 5600MB Eindhoven, The Netherlands; evamagdalen2@gmail.com (E.M.E.); l.brunsveld@tue.nl (L.B.)

* Correspondence: reza.ahmadian@hhu.de; Tel.: +49-21-1811-2384

† These authors contributed equally to this work.

Abstract: SRC homology 3 (SH3) domains are critical interaction modules that orchestrate the assembly of protein complexes involved in diverse biological processes. They facilitate transient protein–protein interactions by selectively interacting with proline-rich motifs (PRMs). A database search revealed 298 SH3 domains in 221 human proteins. Multiple sequence alignment of human SH3 domains is useful for phylogenetic analysis and determination of their selectivity towards PRM-containing peptides (PRPs). However, a more precise functional classification of SH3 domains is achieved by constructing a phylogenetic tree only from PRM-binding residues and using existing SH3 domain–PRP structures and biochemical data to determine the specificity within each of the 10 families for particular PRPs. In addition, the C-terminal proline-rich domain of the RAS activator SOS1 covers 13 of the 14 recognized proline-rich consensus sequence motifs, encompassing differential PRP pattern selectivity among all SH3 families. To evaluate the binding capabilities and affinities, we conducted fluorescence dot blot and polarization experiments using 25 representative SH3 domains and various PRPs derived from SOS1. Our analysis has identified 45 interacting pairs, with binding affinities ranging from 0.2 to 125 micromolar, out of 300 tested and potential new SH3 domain–SOS1 interactions. Furthermore, it establishes a framework to bridge the gap between SH3 and PRP interactions and provides predictive insights into the potential interactions of SH3 domains with PRMs based on sequence specifications. This novel framework has the potential to enhance the understanding of protein networks mediated by SH3 domain–PRM interactions and be utilized as a general approach for other domain–peptide interactions.

Keywords: ARHGAP12; GRB2; NCK1; proline-rich motifs; protein–protein interaction; SH3 domain; signal transduction; SOS1; SRC homology 3; WRCH1/RHOU



Citation: Kazemein Jasemi, N.S.; Mehrabipour, M.; Magdalena Estirado, E.; Brunsveld, L.; Dvorsky, R.; Ahmadian, M.R. Functional Classification and Interaction Selectivity Landscape of the Human SH3 Domain Superfamily. *Cells* **2024**, *13*, 195. <https://doi.org/10.3390/cells13020195>

Academic Editors: Alexander E. Kalyuzhny and Laszlo Buday

Received: 12 December 2023

Revised: 9 January 2024

Accepted: 17 January 2024

Published: 20 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Protein–protein interactions are fundamental to the intricate machinery that controls virtually all biological processes. [1]. Among the diverse array of protein domains that facilitate these interactions, the SRC homology 3 (SH3) domains stand out as central modular units. These compact domains, consisting of approximately 60 amino acids with similar sequences that adopt a compact β -barrel fold made of five β -strands [2], are found predominantly in various signaling proteins and various protein families [3,4]. The selective interactions of the SH3 domain with proline-rich motifs (PRMs) are fundamental for the assembly and orchestration of multiprotein complexes [5]. It is rational that the SH3 domain-containing proteins (SH3DCPs) are involved in a wide variety of biological processes [4], subsequently leading to a substantial influence on a spectrum of diseases,

such as cancers [6], neurological disorders [7], kidney and urinary disorders [8,9], muscle and myopathy disorders [10,11], immune disorders [12,13], and genetic and developmental disorders [4,14,15].

To date, a series of seven types of PRM-binding modules have been reported, including SH3, WW (two highly conserved tryptophan amino acids), EVH1 (Ena/VASP homology domain 1), GYF (glycine-tyrosine-phenylalanine), Profilin, CAP-Gly (cytoskeleton-associated protein-glycine-rich), and UEV (ubiquitin E2 variant) [16,17]. Proline-rich target peptides possess core PRMs with unique properties that influence interaction selectivity. The diversity of PRMs results from the inclusion of one or more proline residues in various combinations within the peptide sequences. The proline side chains and carbonyl groups are exposed at regular intervals, allowing intermolecular hydrogen bonding with PRM-binding domains [16]. Interactions involving PRMs have a low entropic cost of binding due to the restricted rotational freedom of proline residues along the peptide backbone. This restricted flexibility contributes to a higher overall binding energy for complexes involving PRM-containing peptides (or PRPs). This property increases the affinity of PRM interactions and influences ligand recognition [16]. In addition, PRMs can interact with their binding partners in two distinct orientations; this is influenced by the arrangement of non-proline residues either located N- or C-terminal to the core motif, often involving positively charged counterparts (R or K) [16,18].

In the context of SH3 domain interaction with PRMs, a specific contact recognition occurs, where a positively charged PRM residue binds to conserved negatively charged residues in the variable loop region of the SH3 domain, resulting in moderate selectivity and affinity [5,16]. Comparative analysis of SH3 domain binding sites reveals remarkable variability and flexibility in the loop regions, contributing to the specificity and affinity of PRM binding [6,19]. The preference of the SH3 domain for specific PRMs is usually moderate, with affinities typically in the low micromolar range [20–22]. Furthermore, SH3 domains exhibit a broad spectrum of both conventional (PRM-based) and nonconventional selectivity, effectively recognizing a diverse array of protein interactors in a differentiated manner [4,5].

SH3DCPs play a pivotal role in biological processes by facilitating diverse protein–protein interactions that rely on their selectivity and affinity. The intricate nature of protein assembly orchestrated by SH3 domains raises significant questions regarding the underlying selectivity framework governing complex networks of SH3 domain–PRMs interactions [23]. Despite sharing a 25% sequence homology, accurate prediction of the selective PRM recognition by SH3 domains remains a formidable challenge [21,24,25]. In this study, we analyzed the phylogenetic and structure–function relationships of all 298 human SH3 domains, specifically focusing on the sequences of their PRM binding sites. We then performed classification based on their PRM-binding selectivity, organizing them into 10 distinct families. In addition, the distinctive recognition pattern of PRMs within SOS1, a well-established PRM-containing protein, caught our attention. This pattern of selectivity across all SH3 families led us to use SOS1 as a comprehensive model protein to elucidate the recognition mechanisms of established SH3 domains within the human proteome in our research. The binding capabilities and affinities of 25 representative SH3 domains toward 10 SOS1-derived PRPs and 2 reference peptides were carefully evaluated. The reference peptides, RP1, a derivative of SOS1 and part of P3, and RP2, a derivative of the RHO GTPase WRCH1, were used as controls. Our investigation using fluorescence dot blot and polarization techniques revealed a significant finding: out of 300 SH3 domain–peptide combinations, only 45 exhibited binding affinities, which ranged from 0.2 to 125 micromolar. This study pioneers the understanding of the selectivity and affinity of SH3 protein modules for specific PRMs, encompassing a wide range of proteins. This framework lays the foundation for a predictive matrix that enables the anticipation of SH3 domain–PRM-mediated protein–protein interactions within complex cell signaling networks.

2. Materials and Methods

2.1. Bioinformatics, Databases, and Structural Analysis

The sequences of the SH3DCPs were obtained from the UniProt database by combining full-text searches and sequence homology searches performed with the HMMER v3.4 software package. Isolated sequences of the SH3 domains were then extracted from the previously obtained proteins, again using the subprograms of HMMER. Alignment of the SH3 domain sequences was then performed using BioEdit 7.2.5 software, and the resulting phylogenetic tree was constructed using MEGA 10.2.6 software.

Available structures containing SH3 domains were retrieved from the Protein Data Bank (PDB) website using the BLAST program. To further analyze the SH3 domain structure and define its binding residues for interaction with PRMs, Python scripts were used to identify all residues in SH3 structures within 4.0 Å of the PRM bound to it. The information thus obtained was then projected onto the global sequence alignment of all SH3 domains, and only homologous residues potentially contacting PRMs were selected to define the alignment of the PRM-binding residues of SH3 domains. Finally, PRMs were collected from published articles available on the NCBI website. In addition, BLAST analysis of SOS1 PRMs was performed on the NCBI platform.

2.2. Constructs, Peptides, and Proteins

The constructs and peptides employed in our study are listed in Tables S1 and S2, respectively. All fluorescein-labeled PRPs were synthesized and used under the conditions described previously [22]. p3XFLAG-CMV ARHGAP12^{wt} [26] was used to generate a SH3 domain deletion (ARHGAP12^{ΔSH3}). NCK1^{wt}, NCK1^{ΔSH3-3}, and NCK1^{Set-1} (N205D, D206T, D226Q, and P227D) were ordered in pcDNA3.0-Flag vectors from BioCat GmbH, Heidelberg, Germany. HA-SOS1 in the pCGN vector was ordered from addgene (#32920). All SH3 domains of the proteins listed in Table S1 within the pGEX4-T1 vector were expressed in *Escherichia coli* strains CodonPlus, Rosetta, and BL21(DE3) and purified as GST-tagged fusion proteins. The purification process involved affinity chromatography on a glutathione Sepharose column [22]. For subsequent polarization analysis, a portion of these GST fusion proteins underwent cleavage of the GST tag using thrombin (#T6884-1KU, Sigma Aldrich, Taufkirchen, Germany) at 4 °C until achieving full digestion of the fusion protein. Following this cleavage step, the proteins were subjected to further purification and separation by employing size exclusion. All purified proteins underwent analysis via SDS-PAGE and were subsequently stored at −80 °C for further use.

2.3. Pull-Down and Fluorescence Dot Blot Analysis

Pull-down of 10 μM of FITC-labeled peptides (Table S2) with 5 μM of purified GST-SH3 domains (Table S1) was performed using 10 μL of glutathione Sepharose beads (GE Healthcare, Chalfont Saint Giles, UK) in a buffer containing 30 mM Tris-HCl at pH 7.5, 3 mM dithiothreitol, and 5 mM MgCl₂ for 1 h at 4 °C. Purified GST was used as a negative control. After three washes, bound proteins were eluted by incubation in the same buffer containing 20 mM reduced glutathione for 15 min at 4 °C, and the beads were separated by centrifugation. Bound FITC-labeled peptides were detected by dot blot analysis using 1 μL of eluent at an emission wavelength of 600 nm and an Odyssey Fc imaging system (LI-COR Biosciences, Lincoln, NE, USA). Detected signals were quantified densitometrically using LI-COR Image Studio version 5.2 imaging software.

2.4. Fluorescence Polarization

The interaction between fluorescein-labeled proline-rich peptides (0.2 μM) and increasing concentrations of SH3 domains (ranging from 0 to 200 μM) was measured in a buffer (containing 30 mM Tris/HCl (pH 7.5), 5 mM MgCl₂, and 3 mM DTT at 25 °C) in a total volume of 200 μL using a Fluoromax 4 fluorimeter in polarization mode and a quartz glass fluorescence cuvette (Hellma Ultra-Micro Cuvette 105.250-QS, Thermo Fisher Scientific, Waltham, MA USA). Excitation was performed at 470 nm and emission

was measured at 560 nm. Dissociation constants (K_d) were determined by fitting the concentration-dependent binding curve to a quadratic ligand binding equation.

2.5. Transfection and Immunoprecipitation Analysis

CHO-K1 cells were cultured in Dulbecco's modified Eagle's serum (DMEM, Gibco, Waltham, MA, USA) supplemented with 10% FBS (Gibco) and 1% penicillin/streptomycin (Genaxxon, Ulm, Germany). Cells were co-transfected with HA-tagged SOS1 full-length (FL) and FLAG-tagged NCK1^{wt}, NCK1^{ΔSH3-3}, and NCK1^{Set-1} or FLAG-tagged ARHGAP12^{wt} ARHGAP12^{ΔSH3} using a Turbofect reagent (Thermo Fisher Scientific). To perform co-immunoprecipitation assays, the CHO-K1 cells were lysed on ice for 5 min employing a buffer containing 50 mM Tris/HCl (pH 7.4), 100 mM NaCl, 2 mM MgCl₂, 10% glycerol, 20 mM β-glycerophosphate, 1 mM Na₃VO₄, 1% IGPAL (Thermo Fisher Scientific), and 1x protease inhibitor cocktail (Roche, Basel, Switzerland). Lysates were cleared by centrifugation (20,000× *g* at 4 °C for 5 min). Protein concentrations were determined using the Bradford assay. Lysates were then incubated overnight at 4 °C with either anti-Flag M2 agarose beads (Sigma Aldrich) or Protein A-Sepharose beads with anti-Flag antibody and anti-IgG as control. The beads were then washed three times with a wash buffer containing 50 mM Tris-HCl and 150 mM NaCl with 1 mM EDTA. Proteins bound to the beads were eluted with 2.5x Laemmli loading buffer and subjected to SDS-PAGE for further analysis. The primary antibodies used in Western blot analysis included anti-GST (own antibody), anti-Flag (1:1000 WB and 1:50 CO-IP, #F742; and #F3165, both from Sigma), anti-IgG (1:50; # sc-2025, Santa Cruz), anti-SOS1 (1:1000; #sc-256, Santa Cruz), and anti-Vinculin (1:1000; #V9131, Sigma). The secondary antibodies used were purchased from LI-COR (anti-mouse 700 nm: IRDye #926-32213; anti-rabbit 800 nm: IRDye #926-6807).

3. Results

3.1. Sequence–Structure–Function Classification of Human SH3 Domains

In our previous study, we performed a comprehensive survey in which we identified 298 SH3 domains within 221 SH3DCPs spanning a range of 13- to 720-kilodalton proteins [4]. This analysis included a phylogenetic assessment of human SH3DCPs based on their multidomain architecture, providing a convenient functional classification within different physiological pathways. However, this approach did not address the intrinsic PRM selectivity of the SH3 domain itself. Therefore, we set out to comparatively study the sequence–structure–function relationships of human SH3 domains with a focus on three key aspects: the amino acid sequence, three-dimensional (3D) structure, and spatial arrangement of PRM-binding sites, combining bioinformatics with experimental and structural biology.

As a first step, we focused on elucidating the critical aspects of the PRM binding properties of the SH3 domain. To achieve this, we obtained primary sequences covering a collection of 298 human SH3 domains from the UniProt database. These sequences were then aligned and used to construct a phylogenetic tree of the SH3 domain superfamily using MEGA software (version 10.2.6). The resulting phylogenetic tree, designated tree #1 (Figure S1), depicted the evolutionary relationships among human SH3 domains. Examination of this tree revealed a remarkable conservation of key regions essential for the 3D structure of SH3 domains (Figure S2). This finding underscored a robust and consistent sequence–structure relationship spanning specific parts of the SH3 domain responsible for PRP interactions in SH3DCP families [4]. However, a more complex scenario emerged when delving into the comprehensive analysis of SH3 domain–PRP interactions. Despite meticulous exploration of SH3 domain–PRP structures available in the protein database (Table S3) and a comprehensive review of published biochemical data on SH3 domain–PRP interactions (Table S4), no discernible structure–function relationship was revealed from tree #1. Strikingly, PRPs exhibit clustering patterns that are inconsistent with established SH3 domain families. Instead, they were distributed among distantly related families (Figure S1).

A different approach to characterize the SH3 domain–PRM interaction was to perform a second phylogenetic analysis focusing only on the active site regions of the SH3 domain, specifically the residues involved in PRM binding. Unlike the methodology for tree #1, which considered complete SH3 domain sequences, the second phylogenetic tree (tree #2) was constructed exclusively from the PRM-binding residues potentially involved in PRM interactions. The PRM-binding residues within the SH3 domains were inferred from the structures of the SH3 domains in complex with their specific PRPs (Table S3). It is noteworthy that tree #2 shows 10 different families of SH3 domains that can be very well assigned to their respective PRPs (Figures 1, S3 and S4). This strategic approach now allowed us to explore the intricate associations between individual SH3 domains and specific PRMs, as discussed in the following sections.

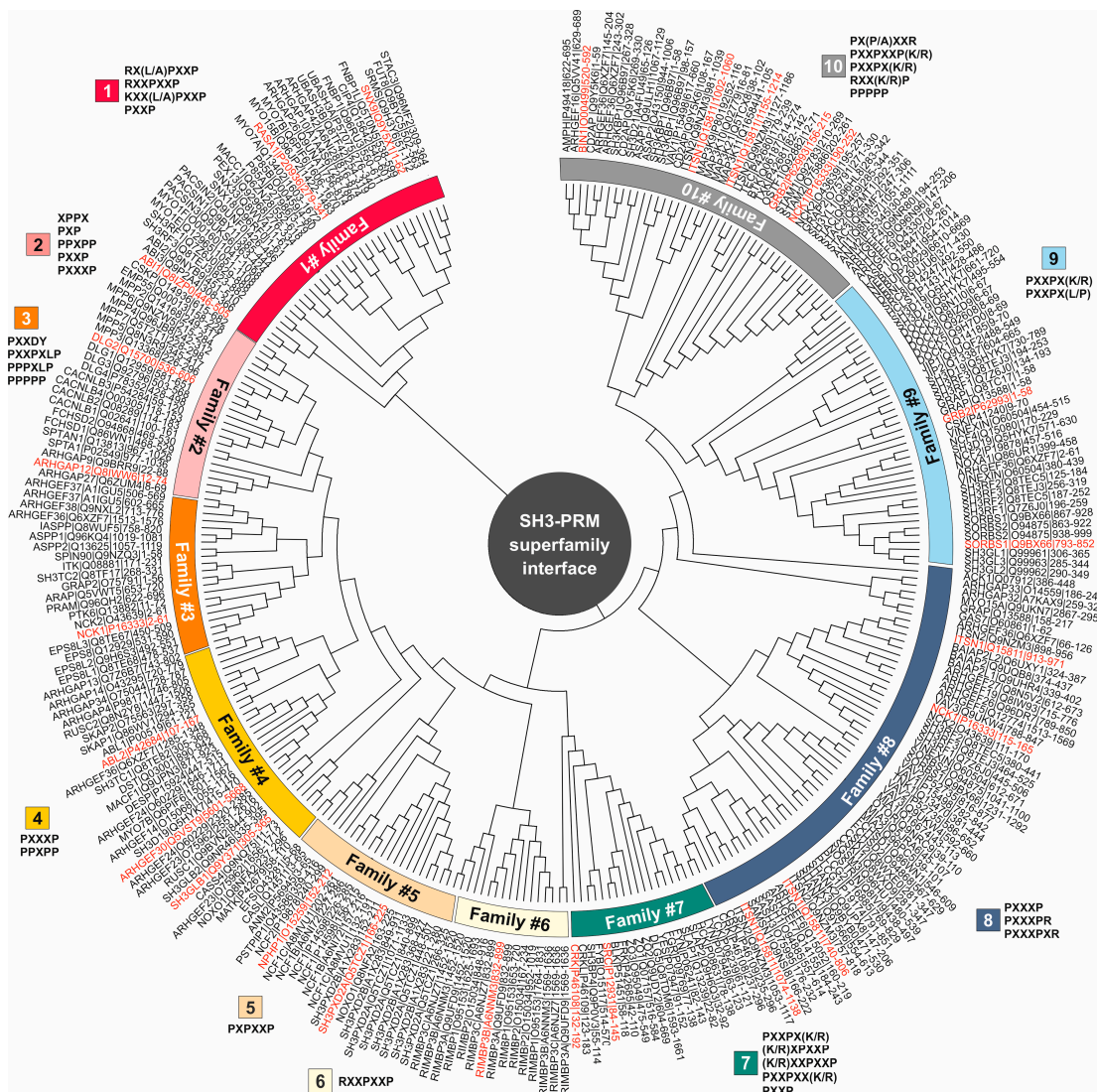


Figure 1. Evolutionary relationships among PRM-binding residues within SH3 domains. The phylogenetic tree (tree #2) was constructed by collecting PRM—binding residues from 298 human SH3 domains, using published SH3 domain-PRM structures (Table S3) and biochemical data (Table S4), and utilizing the MEGA7 software. SH3 domains were systematically classified into ten distinct families based on their interaction properties with specific PRMs as indicated by the color codes. SH3 domains highlighted in red were selected as representatives for further analysis in this study. The protein name is accompanied by the corresponding Uniprot ID and residue span for the SH3 domain.

3.2. PRM Selectivities of Different SH3 Families

Classification of PRMs into distinct families based on published structural and biochemical data (Tables S3 and S4) reveals specific sequence patterns that guide protein–protein interactions (Figure S5). Family 1 includes motifs such as RX(L/A)PXXP, RXXPXXP, KXX(L/A)PXXP, and PXXP, suggesting a diverse yet structured arrangement for interaction. Family 2, characterized by a PPXPPXP consensus, shows patterns such as XPPX, PXP, PPXPP, PXXP, and PXXXXP, indicating a diverse but consistent motif profile. Family 3 sequences, including PXXDY, PXXPXLP, PPPXLP, and PPPPP, show specificity around proline and other residues such as D and Y. Family 4 follows a PXXXXPPXP consensus with specific motifs such as PXXXXP and PPXPP. Family 5 has a specific PXPXXP motif. Notably, family 6 lacks structural and biochemical data. However, RIMBP1/2 (RIM-binding protein 1 and 2) can recognize a potential consensus of RXXPXXP and can likely bind to motifs such as the RQLPQL/VP, RLLPPTP, and RQLPQTP found in RIM1/2 (RAB3-interacting molecule 1 and 2). RIMBP1/2 has been shown to bind and couple RIM1/2 to voltage-gated Ca²⁺ channels [27]. Family 7 shows patterns such as PXXPX(K/R), (K/R)XPXXP, (K/R)XXPXXP, PXXPXX(K/R), and PXXP. Family 8 motifs, including PXXXXP, PXXXXPR, and PXXXXPR, highlight a selective array of proline-rich sequences. Family 9, characterized by PXXPX(K/R) and PXXPX(L/P), shows specificity for proline and amino acid residues such as K/R and L/P. Family 10 presents PX(P/A)XXR, PXXPXXP(K/R), PXXPX(K/R), RXX(K/R)P, and PPPPP motifs, illustrating a specific yet versatile proline-rich arrangement. These results highlight the complex yet diverse nature of PRMs across families controlling specific protein interactions and functions.

We observed an overlap of PRM sequences from families 5 and 6 with family 1, suggesting potential similarities and shared binding motifs within their respective SH3 domain interactions. In our phylogenetic classification based on SH3 domain specificity for PRM, family 1 interacts with RX(L/A)PXXP, RXXPXXP, KXX(L/A)PXXP, and PXXP motifs, whereas family 5 has specificity for PXPXXP and family 6 for RXXPXXP motifs. The biological interpretation of this overlap suggests a potential convergence or similarity in binding preferences among these families despite their specific motifs. Such overlapping PRM sequences imply a nuanced relationship in which different SH3 domain families may exhibit distinct specificities yet recognize certain common motifs. This observation may indicate functional redundancies, cooperative interactions, or shared regulatory pathways among these SH3 domain families in cellular processes.

3.3. Affinity and Selectivity of the SH3 Family Proteins for SOS1 PRP

The intriguing recognition pattern of PRMs observed in SOS1, a particularly PRM-rich protein, caught our attention. SOS1 shows co-occurrence of 13 out of a total of 14 PRMs (including sequences such as PPPP, XPPX, PXP, PXPXP, PPXPP, PXXP, PXXPX[KR], [KR]XXPXXP, PXXPXXP, PXXXXP, PXXXXPXXP, PXXXXPR, and PXXXXXP), as shown in Table S5. This distinct pattern, showing selectivity across all SH3 families, motivated us to use SOS1 as a comprehensive model to uncover the recognition mechanisms of established SH3 proteins in the human proteome. Therefore, 25 SH3 domains from different SH3 families were selected (Figure 1), cloned, purified as GST fusion proteins, and used for PRP binding analysis (Table S1). We selected at least one representative SH3 domain per defined family concerning accessibility and experimental viability (see Tables S3 and S4). In addition, we selected 10 different PRPs from the proline-rich domain (RPD) of the SOS1 protein (Table S5). This collection was designed to cover the full spectrum of PRM types (P1–P10; Table S2). Two reference peptides were included as controls: RP1, a well-studied SOS1 derivative encompassing part of P3, and RP2, a peptide derived from the N-terminal extension of the RHO GTPase WRCH1/RHOU (Table S2). The 12 PRPs were labeled with FITC to assess their binding capacities with purified GST fusion proteins of the 25 SH3 domains using fluorescence dot blot and polarization analysis (Figure 2A).

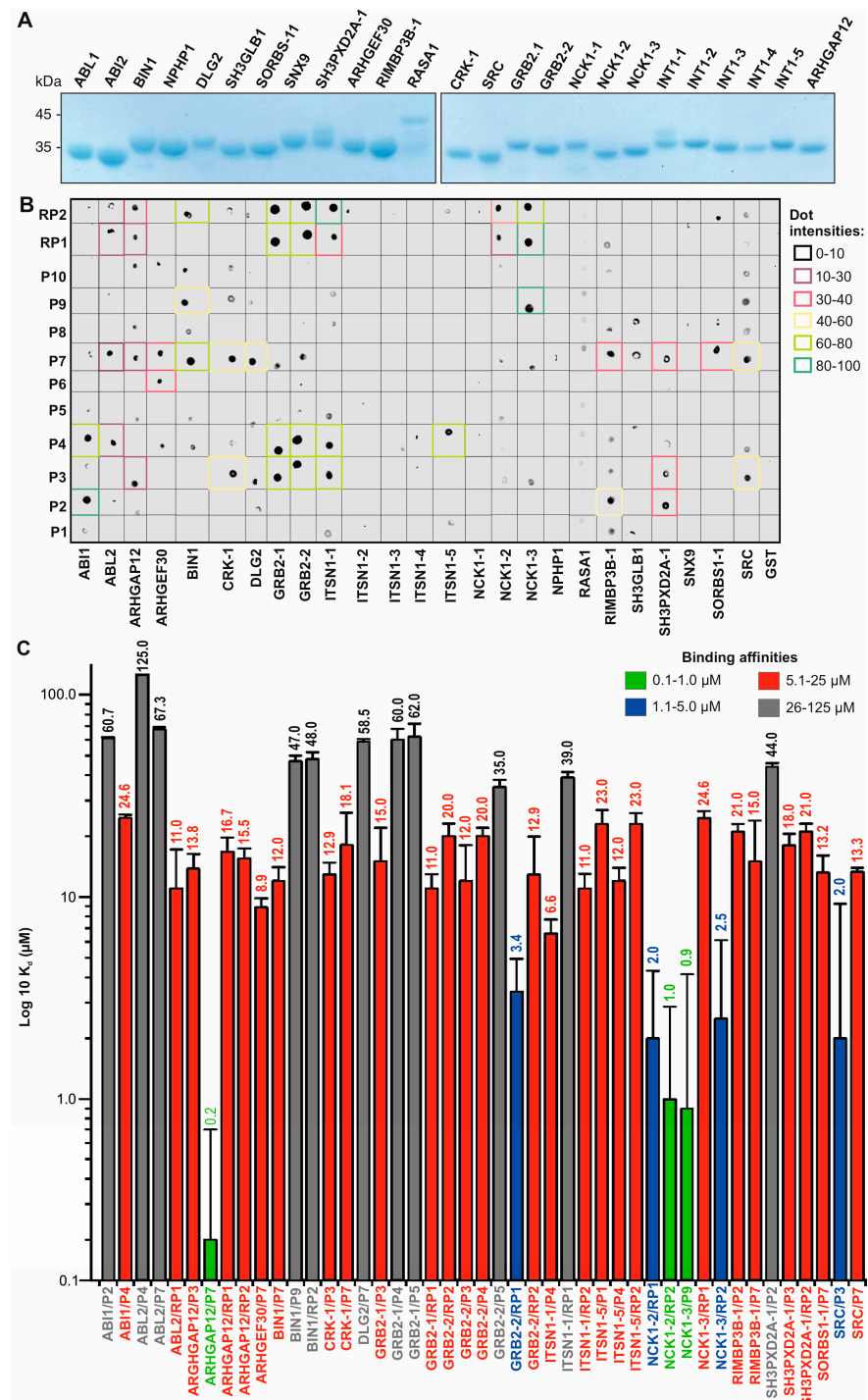


Figure 2. Evaluation of the binding selectivity of SH3 domain representatives with different PRP types. **(A)** Coomassie brilliant blue-stained SDS gels show purified SH3 domains as GST fusion proteins. **(B)** Fluorescence dot blots revealed the variable binding strengths of 12 fluorescent PRPs with 25 GST-SH3 domains. Dot intensities are categorized into five groups ranging from 0 (black) to 100 (dark green). **(C)** Bar graphs show the evaluated dissociation constants (K_d) for the selected SH3 domain-PRP interactions determined by fluorescence polarization (Table S6; Figure S6). The color codes indicate the K_d values, classified into high affinity (green), intermediate affinity (blue), low affinity (red), and very low affinity (black).

The binding of 12 FITC-labeled PRPs to 25 GST-SH3 domains was qualitatively analyzed by combining GST pull-down and dot blot assays. GST protein alone was used as a negative control. In a previous study [22], we showed that fluorescein labeling does not affect the interaction of proline-rich peptides with the SH3 domains. Different binding strengths were observed among the proteins tested, particularly with the P2, P3, P4, P7, and P9 peptides (Figure 2A). The strongest interactions (dot intensity >80) were between ABI1 and P2, ITSN1-1 and RP2, and NCK1-3 and P9 and RP1. In contrast, no PRP binding (dot intensity 0) was detected for NCK1-1, NPHP1, RASA1, SH3GLB-1, SNX9, ITSNS1-2, ITSNS1-3, and ITSNS1-4. These data provide valuable insight into the varying degrees of interaction across the panel of PRPs and SH3 domains tested (Figure 2A).

Fluorescence polarization measurements were performed to determine the binding affinities of SH3 domain–PRP interaction pairs from the dot blot analysis. SH3 proteins were titrated at increasing concentrations against fluorescent PRPs, which were kept at a constant concentration of 0.2 μM . GRB2-2^{W193K}, which is defective in the binding of PRPs such as RP1, was used as a negative control as previously described [22]. Interestingly, none of the PRPs we examined showed any binding for seven SH3 domains: ITSN1-2/-3/-4, NPHP1, RASA1, SH3GLB1, and SNX9 (Table S6). The resulting data (Figure S6) allowed the evaluation of equilibrium dissociation constants (K_d) for 45 interactions between the SH3 domains and the PRPs (Figure 2B; Table S6). In particular, the results confirmed that the peptides P2, P3, P4, P7, and P9 were associated with approximately 17 SH3 domains. The K_d values determined were categorized into four affinity levels (Figure 2B; Table S6): high (0.1 to 1.0 μM ; green), intermediate (1.1 to 5 μM ; blue), low (5.1 to 25 μM ; red), and very low (26 to 125 μM ; black). Whereas previously reported SH3 domain–PRM interactions exhibited micromolar affinities, our results revealed interactions with nanomolar affinities in some cases. The most notable and novel pairs of interaction were ARHGAP12/P7, NCK1-3/P9, and NCK1-2/RP2, which had affinities in the nanomolar range of 0.2, 0.9, and 1.0 μM , respectively.

3.4. Non-Conserved Residues Define the Selectivity and Affinity of SH3 Domain–PRM Interactions

To better understand the role of the residues of SH3 domains in selective binding to PRPs, we have generated a multiple sequence alignment (Figure S7). It highlights the conserved and variable residues that are likely to be critical for the selectivity and affinity of the SH3 domain–PRM interactions. The importance of variable residues was investigated by specifically selecting ARHGAP12 and NCK1-3 for mutational analysis due to their high binding affinities of 0.2 and 0.9 μM for P7 and P9, respectively. In contrast, ABL2 and BIN1 were selected for their very low affinity for P7 and P9, respectively. Two different sets of mutations were generated by substituting a combination of amino acids from ARHGAP12 and NCK1-3 for ABL2 and BIN1 and vice versa (Figure 3A; Table S1).

Comparative fluorescence polarization measurements between wt and mutant SH3 domains (Figure S8) revealed that variable residues determine selectivity and affinity. The determined K_d values of 14 and 23 μM showed a drastic reduction in the binding affinity of the Set-1 mutants of ARHGAP12 and NCK1 by 115-fold and 15.6-fold for P7 and P9, respectively (Figure 3B). This demonstrates the importance of the selected variable residues for the PRP interactions, especially because ABL2^{Set-1} and BIN1^{Set-1} showed a 9- and an 11-fold increase in binding to P7 and P9, respectively (Figure 3B). In light of this result, we decided to investigate another set of variable residues (Set-2; Figure 3A). The binding affinity for ARHGAP12^{Set-2} was reduced 47-fold, indicating the critical role of these residues in determining the selectivity and affinity of ARHGAP12 for P7 (Figure 3B). However, NCK1-3^{Set-2} did not differ from the NCK1-3^{wt} in terms of P9 binding, suggesting that the Set-2 residues are not critical for NCK1-3/P9 interaction (Figure 3B).

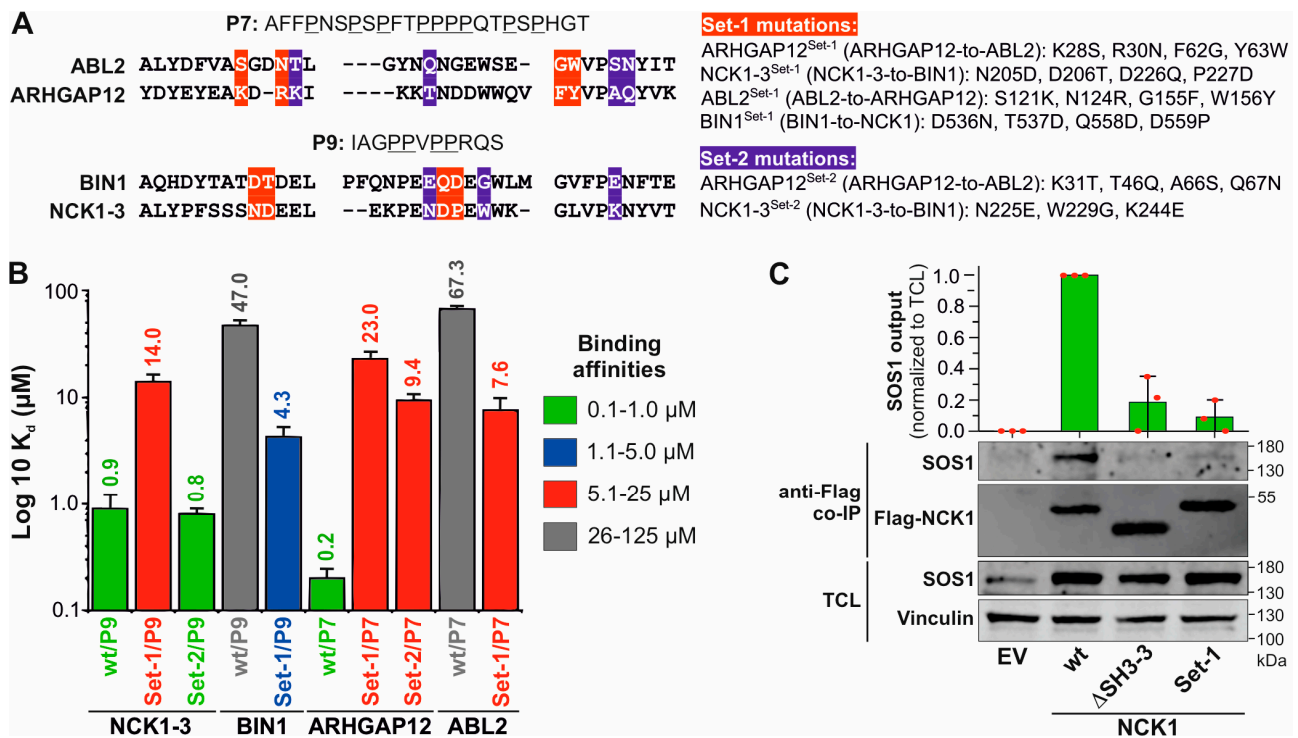


Figure 3. Variable residues determine the specificity and affinity of SH3 domain–PRP interactions. (A) A sequence alignment of the PRM-binding residues of ABL2, ARHGAP12, BIN1, and NCK1-3 is extracted from the alignment shown in Figure S7. Residues in red (Set-1) and blue (Set-2) are variable residues (left panel) and are the subject of mutation analysis (right panel). (B) K_d values determined by fluorescence polarization partially revealed shifts in the binding affinities of the investigated SH3 domain mutants for P7 and P9, respectively, with decreased affinity for ARHGAP12^{Set1} and ARHGAP12^{Set2}, and NCK1^{Set-1} and increased affinity for ABL2^{Set1} and BIN1^{Set-1}. NCK1^{Set-2} had the same K_d value as NCK1^{wt}. (C) SOS1 co-immunoprecipitated (Co-IP) with NCK1^{wt} but not with NCK1-3^{ΔSH3-3} and NCK1^{Set-1} overexpressed in CHO-K1 cells.

To investigate a potential SOS1 binding of ARHGAP12 and NCK1 and to assess the relevance of variable residues in the SH3 domain–PRMs interaction in cells, CHO-K1 cells were co-transfected with wt and mutant variants of NCK1 and ARHGAP12 together with HA-SOS1 containing P7 and P9 at its C-terminal PRD (Figure 4). Co-immunoprecipitation (Co-IP) with anti-Flag beads was performed to investigate the possible interaction of ARHGAP12 and NCK1 with HA-SOS1 (Figures S9 and S10). As shown in Figure 3C, all three NCK1 proteins, wt, ΔSH3-3, and Set-1, were immunoprecipitated, but SOS1 was only co-immunoprecipitated with NCK1^{wt} and not with NCK1^{ΔSH3-3} and NCK1^{Set-1} (Figure 3C). No ARHGAP12-SOS1 interaction was observed in similar experiments (Figure S10). Taken together, our data not only highlight the essential role of these flanking residues in determining the selectivity and affinity of SH3 domains for their cognate PRPs but also provide unprecedented insight into a potential SOS1-NCK1 interaction in cells.

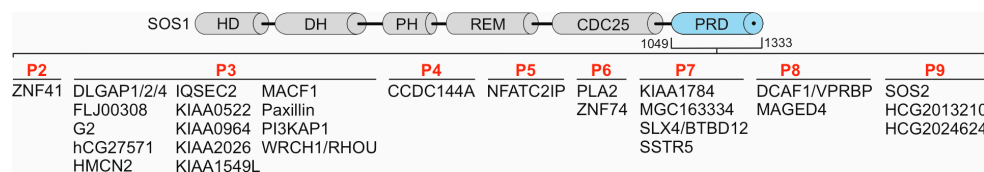


Figure 4. Proteins containing PRMs homologous to P2–P9 derived from the SOS1 PRD. BLAST searches with each SOS1 PRP identified several human proteins (see Table S7) with a high degree of sequence similarity to P2–P9. CDC25, cell division cycle 25; DH, DBL homology domain; HD, histone-like domain; PH, pleckstrin domain; REM, RAS exchange motif; PRD, proline-rich domain.

3.5. SH3 Domain–PRP Relationships beyond SOS1

We used the position-specific iterated BLAST (PSI-BLAST) algorithm to perform an analysis of the SOS1 PRPs in the human proteome using the NCBI (National Center for Biotechnology Information, Bethesda, MD, USA) database as a reference. Our goal was to identify homologous peptides from proteins other than SOS1 as potential interaction partners for the SH3 domains investigated in this study. In our analysis, we considered alignments with percentage identities ranging from 98% to 100% and with E values up to 10, which, although less stringent, may still indicate potential similarities between the SOS1 PRPs and other protein sequences. More than 30 proteins were found with partial sequence identities with P2 to P9 (Table S7; Figures 4 and S11) and 0 with P1 and P10 (Refer to Table S1). Some of these proteins contain multiple PRM repeats (Figure S11), for example, IQSEC2, paxillin, DLGAP1, PI3KAP1, and WRCH1/RHO within P3 (Motif: PVPPPVP), SSTR5 and SLX4 within P7 (Motif: PPPPQTP), DCAF1 and MAGED4 within P8 (Motif: HLPSPPP), and SOS2 and HCG2013210 within P9 (Motif: PPVPPRQ). This suggests that the SH3 interactions with the PRPs characterized in this study go beyond SOS1 as a binding partner, although the binding specificities of the listed proteins (Figure 4) remain to be investigated. However, several studies confirm the interactions of the identified proteins with SH3DCPs, including zinc finger proteins with p130Cas [28], MACF1 with Spectrin [29], DLGAP with DLG [30], WRCH1/RHO with GRB2 [31,32], Paxillin with SRC [33,34], and SSTR5 with Homer, Dynamin, IRSp53, and Cortactin [35].

4. Discussion

SH3 domains are critical in multiple signaling pathways; they interact with diverse proteins involved in apoptosis, proteasomal degradation, endocytosis, and with SRC family protein tyrosine kinases, influencing downstream processes including proliferation, cell survival, growth, actin reorganization, and cell migration [4]. The broad influence of SH3 domains on cellular functions raises fundamental questions about the specificity of their interaction networks. Previous research has also highlighted the importance of proline amino acids in forming the polyproline type II helix (PPII) conformation, which provides a binding pocket for SH3 domain residues, particularly from the RT and n-SRC loops [17]. Despite the discovery of 14 PRM consensus sequences in the human proteome (Table S5), the specificity pattern of the interaction of SH3 domains with PRMs is still unclear. Understanding the molecular basis of SH3 domain–PRM interactions is crucial to gain insight into how these interactions regulate signaling pathways.

Interface residues often play an important role in the functional outcome of protein interactions. Many diseases, including cancer and neurodegenerative disorders, are associated with aberrant protein–protein interactions resulting from mutated interacting residues. Abnormal SH3 domain interactions in cancer fuel dysregulated signaling that drives uncontrolled cell proliferation, survival, tumor growth, metastasis, apoptosis evasion, and resistance to anticancer therapies [4,6]. In neurodegenerative diseases like Alzheimer’s and Parkinson’s, these interactions disrupt signaling pathways, leading to the accumulation of misfolded proteins and neuronal degeneration. Altered SH3 domain interactions in synaptic proteins affect neurotransmission and synaptic plasticity [7]. In autoimmune diseases, aberrant SH3 domain interactions contribute to immune cell activation and tissue

damage, as immune cells mistakenly target healthy tissues [12,13]. Understanding the interface residues involved in SH3 domain–PRM interactions can thus provide insight into disease mechanisms and potential targets for intervention. In drug development, knowledge of interface residues is essential for designing molecules that can disrupt or modulate specific protein–protein interactions. Targeting these residues can lead to the development of therapeutic agents for the treatment of various diseases.

In recent years, several studies have been devoted to elucidating the diverse nature of SH3 domain interactions. Cesareni and colleagues used a novel chip technology to perform high-throughput qualitative analyses, revealing a variety of human SH3 domains that fall into two categories: those characterized by classical proline-rich core motifs accompanied by positively charged amino acids and atypical ones lacking the core motif [20]. Nevertheless, the diversity of PRM selectivity patterns is evident among all human SH3 domains. Furthermore, a comprehensive analysis of SH3 domain interactions in the evolution of four yeast species, *Saccharomyces cerevisiae*, *Ashbya gossypii*, *Candida albicans*, and *Schizosaccharomyces*, revealed that nearly 75 percent of SH3 families identified within the phylogenetic tree have a conserved SH3 specificity profile over 400 million years of evolution [36]. Utilizing the evolutionary relationships of peptide recognition domains in eukaryotes, we identified common structural features and ancestry that allowed us to group SH3 domains into similar binding preference families. This comprehensive investigation aimed to clarify the specificity profiles of SH3 PRMs within the human proteome through categorization based on the phylogenetic tree of SH3DCPs.

To provide an accurate specificity map of SH3 domains, we performed deep phylogenetic analyses coupled with computational analysis of the related structural data. Initial evolutionary analysis of the sequence–structure–function of full-length SH3 domains was unsuccessful due to the presence of SH3 regions that do not interact with PRMs. Within these, each SH3 domain exhibited variation in binding specificity to PRMs, making the characterization of SH3DCPs infeasible. Instead of relying on full-length sequences, we focused on binding residues that directly interact with PRMs, as revealed by sequence alignments coupled with analysis of the published SH3 domain structures in complex with PRMs. This refined phylogenetic approach led to the identification of ten distinct families based on both the structural and biochemical assessments of SH3 domain–PRM interactions and their distribution within the phylogenetic tree. This approach facilitates the assessment of cross-reactivity among SH3 domain recognition sites for PRMs, a phenomenon also observed in previous studies examining the yeast SH3 domain peptide library [37] and including SH3 domains that recognize multiple PRMs. The findings of this study highlight the fact that each SH3 domain family interacts with different but distinct sets of PRMs.

Considering the significant involvement of SOS1 in interactions with SH3DCPs such as GRB2 [38], ITSN1 [39], NCK1 [40], and ABI1 [41], along with its comprehensive coverage of all known PRMs in the human proteome (Tables S2 and S5), we decided to use SOS1 as a model for in-depth exploration of SH3 domain specificity in the realm of polyproline interactions. We performed in vitro studies with 25 representative SH3 domains selected from the phylogenetic tree. We performed low-throughput analyses, including pull-down assays, dot blotting, and fluorescence polarization, to investigate SH3 domain–PRP interactions. These investigations revealed novel interactions that had nanomolar affinities, which were subsequently confirmed by mutational studies.

The general concept of protein association is essential for describing protein–protein interactions in complexes, especially those with weak affinities in the micromolar range or transient interactions such as the SH3 domain with PRMs [42]. Mayer and Saksela noted that the limited selectivity of SH3 domains for PRMs implies that SH3-domain-mediated interactions may be highly dependent on external environmental factors [43]. In certain scenarios, the presence of additional surfaces on either the SH3 domain or the ligand it recognizes, along with the presence of either multiple SH3 domains or different domains within the same protein, or even the co-localization of two partners within a multi-protein complex, can cooperatively enhance SH3 domain–PRM specificity to a

significant degree [44]. This suggests that the low-affinity-region results in our study may be compensated by these scenarios, ultimately increasing the affinity and specificity of SH3 domain–PRM interactions. It has also been reported that the moderate affinities of SH3-domain-mediated interactions imply that the interactions have a high dynamic remodeling potential (rapid off-rates), depending on the subcellular localization and accessible binding partners [43]. This observation is consistent with our polarization data, which showed high K_d values for many low-affinity interactions.

The question of how the specificity of SH3 domain–PRM networks is achieved has been addressed by various research groups. It has been postulated that specificity in cells is not solely encoded by isolated SH3 domain–PRM partners but rather by the context in which the partners are presented as full-length proteins. Dione et al. have shown that the identity of the host protein and the position of the SH3 domains within their host are critical for interaction specificity, cellular functions, and key biophysical processes such as phase separation [45]. In addition, Zarrinpar et al. have shown that isolated SH3 domains can determine the interaction specificity between host SH3 domains [46]. This may also be true for certain high-affinity SH3 domain–PRM interactions, as shown in this study for the newly discovered interactions of NCK1-2, NCK1-3, and ARHGAP12 with WRCH1/RHOA-derived RP2 and SOS1-derived P9 and P7, respectively. NCK1 has been shown to modulate ITSN1-CDC42-WASP-dependent actin polymerization [47]. WRCH1/RHOA, a CDC42 homologous protein, encompasses an extended N-terminus that contains PRMs specific for various adaptor proteins, including GRB2, CRK, and NCK1 [32]. The association of these proteins with WRCH1/RHOA may not only determine its signaling specificity but may also regulate its activity in cells [48]. However, it should be noted that, in some cases, a negative effect of other domains on SH3 domain–PRM binding was observed. Notably, NCK1-3 showed more extensive protein interaction than the full-length NCK1 in immunoprecipitation experiments, possibly indicating a detrimental effect of the C-terminal SH2 domain on specific SH3-domain-mediated interactions. Furthermore, the spectrum of proteins associated with NCK1-3 is not simply the cumulative sum of proteins associated with individual SH3 domains [49].

A closer examination of the PRMs revealed that the positioning of the proline residues plays a critical role in the recognition of the SH3 domain, providing the structural basis for defining interaction specificity. The current results indicate that residues -2 , -1 , $+1$, and $+2$ are critical for the recognition of SH3 PRMs. In addition, adjacent positively charged residues contribute additional features that help stabilize the transient interaction [50]. For example, structural studies of the PI3K SH3 domain in association with the peptide RKLPPRPSK provided evidence for the role of adjacent non-proline residues such as Arg-1, Leu-3, and Arg-6 in contributing to the SH3 domain interaction [51]. Arginine residues at positions $R + 5$, $R + 6$, and $R + 7$ are thought to play an important role in enhancing the affinity of GRB2-SH3 domains for SOS1-PRM by contributing to the overall free energy of the interaction [52]. While the majority of SH3 domains interact with PRMs, there have been documented cases where the SH3 domain deviates from the typical classical proline-rich interaction pattern [4]. For example, the RASA1 (or p120RASGAP) SH3 domain specifically interacts with the catalytic arginine finger of the RHOGAP domain of DLC1, thereby competitively and potently inhibiting its RHOGAP activity [53]. Interestingly, none of the PRPs we examined showed any binding for seven SH3 domains: ITSN1-2/-3/-4, NPHP1, RASA1, and SH3GLB1 (Table S6). While our study comprehensively highlights the major SH3 domain–PRM interactions in the human proteome, the specificity and mechanism of the PRM-independent interaction of SH3 domains remain to be elucidated in further studies.

In our study, SOS1 was used as a PRP model for biophysical and bioinformatic analysis of the SH3 domain–PRM interaction landscape because the SOS1 PRD contains 13 out of 14 different classified proline-rich consensus sequence motifs (Table S5). This alternative model reveals a spectrum of interactions between different SOS1 PRPs and a number of SH3DCPs, including ABI1, ABL2, ARHGAP12, ARHGEF30 (Obscurin/OBSCN),

BIN1, CRK-1, DLG2, GRB2, ITSN1, NCK1, SRC, SH3PXD2A-1, and SORBS1-1. Among them, ABI1 [54], ITSN1 [55], SRC [56], NCK1 [57], GRB2 [58], CRK [59], and SH3PXD2A (TKS5) [60] have been previously established as SOS1 binding partners in cells. ABL1, but not ABL2, has also been shown to interact with SOS1 [61]. Importantly, the precise binding sites for most of these proteins have yet to be investigated. Our study not only elucidates the binding sites of these established SOS1 partners but also uncovers novel interactions, including ABL2, BIN1, DLG2, SORBS1, ARHGEF30, and ARHGAP12. In particular, a high affinity (<0.5 μ M) interaction was observed between ARHGAP12 and the P7 of SOS1, demonstrating the interplay between small GTPase regulators, GAPs, and GEFs. However, immunoprecipitation experiments with overexpressed ARHGAP12 in CHO-K1 cells did not confirm an interaction with SOS1 (Figure S10). The reliability of this result may depend on the expression level and affinity of other interaction partners, like accessory proteins, that could potentially bind more strongly and possibly in a multivalent manner. However, it is important to note that this result does not definitively rule out the existence of this interaction in cells, especially considering examples of reported GEF-GAP interactions. A study using immunoprecipitation and mass spectrometry unravels the intricate protein interaction networks involving the synaptic proteins SYNGAP1 (RASGAP), KALIRIN (RHOGEF), and AGAP2 (ARFGAP) in both the postsynaptic density (PSD) and non-PSD fractions of the adult mouse cortex. This investigation sheds light on their role in the organization of GAP and GEF protein families and their associations with proteins associated with intellectual disability and psychiatric disorders [62]. In conclusion, to confirm the significance and broader implications of these novel findings, additional studies within the cellular context are warranted.

Predicting the potential interaction of SH3 domains with PRMs by considering their sequence specificities, as we did in this study, is a promising approach in the field of molecular biology and protein–protein interactions. In pursuing this goal, we are faced with an interesting challenge: the identification of binding affinities between SH3 domains and peptides containing PRMs. The sequence specificity of these interactions is paramount, as SH3 domains exhibit diverse binding preferences that depend on the PRMs present in the peptides.

5. Conclusions

SH3 domains are small protein interaction modules that are involved in numerous fundamental cellular processes and associated with the development of several diseases, including Joubert syndrome, leukemia, lymphoma, Usher syndrome or non-syndromic deafness, centronuclear myopathy, schizophrenia, and other neurodevelopmental disorders [4]. Over the past three decades, researchers have focused on how members of the SH3DCP superfamily selectively recognize and bind to their associated PRM-containing proteins.

To systematically address this question, we first extracted 298 SH3 domains in 221 SH3DCPs ranging in size from 13 (small monodomain proteins) to 720 (large multidomain proteins) kilodaltons [4]. The subsequent evolutionary multidomain relationship of the SH3DCP superfamily not only allowed us to functionally classify them into thirteen families but also provided new insights into their diverse roles and interactions in cellular signaling processes, as well as their relevance to various diseases when exploiting the modular interactions of SH3 proteins as drug targets [4]. In the present study, we have incorporated the available sequence, structure, and interaction data into a phylogenetic tree (Figure 1) that groups 298 SH3 domains in the human proteome into 10 families related to the PRM binding interface. These families are aligned with the frames necessary for the interaction of their respective potential PRMs. Mutational analysis suggests the critical role of non-conserved sequences within each SH3 family in defining the specificity and affinity of their interactions with specific PRMs. This investigation highlights that the recognition mechanisms of SH3 proteins across the human proteome are not only influenced by PRMs but also by the core binding site within the SH3 domain. The study of the PRM-binding residues of SH3

domains revealed a significant relationship between individual SH3 domains and specific PRMs, culminating in a detailed map of their associations.

In this study, we performed a comprehensive analysis by comparing published biochemical interaction data, available structural information, and sequence alignments. The goal was to identify specificity-determining residues within PRMs that are critical for interacting with different SH3 domains. A phylogenetic tree based solely on the interacting interface of SH3 domains allowed us to categorize distinct families of SH3 domains within the human proteome, each interacting specifically with unique PRMs. Subsequent mutational analysis supported our categorization and hypothesis by demonstrating that the non-conserved interface sequences within each family are critical in defining the specificity of their interaction with PRMs. The different interface residues within each family were found to determine the affinity and specificity of each protein towards PRMs. In particular, the discovery of common PRMs in two different SH3 domain families underscores the importance of other residues (designated X) beyond proline in determining interaction specificity. It is generally accepted that prolines serve as recognition sites and the backbone of interactions, while X residues define specificities. A comparison of the PRM consensus sequences of the SOS1-homologous proteins reveals other common amino acids, such as V in P3 and P9, Q in P7 and P9, R in P3, P4, P5, and P9 (Figure S11). Notably, the third SH3 domain of NCK1 (NCK1-3) tightly bound P9 but none of the other PRPs tested in this study (Table S6); this implies that residues other than valine and arginine in P9 may dictate the specificity of the NCK1–P9 interaction. In addition, the sequence motif ²²⁴ENDPEW of NCK1-3 (Figure 3A) contains three negatively charged residues that may be in electrostatic contact with the R in P9. In contrast, NCK1-1 and NCK1-2, which do not bind P9, contain a lysine instead of glutamate or aspartate within this sequence motif, which counteracts a P9 interaction. An important next step in elucidating the specificity of the interaction of the SH3 domain with the PRM at the atomic level is to analyze the nearly 800 available experimental structures containing SH3 domains. This will be performed by generating homology models and correlating and combining them with the measured affinities and known binding properties of SH3 domain–PRM complexes.

A total of 7 out of the 25 examined SH3 domains showed no interaction with any of the 12 selected PRPs (Table S6). The tested peptides cover 13 of the 14 recognized proline-rich consensus sequence motifs, suggesting that they may bind in a proline-independent manner. (Table S5). It is suggested that these SH3 domains may bind in a proline-independent manner. SH3 domains in several studies exhibit an extended repertoire of binding sequences, known as proline-independent binding, allowing SH3DCPs to mediate a broader array of interactions [4,5]. An example of atypical binding is the SH3 domain of RASA1, the RAS-specific GAP (p120RASGAP), which interacts with the catalytic GAP and kinase domains of DLC1, thereby inhibiting its activity [53,63]. Another example is the selective interaction of the GADS/GRAB2L SH3 domain with an RXXK motif of SLP-76 [64]. These types of SH3 interactions with a non-canonical binding mode add to the complexity of understanding protein–protein interactions involving SH3 domains.

The SH3 domains are modular building blocks across all five kingdoms of life and viruses and play a critical role in facilitating inter- and intramolecular interactions and functional interplay within domain-specific interaction networks. SH3DCPs, except the MIA family with a single SH3 domain, are multi-domain proteins [4]. Several recent studies have shown that SH3 domains have an extended repertoire of binding sequences, known as proline-independent binding [4,17,53]. This allows SH3DCPs to mediate a wider range of interactions. However, a quantitative description of the communication between two different sites in a multivalent protein is still challenging. In some cases, the task reaches another level of complexity, such as the interaction of the two SH3 domains of GRB2 with SOS1. Not only the association of the two functional SH3 domains of GRB2 with SOS1 but also the physical interactions between the two SH3 domains are required to allosterically control SOS1 activation [22]. Such a regulatory mechanism involves a series

of intramolecular interactions that are further amplified by the interaction of GRB2 with upstream ligands [65].

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cells13020195/s1>; Table S1: Proteins used in this study; Table S2: Peptides used in this study; Table S3: Published structures of the SH3 domain–PRM complexes; Table S4: Published dissociation constants (K_d) determined for the SH3 domain–PRP interactions; Table S5: PRM classification and occurrence in SOS1 PRD; Table S6: Dissociation constants (K_d) for the SH3 domain–PRP interactions determined in this study; Table S7: Proteins containing PRMs homologous to peptides 2–9 derived from the SOS1 PR; Figure S1: Evolutionary sequence–structure–function relationships of SH3 domains; Figure S2: Alignment of SH3 domain sequences; Figure S3: PRM-binding residues in human SH3 domains; Figure S4: Exploring evolutionary connections of PRM-binding residues in SH3 domains; Figure S5: Analysis of SH3 domain–PRMs interaction specificity across various SH3 domain families in the human proteome; Figure S6: Fluorescence polarization measurements of SH3 domain interactions with SOS1 and reference peptides; Figure S7: Sequence alignment of PRM-binding residues in representative SH3 domains interacting with specific PRP; Figure S8: Mutational analysis of the SH3 domain–fluorescent PRP interactions using fluorescence polarization. Figure S9: Co-immunoprecipitation of NCK1 with SOS1 in CHO-K1 cells; Figure S10: No co-immunoprecipitation of SOS1 with ARHGAP12 in CHO-K1 cells; Figure S11: SOS1 homologous PRM sequences found in other human proteins.

Author Contributions: M.R.A. designed and coordinated the study. N.S.K.J., M.M. and M.R.A. designed the study and drafted the manuscript. E.M.E. and L.B. synthesized the peptides. N.S.K.J., M.M., R.D. and M.R.A. designed, performed, and analyzed the experiments. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the German Research Foundation (DFG; grant number: AH 92/8-1), the European Network on Noonan Syndrome and Related Disorders (NSEuroNet; grant number: 01GM1602B), and the Foundation for Ageing Research of the Heinrich Heine University (grant number 701.810.845).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the results of this study are available both as supplementary material and, upon reasonable request, are available from the corresponding author.

Acknowledgments: We thank our colleagues Doreen Floß, Roland Piekorz, and Jürgen Scheller for the discussions. We thank Soyoun Kim for providing the p3XFLAG-CMV ARHGAP12 plasmid.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Mabonga, L.; Kappo, A.P. Protein-protein interaction modulators: Advances, successes and remaining challenges. *Biophys. Rev.* **2019**, *11*, 559–581. [[CrossRef](#)] [[PubMed](#)]
- Musacchio, A.; Noble, M.; Pauptit, R.; Wierenga, R.; Saraste, M. Crystal structure of a Src-homology 3 (SH3) domain. *Nature* **1992**, *359*, 851–855. [[CrossRef](#)] [[PubMed](#)]
- Stahl, M.L.; Ferez, C.R.; Kelleher, K.L.; Kriz, R.W.; Knopf, J.L. Sequence similarity of phospholipase C with the non-catalytic region of src. *Nature* **1988**, *332*, 269–272. [[CrossRef](#)] [[PubMed](#)]
- Mehrabipour, M.; Jasemi, N.S.K.; Dvorsky, R.; Ahmadian, M.R. A Systematic Compilation of Human SH3 Domains: A Versatile Superfamily in Cellular Signaling. *Cells* **2023**, *12*, 2054. [[CrossRef](#)] [[PubMed](#)]
- Teyra, J.; Huang, H.; Jain, S.; Guan, X.; Dong, A.; Liu, Y.; Tempel, W.; Min, J.; Tong, Y.; Kim, P.M. Comprehensive analysis of the human SH3 domain family reveals a wide variety of non-canonical specificities. *Structure* **2017**, *25*, 1598–1610.e3. [[CrossRef](#)] [[PubMed](#)]
- Berndt, S.; Gurevich, V.V.; Iverson, T. Crystal structure of the SH3 domain of human Lyn non-receptor tyrosine kinase. *PLoS ONE* **2019**, *14*, e0215140. [[CrossRef](#)] [[PubMed](#)]
- Bircher, J.E.; Koleske, A.J. Trio family proteins as regulators of cell migration and morphogenesis in development and disease—mechanisms and cellular contexts. *J. Cell Sci.* **2021**, *134*, jcs248393. [[CrossRef](#)]
- Hildebrandt, F.; Otto, E.; Rensing, C.; Nothwang, H.G.; Vollmer, M.; Adolphs, J.; Hanusch, H.; Brandis, M. A novel gene encoding an SH3 domain protein is mutated in nephronophthisis type 1. *Nat. Genet.* **1997**, *17*, 149–153. [[CrossRef](#)]

9. Saunier, S.; Calado, J.; Benessy, F.; Silbermann, F.; Heilig, R.; Weissenbach, J.; Antignac, C. Characterization of the NPHP1 locus: Mutational mechanism involved in deletions in familial juvenile nephronophthisis. *Am. J. Hum. Genet.* **2000**, *66*, 778–789. [[CrossRef](#)]
10. Wu, T.; Shi, Z.; Baumgart, T. Mutations in BIN1 associated with centronuclear myopathy disrupt membrane remodeling by affecting protein density and oligomerization. *PLoS ONE* **2014**, *9*, e93060. [[CrossRef](#)]
11. Hohendahl, A.; Roux, A.; Galli, V. Structural insights into the centronuclear myopathy-associated functions of BIN1 and dynamin 2. *J. Struct. Biol.* **2016**, *196*, 37–47. [[CrossRef](#)] [[PubMed](#)]
12. Manso, J.A.; Marcos, T.; Ruiz-Martín, V.; Casas, J.; Alcón, P.; Sánchez Crespo, M.; Bayón, Y.; de Pereda, J.M.; Alonso, A. PSTPIP1-LYP phosphatase interaction: Structural basis and implications for autoinflammatory disorders. *Cell Mol. Life Sci.* **2022**, *79*, 131. [[CrossRef](#)] [[PubMed](#)]
13. Starnes, T.W.; Bennin, D.A.; Bing, X.; Eickhoff, J.C.; Grahf, D.C.; Bellak, J.M.; Seroogy, C.M.; Ferguson, P.J.; Huttenlocher, A. The F-BAR protein PSTPIP1 controls extracellular matrix degradation and filopodia formation in macrophages. *Blood* **2014**, *123*, 2703–2714. [[CrossRef](#)] [[PubMed](#)]
14. Qin, Y.; Du, Y.; Chen, L.; Liu, Y.; Xu, W.; Liu, Y.; Li, Y.; Leng, J.; Wang, Y.; Zhang, X.Y.; et al. A recurrent SHANK1 mutation implicated in autism spectrum disorder causes autistic-like core behaviors in mice via downregulation of mGluR1-IP3R1-calcium signaling. *Mol. Psychiatry* **2022**, *27*, 2985–2998. [[CrossRef](#)] [[PubMed](#)]
15. Leblond, C.S.; Heinrich, J.; Delorme, R.; Proepper, C.; Betancur, C.; Huguet, G.; Konyukh, M.; Chaste, P.; Ey, E.; Rastam, M. Genetic and functional analyses of SHANK2 mutations suggest a multiple hit model of autism spectrum disorders. *PLoS Genet.* **2012**, *8*, e1002521. [[CrossRef](#)] [[PubMed](#)]
16. Ball, L.J.; Kühne, R.; Schneider-Mergener, J.; Oschkinat, H. Recognition of Proline-Rich Motifs by Protein-Protein-Interaction Domains. *Angew. Chem. Int. Ed.* **2005**, *44*, 2852–2869. [[CrossRef](#)] [[PubMed](#)]
17. Li, S.S.-C. Specificity and versatility of SH3 and other proline-recognition domains: Structural basis and implications for cellular signal transduction. *Biochem. J.* **2005**, *390*, 641–653. [[CrossRef](#)]
18. Cesareni, G.; Gimona, M. *Modular Protein Domains*; Wiley: Hoboken, NJ, USA, 2005.
19. Cámara-Artigas, A.; Martínez-Rodríguez, S.; Ortiz-Salmerón, E.; Martín-García, J.M. 3D domain swapping in a chimeric c-Src SH3 domain takes place through two hinge loops. *J. Struct. Biol.* **2014**, *186*, 195–203. [[CrossRef](#)]
20. Carducci, M.; Perfetto, L.; Briganti, L.; Paoluzi, S.; Costa, S.; Zerweck, J.; Schutkowski, M.; Castagnoli, L.; Cesareni, G. The protein interaction network mediated by human SH3 domains. *Biotechnol. Adv.* **2012**, *30*, 4–15. [[CrossRef](#)]
21. Panni, S.; Dente, L.; Cesareni, G. In vitro evolution of recognition specificity mediated by SH3 domains reveals target recognition rules. *J. Biol. Chem.* **2002**, *277*, 21666–21674. [[CrossRef](#)]
22. Kazeminejad, N.S.; Herrmann, C.; Magdalena Estirado, E.; Gremer, L.; Willbold, D.; Brunsveld, L.; Dvorsky, R.; Ahmadian, M.R. The intramolecular allostery of GRB2 governing its interaction with SOS1 is modulated by phosphotyrosine ligands. *Biochem. J.* **2021**, *478*, 2793–2809. [[CrossRef](#)] [[PubMed](#)]
23. Mayer, B.J. SH3 domains: Complexity in moderation. *J. Cell Sci.* **2001**, *114*, 1253–1263. [[CrossRef](#)] [[PubMed](#)]
24. Kay, B.K. SH3 domains come of age. *FEBS Lett.* **2012**, *586*, 2606–2608. [[CrossRef](#)]
25. Landgraf, C.; Panni, S.; Montecchi-Palazzi, L.; Castagnoli, L.; Schneider-Mergener, J.; Volkmer-Engert, R.; Cesareni, G. Protein interaction networks by proteome peptide scanning. *PLoS Biol.* **2004**, *2*, e14. [[CrossRef](#)] [[PubMed](#)]
26. Bae, D.J.; Seo, J.; Kim, S.Y.; Park, S.Y.; Do Yoo, J.; Pyo, J.H.; Cho, W.; Cho, J.Y.; Kim, S.; Kim, I.S. ArhGAP12 plays dual roles in Stabilin-2 mediated efferocytosis: Regulates Rac1 basal activity and spatiotemporally turns off the Rac1 to orchestrate phagosome maturation. *Biochim. Biophys. Acta Mol. Cell Res.* **2019**, *1866*, 1595–1607. [[CrossRef](#)] [[PubMed](#)]
27. Hibino, H.; Pironkova, R.; Onwumere, O.; Vologodskaya, M.; Hudspeth, A.J.; Lesage, F. RIM binding proteins (RBPs) couple Rab3-interacting molecules (RIMs) to voltage-gated Ca(2+) channels. *Neuron* **2002**, *34*, 411–423. [[CrossRef](#)] [[PubMed](#)]
28. Nakamoto, T.; Yamagata, T.; Sakai, R.; Ogawa, S.; Honda, H.; Ueno, H.; Hirano, N.; Yazaki, Y.; Hirai, H. CIZ, a zinc finger protein that interacts with p130(cas) and activates the expression of matrix metalloproteinases. *Mol. Cell Biol.* **2000**, *20*, 1649–1658. [[CrossRef](#)]
29. Salem, D.; Fecek, R.J. Role of microtubule actin crosslinking factor 1 (MACF1) in bipolar disorder pathophysiology and potential in lithium therapeutic mechanism. *Transl. Psychiatry* **2023**, *13*, 221. [[CrossRef](#)]
30. Rasmussen, A.H.; Rasmussen, H.B.; Silahatoglu, A. The DLGAP family: Neuronal expression, function and role in brain disorders. *Mol. Brain* **2017**, *10*, 43. [[CrossRef](#)]
31. Bustelo, X.R.; Sauzeau, V.; Berenjano, I.M. GTP-binding proteins of the Rho/Rac family: Regulation, effectors and functions in vivo. *Bioessays* **2007**, *29*, 356–370. [[CrossRef](#)]
32. Risse, S.L.; Vaz, B.; Burton, M.F.; Aspenström, P.; Piekorz, R.P.; Brunsveld, L.; Ahmadian, M.R. SH3-mediated targeting of Wrch1/RhoU by multiple adaptor proteins. *Biol. Chem.* **2013**, *394*, 421–432. [[CrossRef](#)] [[PubMed](#)]
33. López-Colomé, A.M.; Lee-Rivera, I.; Benavides-Hidalgo, R.; López, E. Paxillin: A crossroad in pathological cell migration. *J. Hematol. Oncol.* **2017**, *10*, 50. [[CrossRef](#)] [[PubMed](#)]
34. Webb, D.J.; Schroeder, M.J.; Brame, C.J.; Whitmore, L.; Shabanowitz, J.; Hunt, D.F.; Horwitz, A.R. Paxillin phosphorylation sites mapped by mass spectrometry. *J. Cell Sci.* **2005**, *118*, 4925–4929. [[CrossRef](#)] [[PubMed](#)]

35. Kreienkamp, H.-J.; Liew, C.W.; Bächner, D.; Mameza, M.-G.; Soltau, M.; Quitsch, A.; Christenn, M.; Wente, W.; Richter, D. Physiology of somatostatin receptors: From genetics to molecular analysis. In *Somatostatin*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 185–202.
36. Verschuere, E.; Spiess, M.; Gkourtsa, A.; Avula, T.; Landgraf, C.; Mancilla, V.T.; Huber, A.; Volkmer, R.; Winsor, B.; Serrano, L. Evolution of the SH3 domain specificity Landscape in Yeasts. *PLoS ONE* **2015**, *10*, e0129229. [[CrossRef](#)] [[PubMed](#)]
37. Cesareni, G.; Panni, S.; Nardelli, G.; Castagnoli, L. Can we infer peptide recognition specificity mediated by SH3 domains? *FEBS Lett.* **2002**, *513*, 38–44. [[CrossRef](#)]
38. Chook, Y.M.; Gish, G.D.; Kay, C.M.; Pai, E.F.; Pawson, T. The Grb2-mSos1 complex binds phosphopeptides with higher affinity than Grb2. *J. Biol. Chem.* **1996**, *271*, 30472–30478. [[CrossRef](#)]
39. Herrero-Garcia, E.; O’Byrne, J.P. Intersectin scaffold proteins and their role in cell signaling and endocytosis. *Biochim. Biophys. Acta (BBA)-Mol. Cell Res.* **2017**, *1864*, 23–30. [[CrossRef](#)]
40. Alfaidi, M.; Scott, M.L.; Orr, A.W. Sinner or Saint?: Nck adaptor proteins in vascular biology. *Front. Cell Dev. Biol.* **2021**, *9*, 688388. [[CrossRef](#)]
41. Yu, X.; Liang, C.; Zhang, Y.; Zhang, W.; Chen, H. Inhibitory short peptides targeting EPS8/ABI1/SOS1 tri-complex suppress invasion and metastasis of ovarian cancer cells. *BMC Cancer* **2019**, *19*, 878. [[CrossRef](#)]
42. Hahn, S.; Kim, D. Transient protein-protein interaction of the SH3-peptide complex via closely located multiple binding sites. *PLoS ONE* **2012**, *7*, e32804. [[CrossRef](#)]
43. Mayer, B.J.; Saksela, K. SH3 domains. In *Modular Protein Domains*; Wiley: Hoboken, NJ, USA, 2005; pp. 37–58.
44. Dionne, U.; Percival, L.J.; Chartier, F.J.; Landry, C.R.; Bisson, N. SRC homology 3 domains: Multifaceted binding modules. *Trends Biochem. Sci.* **2022**, *47*, 772–784. [[CrossRef](#)] [[PubMed](#)]
45. Dionne, U.; Bourgault, É.; Dubé, A.K.; Bradley, D.; Chartier, F.J.M.; Dandage, R.; Dibyachintan, S.; Després, P.C.; Gish, G.D.; Pham, N.T.H.; et al. Protein context shapes the specificity of SH3 domain-mediated interactions in vivo. *Nat. Commun.* **2021**, *12*, 1597. [[CrossRef](#)] [[PubMed](#)]
46. Zarrinpar, A.; Park, S.-H.; Lim, W.A. Optimization of specificity in a cellular protein interaction network by negative selection. *Nature* **2003**, *426*, 676–680. [[CrossRef](#)]
47. Humphries, A.C.; Donnelly, S.K.; Way, M. Cdc42 and the Rho GEF intersectin-1 collaborate with Nck to promote N-WASP-dependent actin polymerisation. *J. Cell Sci.* **2014**, *127 Pt 3*, 673–685. [[CrossRef](#)] [[PubMed](#)]
48. Shutes, A.; Berzat, A.C.; Cox, A.D.; Der, C.J. Atypical mechanism of regulation of the Wrch-1 Rho family small GTPase. *Curr. Biol.* **2004**, *14*, 2052–2056. [[CrossRef](#)] [[PubMed](#)]
49. Lettau, M.; Pieper, J.; Gerneth, A.; Lengel-Janssen, B.; Voss, M.; Linkermann, A.; Schmidt, H.; Gelhaus, C.; Leippe, M.; Kabelitz, D.; et al. The adapter protein Nck: Role of individual SH3 and SH2 binding modules for protein interactions in T lymphocytes. *Protein Sci.* **2010**, *19*, 658–669. [[CrossRef](#)] [[PubMed](#)]
50. Ahmad, M.; Gu, W.; Helms, V. Mechanism of fast peptide recognition by SH3 domains. *Angew. Chem. Int. Ed.* **2008**, *47*, 7626–7630. [[CrossRef](#)]
51. Yu, H.; Chen, J.K.; Feng, S.; Dalgarno, D.C.; Brauer, A.W.; Schrelber, S.L. Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* **1994**, *76*, 933–945. [[CrossRef](#)]
52. McDonald, C.B.; Seldeen, K.L.; Deegan, B.J.; Farooq, A. SH3 domains of Grb2 adaptor bind to PXP motifs within the Sos1 nucleotide exchange factor in a discriminate manner. *Biochemistry* **2009**, *48*, 4074–4085. [[CrossRef](#)]
53. Jaiswal, M.; Dvorsky, R.; Amin, E.; Risse, S.L.; Fansa, E.K.; Zhang, S.-C.; Taha, M.S.; Gauhar, A.R.; Nakhaei-Rad, S.; Kordes, C. Functional cross-talk between ras and rho pathways: A Ras-specific GTPase-activating protein (p120RasGAP) competitively inhibits the RhoGAP activity of deleted in liver cancer (DLC) tumor suppressor by masking the catalytic arginine finger. *J. Biol. Chem.* **2014**, *289*, 6839–6849. [[CrossRef](#)]
54. Fan, P.-D.; Goff, S.P. Abl interactor 1 binds to sos and inhibits epidermal growth factor-and v-Abl-induced activation of extracellular signal-regulated kinases. *Mol. Cell. Biol.* **2000**, *20*, 7591–7601. [[CrossRef](#)] [[PubMed](#)]
55. Tong, X.K.; Hussain, N.K.; de Heuvel, E.; Kurakin, A.; Abi-Jaoude, E.; Quinn, C.C.; Olson, M.F.; Marais, R.; Baranes, D.; Kay, B.K.; et al. The endocytic protein intersectin is a major binding partner for the Ras exchange factor mSos1 in rat brain. *EMBO J.* **2000**, *19*, 1263–1271. [[CrossRef](#)] [[PubMed](#)]
56. Qian, X.; Esteban, L.; Vass, W.C.; Upadhyaya, C.; Papageorge, A.G.; Yienger, K.; Ward, J.M.; Lowy, D.R.; Santos, E. The Sos1 and Sos2 Ras-specific exchange factors: Differences in placental expression and signaling properties. *Embo J.* **2000**, *19*, 642–654. [[CrossRef](#)] [[PubMed](#)]
57. Hu, Q.; Milfay, D.; Williams, L.T. Binding of NCK to SOS and activation of ras-dependent gene expression. *Mol. Cell Biol.* **1995**, *15*, 1169–1174. [[CrossRef](#)] [[PubMed](#)]
58. Chardin, P.; Camonis, J.H.; Gale, N.W.; Van Aelst, L.; Schlessinger, J.; Wigler, M.H.; Bar-Sagi, D. Human Sos1: A guanine nucleotide exchange factor for Ras that binds to GRB2. *Science* **1993**, *260*, 1338–1343. [[CrossRef](#)] [[PubMed](#)]
59. Birge, R.B.; Kalodimos, C.; Inagaki, F.; Tanaka, S. Crk and CrkL adaptor proteins: Networks for physiological and pathological signaling. *Cell Commun. Signal.* **2009**, *7*, 13. [[CrossRef](#)] [[PubMed](#)]
60. Rufer, A.C.; Rumpf, J.; von Holleben, M.; Beer, S.; Rittinger, K.; Groemping, Y. Isoform-selective interaction of the adaptor protein Tks5/FISH with Sos1 and dynamin. *J. Mol. Biol.* **2009**, *390*, 939–950. [[CrossRef](#)]

61. Sini, P.; Cannas, A.; Koleske, A.J.; Di Fiore, P.P.; Scita, G. Abl-dependent tyrosine phosphorylation of Sos-1 mediates growth-factor-induced Rac activation. *Nat. Cell Biol.* **2004**, *6*, 268–274. [[CrossRef](#)]
62. Wilkinson, B.; Li, J.; Coba, M.P. Synaptic GAP and GEF complexes cluster proteins essential for GTP signaling. *Sci. Rep.* **2017**, *7*, 5272. [[CrossRef](#)]
63. Chau, J.E.; Vish, K.J.; Boggon, T.J.; Stiegler, A.L. SH3 domain regulation of RhoGAP activity: Crosstalk between p120RasGAP and DLC1 RhoGAP. *Nat. Commun.* **2022**, *13*, 4788. [[CrossRef](#)]
64. Berry, D.M.; Nash, P.; Liu, S.K.; Pawson, T.; McGlade, C.J. A high-affinity Arg-X-X-Lys SH3 binding motif confers specificity for the interaction between Gads and SLP-76 in T cell signaling. *Curr. Biol.* **2002**, *12*, 1336–1341. [[CrossRef](#)] [[PubMed](#)]
65. Kazeminejad, N.S.; Ahmadian, M.R. Allosteric regulation of GRB2 modulates RAS activation. *Small GTPases* **2022**, *13*, 282–286. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.