2-1-2019

Engineered SH2 domains with tailored specificities and enhanced affinities for phosphoproteome analysis

Gianluca Veggiani
University of Toronto

Haiming Huang
University of Toronto

Bradley P. Yates
University of Toronto

Jiefei Tong
Hospital for Sick Children University of Toronto

Tomonori Kaneko
Schulich School of Medicine & Dentistry

See next page for additional authors

THE PROTEIN SOCIETY

Engineered SH2 domains with tailored specificities and enhanced affinities for phosphoproteome analysis

Gianluca Veggiani,[1,2] Haiming Huang,[1,2] Bradley P. Yates,[1,2] Jiefei Tong,[3] Tomonori Kaneko,[4] Rakesh Joshi,[4] Shawn S.C. Li,[4] Michael F. Moran,[3,5,6] Gerald Gish,[7] and Sachdev S. Sidhu[1,2,5]*

[1]Department of Biochemistry, University of Toronto, Toronto, Ontario, M5S 1A8, Canada
[2]Donnelly Centre for Cellular and Biomolecular Research, Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario, M5S3E1, Canada
[3]Program in Molecular Structure and Function, The Hospital for Sick Children, Department of Molecular Genetics, University of Toronto, Toronto, Ontario, M5G 0A4, Canada
[4]Department of Biochemistry, Siebens-Drake Medical Research Institute, Schulich School of Medicine and Dentistry, University of Western Ontario, London, Ontario, N6A 5C1, Canada
[5]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, M5S 3E1, Canada
[6]The Hospital for Sick Children, SPARC Biocentre, Toronto, Ontario, M5G 0A4, Canada
[7]Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, Ontario, M5G 1X5, Canada

Received 7 October 2018; Accepted 12 November 2018
DOI: 10.1002/pro.3551
Published online 00 Month 2018 proteinscience.org

Abstract: Protein phosphorylation is the most abundant post-translational modification in cells. Src homology 2 (SH2) domains specifically recognize phosphorylated tyrosine (pTyr) residues to mediate signaling cascades. A conserved pocket in the SH2 domain binds the pTyr side chain and the EF and BG loops determine binding specificity. By using large phage-displayed libraries, we engineered the EF and BG loops of the Fyn SH2 domain to alter specificity. Engineered SH2 variants exhibited distinct specificity profiles and were able to bind pTyr sites on the epidermal growth factor receptor, which were not recognized by the wild-type Fyn SH2 domain. Furthermore, mass spectrometry showed that SH2 variants with additional mutations in the pTyr-binding pocket that enhanced affinity were highly effective for enrichment of diverse pTyr peptides within the human proteome. These results showed that engineering of the EF and BG loops could be used to tailor SH2 domain specificity, and SH2 variants with diverse specificities and high affinities for pTyr residues enabled more comprehensive analysis of the human phosphoproteome.

Statement: Src Homology 2 (SH2) domains are modular domains that recognize phosphorylated tyrosine embedded in proteins, transducing these post-translational modifications into cellular responses. Here we used phage display to engineer hundreds of SH2 domain variants with altered binding specificities and enhanced affinities, which enabled efficient and differential enrichment of the human phosphoproteome for analysis by mass spectrometry. These engineered SH2 domain variants will be useful tools for elucidating the molecular determinants governing SH2 domains binding specificity and for enhancing analysis and understanding of the human phosphoproteome.

Keywords: SH2 domain; phage display; protein engineering; phosphoproteomics

Abbreviations: ELISA, enzyme-linked immunosorbent assay; GST, glutathione S-transferase; MS, mass spectrometry; pTyr, phospho-tyrosine; SH2, Src Homology 2 domain.

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Canadian Institutes of Health Research MOP-93684.

**Correspondence to*: Sachdev S. Sidhu, Department of Biochemistry, University of Toronto, Toronto, ON, Canada M5S 1A8. E-mail: sachdev.sidhu@utoronto.ca

Introduction

Cell signaling relies on highly coordinated and regulated networks of protein–protein interactions to efficiently respond to environmental stimuli.[1] Phosphorylation, the addition of a phosphate group to a protein, is the most frequent post-translational modification,[2] and it acts as a molecular switch to regulate dynamic protein–protein interactions.

To identify phosphorylated targets and assess phosphorylation states, cells rely on specialized modular domains that bind to specific phosphorylated sequences in proteins.[3,4] In metazoans, members of the large Src Homology 2 (SH2) domain structural family control cellular signaling cascades by binding with moderate affinity to specific phosphorylated tyrosine (pTyr) residues in proteins.[5] The expansion of the SH2 domain repertoire appears to have enabled an increased sophistication in pTyr-mediated signaling,[5] and in turn, this may have facilitated the transition from unicellular to multicellular organisms.[6] For example, the unicellular yeast *Saccharomyces cerevisiae* contains only a single SH2 domain,[7] whereas 112 human proteins contain 122 SH2 domains, which regulate numerous signaling pathways that are essential for normal cell function and have been implicated in many diseases.[8,9]

The SH2 fold is comprised of a β-sheet flanked by two alpha helices [Fig. 1(a)]. The recognition of a pTyr peptide ligand by an SH2 domain can be described as being two-pronged, wherein a conserved Arg residue at the base of the pTyr-binding pocket coordinates the pTyr side chain, and a cleft on the surface of the domain interacts with other ligand residues to confer specificity [Fig. 1(b)].[7,10,11] Interactions with the pTyr side chain contribute roughly half of the total free energy of the SH2–ligand interaction,[12] and interactions are typically of moderate affinities in the low micromolar range.[13] Aside from the conserved interactions with the pTyr side chain, additional interactions between the SH2 domain are mediated mainly by the variable EF and BG loops, which flank the hydrophobic cleft and define specificity for residues C-terminal to the pTyr.[11,14]

Many structural studies and screens with phosphopeptide libraries have investigated the binding specificities of SH2 domains,[15,16] which have been shown to recognize three major types of ligands[17–20] and have been grouped into three corresponding specificity classes. Class 1 domains are defined by specificity for the consensus pYξξΦ (where ξ and Φ denote hydrophilic and hydrophobic residues, respectively), and are further divided into four subgroups, with Class 1c domains recognizing an asparagine at the second position following pTyr, (P + 2, pY-x-N). In Class 1c domains, exemplified by the Grb2 SH2 domain, a bulky residue at the first position of the EF loop (Trp^{121}) blocks access of the ligand to the binding pocket, forcing it to adopt a Type I β-turn that enables contact with an Asn residue at P + 2 [Fig. 1(c)].[14,21] Class 2 domains preferably recognize a proline or aliphatic residues (Ψ) at the third position following pTyr (P + 3, pY-x-x-P/Ψ). For example, in the the Fyn SH2 domain (Fyn-SH2), the EF loop adopts an open conformation enabling access of a hydrophobic P + 3 residue to the hydrophobic cleft, while Leu^{239} in the BG loop blocks the P + 4 binding pocket [Fig. 1(d)].[14,22] Class 3 domains recognize a hydrophobic residue at the fourth position following pTyr (P + 4, pY-x-x-x-Φ).[17] The BRDG1 SH2 domain exemplifies a Class 3 domain, in which Leu^{240} blocks the P + 3 binding pocket and an open P + 4 pocket accommodates a Leu side chain from the ligand [Fig. 1(e)].[14]

Since the EF and BG loops work together to dictate SH2 domain specificity by controlling access to binding pockets, we employed a phage display approach that diversified these loops to develop SH2 domain variants with altered specificities. We generated libraries of Fyn-SH2 variants with diverse EF and BG loops and screened for binding to a panel of pTyr peptides representing diverse specificities. We profiled the binding specificities of selected Fyn-SH2 variants and identified altered binding specificities that enabled recognition of ligand classes that were not recognized by wild-type Fyn-SH2 (Fyn-SH2.wt). When used for the analysis of the human phosphoproteome by mass spectrometry, Fyn-SH2 variants with altered specificities enabled the isolation of distinct phosphorylated proteins, confirming the robustness of loop engineering for reshaping SH2 domain specificity.

Results

Fyn-SH2 variants with altered binding specificities

The EF and BG loops work together to dictate pTyr ligand selectivity of SH2 domains.[14] Thus, to modulate specificity, we created two phage-displayed libraries of Fyn-SH2 variants in which positions within the EF and BG loops were diversified with degenerate codons encoding for all 20 genetically encoded amino acids [Fig. 1(a)]. In both libraries, three codons within the EF loop were replaced by three degenerate codons. The two libraries differed in that three codons in the BG loop were replaced by either three degenerate codons (Library 1) or by zero to five degenerate codons (Library 2). Length variation was introduced in the BG loop of Library 2 based on the observation that longer BG loops in some SH2 domains contribute to extended binding surfaces that improve interactions with pTyr ligands.[23,24]

To assemble a panel of SH2 variants with diverse specificities, the two libraries were cycled through rounds of binding selections with 19 biotinylated pTyr peptides immobilized on streptavidin-coated

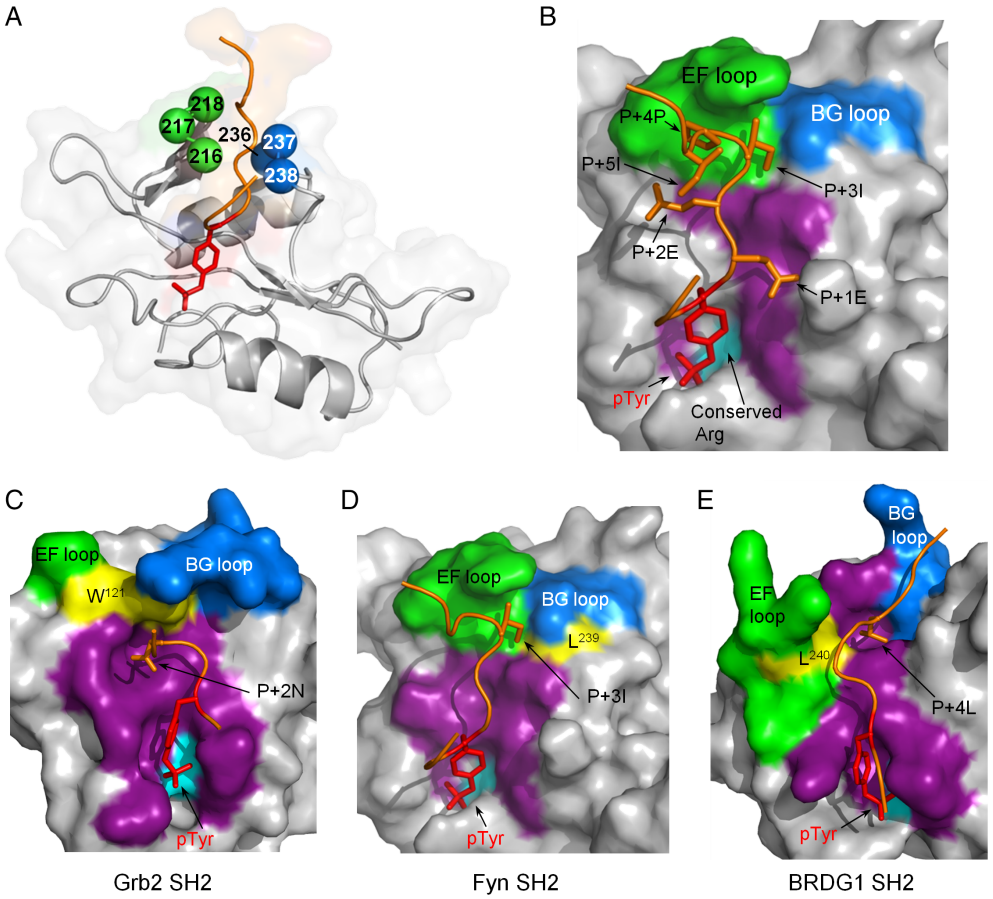

Figure 1. Fyn-SH2 library design. (a) The Fyn-SH2 main chain is shown as a gray ribbon, the main chain of the pTyr peptide ligand (EPQpYEEIPIYL) is colored orange (PDB entry 1AOU), and the pTyr side chain is shown as sticks colored red. Residues that were diversified in the library are shown as numbered spheres colored green (EF loop) or blue (BG loop). In Libraries 1 and 2, the codons encoding the three BG loop residues were replaced by three or zero to five degenerate codons, respectively, and in both libraries, the codons encoding the three EF loop residues were replaced by three degenerate codons. (b) Surface representation of Fyn-SH2 in complex with the pTyr peptide ligand (PDB entry 1AOU). The EF and BG loops are colored green or blue, respectively. The conserved Arg^{176} that coordinates the pTyr side chain (red sticks) is colored cyan. Residues forming the hydrophobic cleft interacting with the ligand are colored purple. The peptide ligand backbone is shown as an orange tube and side chains are shown as sticks. (c–e) Surface representations of the SH2 domains of (c) Grb2 (PDB entry 3WA4), (d) Fyn (PDB entry 1AOU) and (e) BRDG1 (PDB entry 3MAZ) in complex with P+ 2N (SDpYMNMTP), P+ 3I (EPQpYEEIPIYL) or P+ 4L (ANSpYENVLIAK) ligands, respectively. SH2 domain surfaces and peptide ligands are shown and colored as in Panel B, and in addition, key residues in the EF and BG loops that dictate specificity are colored yellow.

plates, which represented diverse natural ligands spanning the known specificity classes (Table I). Following five rounds of selections, positive clones were identified by clonal phage ELISAs as those that exhibited strong signals on wells containing immobilized pTyr peptides but not on wells containing streptavidin only. Approximately 12 positive clones from each library selected against each peptide were subjected to DNA sequencing, resulting in a total of 152 unique Fyn-SH2 variants (Fig. 2 and Table S1). In accordance with the library designs, the variants contained diverse sequences only in the EF and BG loops, and significant length diversity was observed in the BG loop region. In many cases, sequence consensus was observed amongst variants from binding selections for the same peptide, and also, amongst variants selected against peptides of the same class, indicating that these variants likely use similar binding mechanisms (Fig. 2).

Affinity and specificity analysis of Fyn-SH2 variants

We characterized in detail Fyn-SH2.wt and six variants that were chosen to represent diverse loop sequences and binding preferences (Table II). We assessed the affinities and specificities of the domains by fluorescence polarization with a series of pTyr peptides covering a broad range of specificity classes (Table III). The panel included three peptides representing a prototypical P+ 2N, P+ 3I or P+ 4L ligand, and four peptides representing pTyr sites in the epidermal growth factor receptor (EGFR). Fyn-SH2.wt bound to all the prototypical peptides with affinities that agreed with values previously reported for

Table I. *pTyr Peptides used for Fyn-SH2 Variant Selections*

Peptides	Sequence											Binding Class	Motif	Source
1	E	P	Q	pY	E	E	I	E	E	–	–	2	P + 3I	TM antigen
2	T	T	E	pY	S	E	I	K	I	H	T	2	P + 3I	SIG11 pY668
3	E	N	L	pY	E	G	L	N	L	D	D	2	P + 3 L	CD79A pY188
4	E	S	I	pY	E	V	L	G	M	Q	Q	2	P + 3 L	CEA20 pY578
5	P	Q	R	pY	L	V	I	Q	G	D	–	2	P + 3I	EGFR pY978
6	A	D	E	pY	L	I	P	Q	Q	G	–	2	P + 3P	EGFR pY1016
7	D	P	H	pY	Q	D	P	H	S	T	–	2	P + 3P	EGFR pY1125
8	D	H	Q	pY	Y	N	D	A	P	G	–	1c	P + 2 N	ShcA pY239
9	V	P	E	pY	I	N	Q	S	V	P	–	1c	P + 2 N	EGFR pY1138
10	Q	P	E	pY	V	N	Q	A	D	V	–	1c	P + 2 N	ErbB2 pY1139
11	E	P	L	pY	L	N	T	F	A	N	–	1c	P + 2 N	ErbB4 pY1208
12	A	E	L	pY	S	N	A	A	P	V	–	1c	P + 2 N	PDGFRβ pY716
13	E	P	Q	pY	E	E	E	L	E	–	–	3	P + 4 L	TM antigen
14	Q	D	T	pY	E	T	H	L	E	T	–	3	P + 4 L	TRAF7 pY275
15	N	P	D	pY	Q	Q	D	F	F	P	–	3	P + 4F	EGFR pY1172
16	R	N	D	pY	D	D	T	I	P	I	–	3	P + 4I	MALT1 pY470
17	S	N	F	pY	R	A	L	M	D	E	–	3/4	P + 3 L, P + 4 M	EGFR pY998
18	N	P	V	pY	H	N	Q	P	L	N	–	1c/3	P + 2 N, P + 4P	EGFR pY1092
19	N	P	E	pY	L	N	T	V	Q	P	–	1c/3	P + 2 N, P + 4 V	EGFR pY1110

The sequences of each pTyr peptide used as bait in phage display experiments are reported. Gray shading highlights residues defining the binding class and specificity motif are indicated. Each peptide was derived from a natural protein source, as indicated. The following abbreviations are used: TM antigen, polyomavirus middle T-antigen; SIG11, sialic acid-binding Ig-like lectin 11; CD97A, cluster of differentiation 97A; CEA20, carcinoembryonic antigen-related cell adhesion Molecule 20; EGFR, epidermal growth factor receptor; ShcA, Src homology and Collagen A; ErbB2, receptor tyrosine-protein kinase erbB-2; ErbB4, receptor tyrosine-protein kinase erbB-4; PDGFRβ, beta-type platelet-derived growth factor receptor; TRAF7, TNF receptor associated Factor 7; MALT1, Mucosa-associated lymphoid tissue lymphoma translocation Protein 1.

similar peptides,[14,25] but it did not bind to any of the EGFR peptides. Each variant bound to the peptide it was selected for and exhibited distinct binding specificity profiles that ranged from the broad specificity of v29, which recognized six of seven peptides, to the highly specific v17, which recognized its cognate peptide only. Taken together, the fluorescence polarization and peptide array assays showed that we succeeded in generating SH2 domain variants with diverse specificities by altering EF and BG loop sequences.

Phosphoproteome enrichment with Fyn-SH2 variants

Because SH2 domains bind to pTyr proteins and peptides, they can be used as affinity reagents for the enrichment of phosphoproteomes for mass spectrometry (MS) analysis. However, these applications have been limited by the modest affinities of natural SH2 domains.[26] We previously developed "superbinder" Fyn and Src SH2 domains with high affinities for a broad spectrum of pTyr peptides by mutating three residues in the pTyr-binding pocket.[27] Recently, these superbinders have been used for ultra-deep MS analysis of phosphoproteomes.[28,29] To advance this methodology, we investigated whether coverage of the human phosphoproteome could be improved further by combining superbinder mutations with our loop variants to develop SH2 domains with diverse specificities and high affinities for pTyr peptides. We engineered superbinder versions of Fyn-SH2.wt (Fyn-SH2.s) and five variants (v5s, v25s, v27s, v28s, v29s), and confirmed high affinity binding to a panel of pTyr peptides by fluorescence polarization assays (Table IV).

We compared Fyn-SH2.wt, Fyn-SH2.s and three of the superbinder variants (v5s, v25s, v28s) for the ability to enrich phosphorylated peptides from orthovanadate-treated HeLa cells for MS analysis [Fig. 3(a)]. Proteins in cell lysates were digested with trypsin, labeled with tandem mass tags, captured with SH2 domains, and subjected to MS analysis. All superbinders greatly enhanced the enrichment of phosphopeptides compared with Fyn-SH2.wt, and Fyn-SH2.s exhibited the broadest pTyr peptide coverage [Fig. 3(b)]. Despite their high sequence similarity (Table II), the analyzed superbinder Fyn-SH2 variants exhibited different profiles for enrichment of pTyr peptides [Fig. 3(b)]. Correlation analysis showed that v5s and v28s isolated phosphopeptides in a very similar manner, whereas v25s and v28s showed the highest degree of difference (Fig. S1). This is consistent with the high sequence similarity of the BG loops of v5s and v28s (Table II), indicating that these loops may control access to the binding pocket in a similar manner.

This was further assessed by comparing the peptide sets that were commonly isolated by different variants. In comparison to v25s, v5s and v28s isolated 18 and 16 pTyr peptides more efficiently, and shared 55% of these [Fig. 3(c), top left], thus confirming similar binding specificities (Fig. S1). When compared to v28s, v25s and v5s showed a 36% overlap,

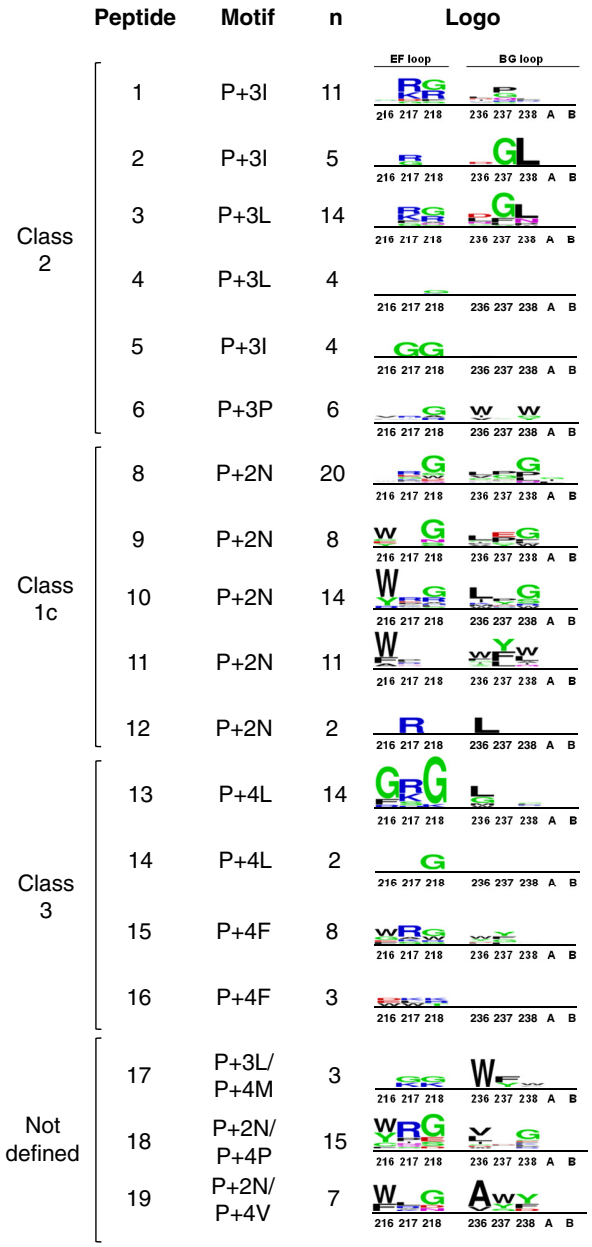

Figure 2. Sequence conservation of Fyn-SH2 variants. A total of 152 unique Fyn-SH2 variants were grouped into five classes according to the nature of the pTyr peptide used for their selection from phage-displayed libraries (Table I). A number of unique variants isolated for each peptide (*n*) were aligned and the alignments were used to derive sequence logos from the EF and BG loop amino acid frequencies using the WebLogo tool.[47] Only one variant was selected for Peptide 7, and thus, a logo could not be determined in this case.

indicating a moderately similar specificity [Fig. 3(c), top middle]. In contrast, none of the pTyr peptides that v25s and v28s enriched more efficiently than v5s were shared [Fig. 3(c), top right], confirming very different binding specificities (Fig. S1).

Similar patterns were observed for peptide sets with a log2-fold enrichment lower than a reference Fyn-SH2 variant [Fig. 3(c), bottom], proving that all three domains have significant specificity differences.

In a second experiment, we first captured intact proteins with an SH2 domain, digested with trypsin, and then performed MS analysis.

For 427 proteins that were log2-fold or greater enriched for Fyn-SH2.s relative to negative control beads, we compared peptide abundance for each of the SH2 domains and an anti-pTyr antibody relative to Fyn-SH2.s [Fig. 3(d)]. This analysis showed that Fyn-SH2.wt and the anti-pTyr antibody poorly captured most of the proteins isolated by Fyn-SH2.s. Moreover, a superbinder version of the Grb2 SH2 domain (Grb2-SH2.s) and the five superbinder Fyn-SH2 variants exhibited significant differences in the capture efficiency for these proteins, amongst themselves and in comparison with Fyn-SH2.s. Many of the proteins were less efficiently captured by Grb2-SH2.s and the superbinder Fyn-SH2 variants than by Fyn-SH2.s, but a subset was more efficiently captured by particular domains [Fig. 3(d)]. Analysis of this subset showed that only four proteins were common to two or more variants (Table S2), indicating that the Fyn-SH2 variants have distinct specificity profiles, and that their combination could improve the pTyr sequence coverage.

Taken together, these results show that superbinder SH2 domains are much more efficient than Fyn-SH2.wt or an anti-pTyr antibody for phosphoproteome enrichment. Moreover, they further highlight the altered specificities of the Fyn-SH2 variants and suggest that combinations of superbinder variants with Fyn-SH2.s could serve as even more efficient phosphoproteome capture reagents than Fyn-SH2.s alone.

Discussion

SH2 domains are key components of cellular signaling pathways,[8,30] and they function by specific recognition of pTyr sites in partner proteins. In the context of complex networks of kinases and phosphatases, correct targeting requires highly specialized and specific interactions. The EF and BG loops of SH2 domains have been shown to be key elements responsible for conferring specificity.[14] Although recognition of pTyr sites is mainly achieved through interactions of these loops with the residues C-terminal to the pTyr moiety,[31] these interactions are not the sole contributors to specificity, and recent studies have revealed a higher degree of complexity.[32]

Substitutions within the SH2 domain core, together with diversification of EF and BG loop positions, enabled the development of an SH2 domain with dual specificity.[33] Moreover, incorporation of non-natural amino acids at a core position demonstrated the importance of the core for dictating SH2 domain specificity.[34] Single substitutions in the EF and BG loops altered the specificities of Src,[35] BRDG1 and Fyn SH2 domains,[36] but the loops act cooperatively and multiple mutations may greatly

Table II. *Fyn-SH2 Variants Subjected to Affinity Assays*

	EF loop			BG loop							
Variants	216	217	218	236	237	238	a	b	Selection peptide	Motif	Source
Fyn wt	T	T	R	A	G	L	–	–			
v5	W	L	G	V	P	G	S	–	9	P + 2 N	EGFR pY1138
v17	G	R	G	–	–	–	–	–	13	P + 4 L	TM antigen
v25	P	G	G	W	Y	W	–	–	5	P + 3I	EGFR pY978
v27	V	R	G	W	Y	W	–	–	6	P + 3P	EGFR pY1016
v28	R	R	R	L	P	G	–	–	7	P + 3P	EGFR pY1125
v29	W	R	G	W	Y	W	–	–	15	P + 4F	EGFR pY1172

Sequences are shown for the EF and BG loop regions that were diversified in the libraries for variants that were subjected to affinity analysis (Tables III and IV). Insertions in the BG loop relative to Fyn-SH2.wt are labeled with letters, whereas dashes indicate gaps in the alignment. The binding specificity motif and source of pTyr peptides used to isolate a specific variant in phage-display experiments are reported as in Table I.

expand the range of specificities that can be supported by the SH2 fold. Thus, we performed combinatorial diversification of the EF and BG loops to more extensively explore the potential for SH2 domain specificity engineering.

Our approach was successful in generating numerous Fyn-SH2 variants from selections for binding to 19 distinct ligands representing various specificity classes. Assays with fluorescence polarization of synthetic peptide ligands showed that many variants exhibited altered specificities and this was further confirmed by mass spectrometry. However, the observed specificities did not always match the specificities expected from the peptides used for selection, and many domains exhibited broad specificities. Similar binding profiles were previously encountered in studies of PDZ and SH3 domain variants derived by phage display, where hundreds of variants were selected for binding to diverse targets but the specificities were typically broad.[37–40] These issues may be addressed by further optimization of library design and selection strategies. In particular, including additional positions beyond the EF and BG loops in the diversification strategy may yield domains that can establish more precise contacts with ligands and thus confer greater specificity. Moreover, selections can be made more stringent by adding competitor peptides to the phage pool to remove variants with broad or unwanted specificities, and by increasing the stringency of the washing procedures to remove less tightly bound variants.

Nevertheless, even with non-optimal library design and selection strategies, we were able to generate Fyn-SH2 variants with diverse specificities, as demonstrated by the acquired ability of a number of Fyn-SH2 variants in binding pTyr sites on EGFR. Therefore, the modulation of the EF and BG could be used as a general strategy to develop SH2 variants with tailored specificity to rewire cell signaling pathways for synthetic biology. For example, Fyn-SH2 variants able to bind clinically relevant pTyr sites on EGFR have the potential to compete with dysregulated EGFR protein partners, thus disrupting aberrant signaling pathways.

We converted Fyn-SH2 variants into superbinders with extremely high affinities for pTyr peptides simply by transferring three additional substitutions from a previously engineered superbinder SH2 domain.[27] As previously reported Fyn and other SH2 domains have the propensity to form dimers at elevated concentrations.[41] However, given the low concentration range used in our binding assays, the high binding affinity observed is likely dependent on the formation of an improved pTyr-binding pocket.

We were able to use the superbinder Fyn-SH2 variants to enrich for diverse sequences within the human phosphoproteome with much greater efficiency than with Fyn-SH2.wt or an anti-pTyr antibody. Furthermore, different Fyn-SH2 variants exhibited distinct patterns of enriched sequences, suggesting that combinations of superbinder SH2 variants with distinct binding selectivity are likely to be even more effective than single superbinders for broad enrichment of the phosphoproteome for MS analysis. Thus, by applying improved strategies for specificity engineering with established means for making superbinders, it should be possible to develop an optimized toolkit of SH2 superbinder variants tailored for phosphoproteome research.

Methods

Library construction and selection of Fyn-SH2 variants

For the construction of phage-displayed libraries, combinatorial site-directed mutagenesis of a phagemid designed for the phage display of Fyn-SH2 was performed.

Positions in the EF and BG loop sequences were simultaneously mutated with a "hard randomization" strategy, as described.[42] Libraries were constructed by oligonucleotide-directed mutagenesis,[43] using a set of mutagenic oligonucleotides containing degenerate NNK (N = A/G/C/T, K = G/T) codons at positions to be diversified.

Table III. *Affinities of Fyn-SH2 Variants for Synthetic Peptides*

Selection peptide	Sequence										Source	Motif	K_D (nM)				
													Fyn-SH2.wt	v17	v25	v28	v29
–	E	P	Q	pY	E	N	E	E	E	–	Synthetic	P + 2 N	350 ± 20	–	1900 ± 700	140 ± 20	120 ± 40
1	E	P	Q	pY	E	E	I	E	E	–	TM antigen	P + 3I	90 ± 30	–	570 ± 50	165 ± 20	500 ± 300
13	E	P	Q	pY	E	E	E	L	E	–	TM antigen	P + 4 L	1300 ± 500	1040 ± 100	1070 ± 100	470 ± 200	500 ± 100
5	P	Q	R	pY	L	V	I	Q	G	D	EGFR pY 978	P + 3I	–	–	1100 ± 100	–	–
6	A	D	E	pY	L	I	P	Q	Q	G	EGFR pY 1016	P + 3P	–	–	–	–	1000 ± 200
7	D	P	H	pY	Q	D	P	H	S	T	EGFR pY 1125	P + 3P	–	–	–	960 ± 300	3500 ± 800
15	N	P	D	pY	Q	Q	D	F	F	P	EGFR pY 1172	P + 4F	–	–	–	1300 ± 200	3300 ± 300

Affinities were determined by fluorescence polarization. Gray shading indicates the pTyr peptide that was used in binding selections from which each variant was obtained. Peptides used in binding selections are indicated as in Table I. Dashes indicate no detectable binding. Sequences of the variants are shown in Table II.

Table IV. *Affinities of Superbinder Fyn-SH2 Variants for Synthetic Peptides*

Selection peptide	Sequence										Source	Motif	K_D (nM)					
													Fyn-SH2.s	v5s	v25s	v27s	v28s	v29s
–	E	P	Q	pY	E	N	E	E	E	–	Synthetic	P + 2 N	40 ± 10	50 ± 10	60 ± 10	30 ± 10	40 ± 10	30 ± 5
1	E	P	Q	pY	E	E	I	E	E	–	TM antigen	P + 3I	20 ± 5	40 ± 10	30 ± 10	30 ± 5	30 ± 10	30 ± 10
13	E	P	Q	pY	E	E	E	L	E	–	TM antigen	P + 4 L	50 ± 10	80 ± 20	50 ± 10	50 ± 10	60 ± 10	70 ± 20
5	P	Q	R	pY	L	V	I	Q	G	D	EGFR pY 978	P + 3I	30 ± 10	70 ± 10	20 ± 5	50 ± 5	70 ± 10	70 ± 10
6	A	D	E	pY	L	I	P	Q	Q	G	EGFR pY 1016	P + 3P	80 ± 15	235 ± 80	400 ± 100	85 ± 10	200 ± 60	170 ± 30
7	D	P	H	pY	Q	D	P	H	S	T	EGFR pY 1125	P + 3P	45 ± 10	70 ± 20	70 ± 20	30 ± 5	50 ± 10	80 ± 20
15	N	P	D	pY	Q	Q	D	F	F	P	EGFR pY 1172	P + 4F	50 ± 10	60 ± 10	55 ± 10	40 ± 5	50 ± 10	50 ± 10

Affinities were determined by fluorescence polarization. Gray shading indicates the pTyr peptide that was used in binding selections from which each variant was obtained (selection peptide, reported as in Table I). Dashes indicate no detectable binding. Sequences of the variants are shown in Table II.

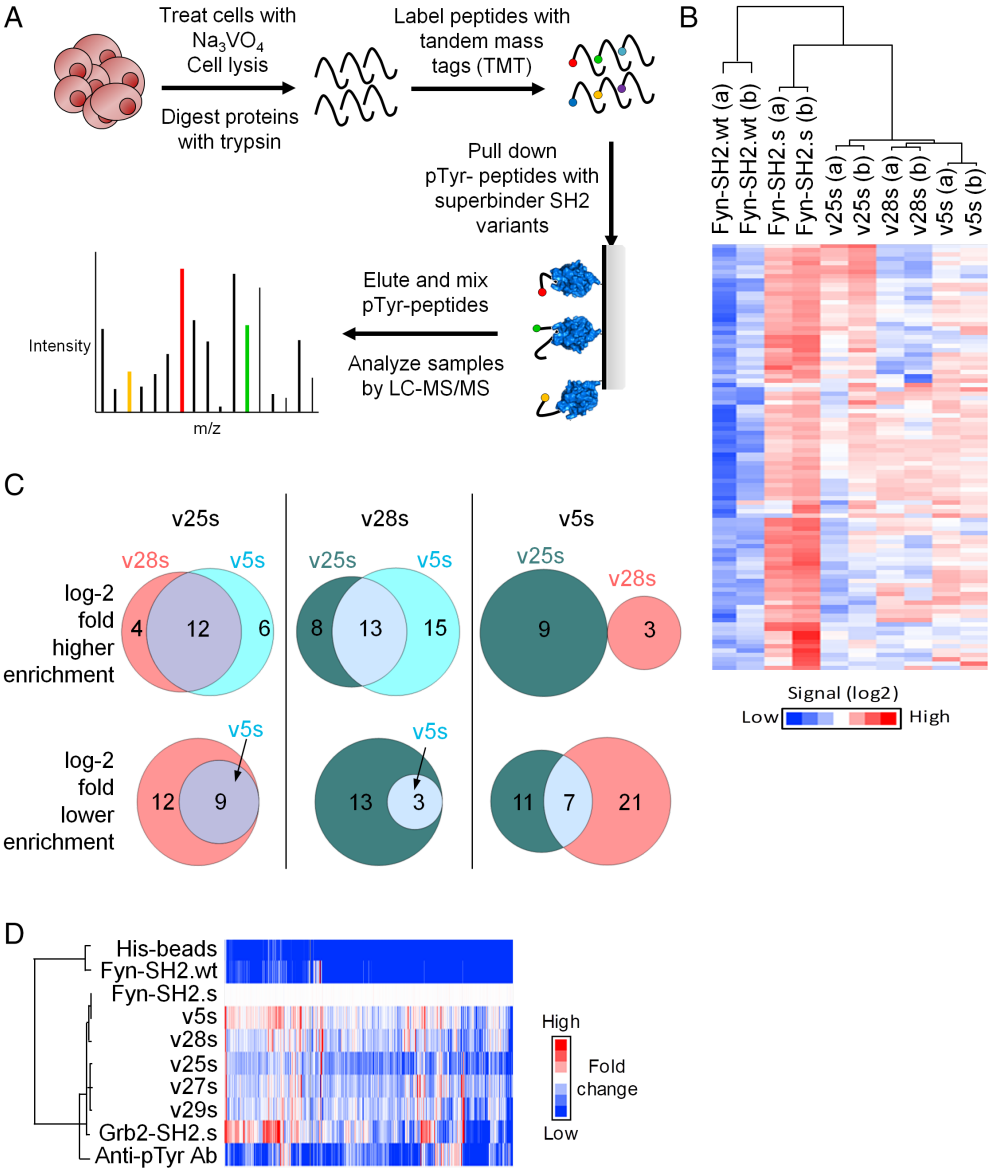

Figure 3. Enrichment of the human phosphoproteome with superbinder Fyn-SH2 variants. (a) Workflow of pTyr peptide enrichment by superbinder Fyn-SH2 variants. Proteins from orthovanadate-treated HeLa cells were digested with trypsin, and peptides were labeled with 10-multiplex tandem mass tags (TMTs). TMT-labeled pTyr peptides were enriched by an immobilized superbinder Fyn-SH2 variant, eluted, pooled and analyzed by LC–MS/MS. (b) Heat map depicting un-supervised clustering of 99 pTyr peptides enriched by Fyn-SH2.wt or superbinder Fyn-SH2 variants. Samples were analyzed in duplicate in two independent experiments (a,b). (c) Venn diagrams comparing the pTyr peptide enrichment profiles of superbinder Fyn-SH2 variants v25s, v28s, and v5s. The top and bottom diagrams show pTyr peptide sets with log2-fold higher or lower enrichment, respectively, than the reference variant that is indicated above each pair of diagrams. (d) Un-supervised clustering of phosphoproteins pulled down by an anti-pTyr antibody, Fyn-SH2.wt or superbinder SH2 variants, prior to trypsin digestion and labeling with TMT. As a control for non-specific binding, TMT-labeled peptides were loaded onto the resin used to immobilize the SH2 variants (His-beads). The heat map displays a cluster of 427 proteins, which showed a log2-fold enrichment two times greater for Fyn-SH2.s compared with the His-beads control. The heat scale depicts changes in fold enrichment relative to Fyn-SH2.s.

For selection of Fyn-SH2 variants, each biotinylated pTyr peptide was immobilized in 96-well Nunc-Immuno MAXISORP plates (Thermo Scientific) coated with streptavidin (New England Biolabs), and phage pools representing the libraries were cycled through five rounds of binding selections with the immobilized pTyr peptide, as described.[44] Phage ELISAs were performed to identify positive clones able to bind to the biotinylated pTyr peptide but not to streptavidin, and the amino acid sequences of positive Fyn-SH2 variants were decoded by DNA sequencing.

Purification of Fyn-SH2 variant proteins

Plasmids encoding SH2 domains fused to the C-terminus of glutathione S-transferase (GST) were transformed into *Escherichia coli* BL21(DE3), and

single colonies were used to inoculate 5 mL 2YT medium containing 0.1 mg/mL carbenicillin. Cultures were grown overnight at 37°C with 200 rpm shaking, diluted 1:200 in 2YT medium containing 0.1 mg/mL carbenicillin, grown at 37°C with 200 rpm shaking to OD_{600} 0.6–0.8, and induced with 0.1 m*M* IPTG at 18°C with 200 rpm shaking for 18 h. Cultures were pelleted and resuspended in 10 mL Lysis Buffer (50 m*M* Tris–HCl pH 7.8, 300 m*M* NaCl, 1 m*M* Phenylmethylsulfonyl fluoride (PMSF), and protease inhibitor cocktail (Sigma-Aldrich)). Cells were lysed by sonication, and protein purification was performed by standard methods with Ni-NTA resin (Qiagen), eluting proteins with an imidazole buffer gradient ranging from 30 to 300 m*M*. The purity of eluted fractions was assessed by SDS-PAGE, and the buffer was exchanged by dialysis at 4°C into PBS pH 7.4. Protein concentrations were determined from OD_{280} measurements with extinction coefficient from ExPASy ProtParam.[45]

Peptide synthesis

Solid phase peptide synthesis was performed using 9-fluorenylmethoxycarbonyl chemistry on Rink amide MBHA resin (Novabiochem) on a Prelude peptide synthesizer (Protein Technologies, Inc.). Each peptide N-terminus was functionalized directly with 5-(and-6)-carboxyfluorescein (ThermoFisher Scientific) or with biotin through a linker composed of two ε-aminocaproic acids (Bachem). All peptides were purified using C-18 reverse phase HPLC (Waters) and authenticity was confirmed by mass spectrometry on an Orbitrap Elite (ThermoFisher). Unless stated otherwise, peptide sequences were derived from human proteins.

Fluorescence polarization binding experiments

Binding measurements were performed in a 96-well plate in FP buffer (20 m*M* HEPES pH 7.5, 150 m*M* NaCl, 1 m*M* EDTA, 0.01 mg/mL BSA, 0.3% BRIJ-35) by mixing 20 nM FITC-labeled pTyr peptide with serial dilutions of Fyn-SH2 variant protein ranging from 0.5 to 9 μM. For Fyn-SH2.wt or superbinder Fyn-SH2 variants, the concentrations ranged from 0.1 to 1.8 μM or 0.023 to 2 μM, respectively. Samples were equilibrated at room temperature for 30 min before reading plates on an Analyst HT Plate Reader (Molecular Devices) using an excitation filter of 485 nm and an emission filter of 530 nm. Dissociation constants were determined with Prism (GraphPad Software Inc) using a one-site total binding model.

Cell culture

HeLa cells were from American Type Culture Collection and were grown at 37°C with 5% CO_2 in Dulbecco's Modified Eagle Medium (DMEM) (Life Technologies) containing 10% (vol/vol) FBS (Sigma-Aldrich) and 50 U/mL penicillin and 50 μg/mL streptomycin (Sigma-Aldrich).

Phosphoproteome enrichment with SH2 variants

HeLa cells were treated with 1 m*M* sodium orthovanadate (Sigma), incubated at 37°C with 5% CO_2 for 20 min, and washed twice with ice cold PBS. Upon addition of Lysis buffer (20 m*M* HEPES pH 8.0, 8 M urea, and phosphatase inhibitors (ThermoFisher Scientific), HeLa cells were scraped from plates, gently sonicated, and spun at 15,000*g* for 15 min at 4°C to remove cell debris. Supernatant was collected and reduced by addition of 5 m*M* Tris(2-Carboxyethyl) phosphine hydrochloride (TCEP) (Sigma) for 60 min at 55°C prior to alkylation with 10 m*M* 2-Chloroacetamide (Sigma) for 15 min at room temperature in the dark. For enrichment of pTyr peptides following alkylation, protein extracts were digested overnight at room temperature in 50 m*M* ammonium bicarbonate pH 8.5 by adding 25 μg TPCK-trypsin (Pierce) per 100 μg protein.

Digested peptides were quantified by BCA assay (Pierce), and buffer was evaporated by centrifugation at 25°C for 20 min using a Savant™ SPD131DDA SpeedVac™ Concentrator (ThermoFisher Scientific). Samples were resuspended in 100 m*M* Triethylammonium bicarbonate buffer (TEAB) pH 8.5, split into 10 vials and labeled for 1 h at room temperature with amine-reactive 10-multiplex tandem mass tags (TMT) (ThermoFisher Scientific) according to the manufacturer's instructions. The reaction was quenched by addition of 5% hydroxylamine (Sigma) for 15 min at room temperature. TMT-labeled peptides were pooled together in one vial and stored at −80°C.

To isolate TMT-labeled pTyr peptides, 50 μg His-tagged Fyn-SH2 variant protein was immobilized onto 25 μL Ni-NTA agarose resin (Qiagen) previously equilibrated with 1 mL binding buffer (50 m*M* Na_2HPO_4 pH 7.4, 50 m*M* NaCl). 100 μg TMT-labeled peptides were applied to each bead-immobilized Fyn-SH2 variant and incubated with end-over-end rotation for 4 h. Beads were collected by centrifugation at 2000*g* at 4°C for 5 min and washed three times with 1 mL binding buffer. Bound TMT-labeled pTyr peptides were eluted from beads by addition of 50 μL binding buffer containing 50 m*M* phenylphosphate (Sigma) (Elution buffer). Samples were incubated in elution buffer for 10 min at room temperature, and eluted pTyr peptides were collected following precipitation of Ni-NTA beads by centrifugation at 2000*g* for 5 min at 4°C. pTyr peptides were transferred to a new tube and adjusted to 1% TFA (Sigma).

To assess enrichment of phosphorylated proteins by Fyn-SH2 variants, cell extracts were prepared as described above. Following protein reduction and alkylation, samples were subjected to affinity purification by applying 2 mg cell extract to Fyn-SH2

variants immobilized onto Ni-NTA agarose resin as previously described. As a control, cell extracts were loaded onto empty Ni-NTA beads (His beads) and beads previously immobilized with 50 μg anti-pTyr antibody. After incubation for 3 h and elution as described above, samples were digested with TPCK-trypsin and labeled with 10-multiplex TMT labels as described above. TMT-labeled peptides were pooled together and adjusted to 1% TFA. Following the enrichment step, the pTyr peptide mixtures were desalted and concentrated using C18 Ziptips (Sigma) according to manufacturer's instructions and stored at −80°C.

LC–MS/MS analysis

MS analysis was performed on an Orbitrap Elite mass spectrometer (ThermoFisher Scientific) coupled to an EASY nlC 1000 chromatography system (ThermoFisher Scientific). Peptides were separated by liquid chromatography with a 2-μm C18 column (Thermo Scientific) at a flow rate of 250 nL/min for 120 min using a 0–40% acetonitrile gradient. Eluted peptides were injected into the mass spectrometer, and data were acquired at a 70,000 resolution with a *m/z* 400. Mass spectra were acquired in full scan mode with HCD (high-energy collision dissociation) fragmentation. Acquired data were analyzed by MaxQuant software[46] for identification and quantification on Swiss-Prot database.

Acknowledgments

This work was supported by the Canadian Institutes of Health Research grant MOP-93684 to S.S.S. SSCL held a Canada Research Chair in Functional Genomics and Cellular Proteomics.

Conflict of interest

The authors declare no competing financial interests.

References

1. Scott JD, Pawson T (2009) Cell signaling in space and time: where proteins come together and when they're apart. Science 326:1220–1224.
2. Khoury GA, Baliban RC, Floudas CA (2011) Proteome-wide post-translational modification statistics: frequency analysis and curation of the Swiss-prot database. Sci Rep 1:90.
3. Cohen GB, Ren R, Baltimore D (1995) Modular binding domains in signal transduction proteins. Cell 80: 237–248.
4. Seet BT, Dikic I, Zhou M-M, Pawson T (2006) Reading protein modifications with interaction domains. Nat Rev Mol Cell Biol 7:473–483.
5. Liu BA, Jablonowski K, Raina M, Arcé M, Pawson T, Nash PD (2006) The human and mouse complement of SH2 domain proteins—establishing the boundaries of phosphotyrosine signaling. Mol Cell 22:851–868.
6. Shiu S-H, Li W-H (2004) Origins, lineage-specific expansions, and multiple losses of tyrosine kinases in eukaryotes. Mol Biol Evol 21:828–840.
7. Liu BA, Nash PD (2012) Evolution of SH2 domains and phosphotyrosine signalling networks. Phil Trans R Soc Lond B Biol Sci 367:2556–2573.
8. Filippakopoulos P, Müller S, Knapp S (2009) SH2 domains: modulators of nonreceptor tyrosine kinase activity. Curr Opin Struct Biol 19:643–649.
9. Hendriks WJAJ, Elson A, Harroch S, Pulido R, Stoker A, den Hertog J (2013) Protein tyrosine phosphatases in health and disease. FEBS J 280:708–730.
10. Ladbury JE, Arold ST (2011) Energetics of Src homology domain interactions in receptor tyrosine kinase-mediated signaling. Methods Enzymol 488:147–183.
11. Eck MJ, Shoelson SE, Harrison SC (1993) Recognition of a high-affinity phosphotyrosyl peptide by the Src homology-2 domain of p56lck. Nature 362:87–91.
12. Bradshaw JM, Waksman G (1999) Calorimetric examination of high-affinity Src SH2 domain-tyrosyl phosphopeptide binding: dissection of the phosphopeptide sequence specificity and coupling energetics. Biochemistry 38:5147–5154.
13. Ladbury JE, Lemmon MA, Zhou M, Green J, Botfield MC, Schlessinger J (1995) Measurement of the binding of tyrosyl phosphopeptides to SH2 domains: a reappraisal. Proc Natl Acad Sci USA 92:3199–3203.
14. Kaneko T, Huang H, Zhao B, Li L, Liu H, Voss CK, Wu C, Schiller MR, Li SS-C (2010) Loops govern SH2 domain specificity by controlling access to binding pockets. Sci Signal 3:ra34–ra34.
15. Bradshaw JM, Mitaxov V, Waksman G (1999) Investigation of phosphotyrosine recognition by the SH2 domain of the Src kinase. J Mol Biol 293:971–985.
16. Machida K, Thompson CM, Dierck K, Jablonowski K, Kärkkäinen S, Liu B, Zhang H, Nash PD, Newman DK, Nollau P, Pawson T, Renkema GH, Saksela K, Schiller MR, Shin D-G, Mayer BJ (2007) High-throughput phosphotyrosine profiling using SH2 domains. Mol Cell 26:899–915.
17. Huang H, Li L, Wu C, Schibli D, Colwill K, Ma S, Li C, Roy P, Ho K, Songyang Z, Pawson T, Gao Y, Ss-C L (2008) Defining the specificity space of the human SRC homology 2 domain. Mol Cell Proteomics 7:768–784.
18. Rodriguez M, Li SS-C, Harper JW, Songyang Z (2004) An oriented peptide array library (OPAL) strategy to study protein–protein interactions. J Biol Chem 279: 8802–8807.
19. Songyang Z, Shoelson SE, Chaudhuri M, Gish G, Pawson T, Haser WG, King F, Roberts T, Ratnofsky S, Lechleider RJ (1993) SH2 domains recognize specific phosphopeptide sequences. Cell 72:767–778.
20. Wavreille A-S, Garaud M, Zhang Y, Pei D (2007) Defining SH2 domain and PTP specificity by screening combinatorial peptide libraries. Methods 42:207–219.
21. Rahuel J, Gay B, Erdmann D, Strauss A, Garcia-Echeverría C, Furet P, Caravatti G, Fretz H, Schoepfer J, Grütter MG (1996) Structural basis for specificity of Grb2-SH2 revealed by a novel ligand binding mode. Nat Struct Biol 3:586–589.
22. Mulhern TD, Shaw GL, Morton CJ, Day AJ, Campbell ID (1997) The SH2 domain from the tyrosine kinase Fyn in complex with a phosphotyrosyl peptide reveals insights into domain stability and binding specificity. Structure 5:1313–1323.
23. Nolte RT, Eck MJ, Schlessinger J, Shoelson SE, Harrison SC (1996) Crystal structure of the PI 3-kinase p85 amino-terminal SH2 domain and its phosphopeptide complexes. Nat Struct Biol 3:364–374.
24. Zadjali F, Pike ACW, Vesterlund M, Sun J, Wu C, Li SSC, Rönnstrand L, Knapp S, Bullock AN, Flores-Morales A (2011) Structural basis for c-KIT inhibition

by the suppressor of cytokine signaling 6 (SOCS6) ubiquitin ligase. J Biol Chem 286:480–490.
25. Kaneko T, Huang H, Cao X, Li X, Li C, Voss C, Sidhu SS, Li SSC (2012) Superbinder SH2 domains act as antagonists of cell signaling. Sci Signal 5:ra68.
26. Jones RB, Gordus A, Krall JA, MacBeath G (2006) A quantitative protein interaction network for the ErbB receptors using protein microarrays. Nature 439: 168–174.
27. Kaneko T, Huang H, Cao X, Li X, Li C, Voss C, Sidhu SS, Li SSC (2012) Superbinder SH2 domains act as antagonists of cell signaling. Sci Signal 5:ra68.
28. Bian Y, Li L, Dong M, Liu X, Kaneko T, Cheng K, Liu H, Voss C, Cao X, Wang Y, Litchfield D, Ye M, Li SSC, Zou H (2016) Ultra-deep tyrosine phosphoproteomics enabled by a phosphotyrosine superbinder. Nat Chem Biol 12:959–966.
29. Tong J, Cao B, Martyn GD, Krieger JR, Taylor P, Yates B, Sidhu SS, Li SSC, Mao X, Moran MF (2017) Protein-phosphotyrosine proteome profiling by superbinder-SH2 domain affinity purification mass spectrometry, sSH2-AP-MS. Proteomics 17:1600360.
30. Pawson T, Gish GD (1992) SH2 and SH3 domains: from structure to function. Cell 71:359–362.
31. Gan W, Roux B (2009) Binding specificity of SH2 domains: insight from free energy simulations. Proteins 74:996–1007.
32. Liu BA, Jablonowski K, Shah EE, Engelmann BW, Jones RB, Nash PD (2010) SH2 domains recognize contextual peptide sequence information to determine selectivity. Mol Cell Proteomics 9:2391–2404.
33. Malabarba MG, Milia E, Faretta M, Zamponi R, Pelicci PG, Di Fiore PP (2001) A repertoire library that allows the selection of synthetic SH2s with altered binding specificities. Oncogene 20:5186–5194.
34. Virdee S, Macmillan D, Waksman G (2010) Semisynthetic Src SH2 domains demonstrate altered phosphopeptide specificity induced by incorporation of unnatural lysine derivatives. Chem Biol 17:274–284.
35. Marengere LE, Songyang Z, Gish GD, Schaller MD, Parsons JT, Stern MJ, Cantley LC, Pawson T (1994) SH2 domain specificity and activity modified by a single residue. Nature 369:502–505.
36. Kaneko T, Sidhu SS, Li SSC (2011) Evolving specificity from variability for protein interaction domains. Trends Biochem Sci 36:183–190.
37. Ernst A, Appleton BA, Ivarsson Y, Zhang Y, Gfeller D, Wiesmann C, Sidhu SS (2014) A structural portrait of the PDZ domain family. J Mol Biol 426:3509–3519.
38. Ivarsson Y, Arnold R, McLaughlin M, Nim S, Joshi R, Ray D, Liu B, Teyra J, Pawson T, Moffat J, Li SSC, Sidhu SS, Kim PM (2014) Large-scale interaction profiling of PDZ domains through proteomic peptide-phage display using human and viral phage peptidomes. Proc Natl Acad Sci 111:2542–2547.
39. Hiipakka M, Saksela K (2007) Versatile retargeting of SH3 domain binding by modification of non-conserved loop residues. FEBS Lett 581:1735–1741.
40. Hiipakka M, Poikonen K, Saksela K (1999) SH3 domains with high affinity and engineered ligand specificities targeted to HIV-1 Nef. J Mol Biol 293:1097–1106.
41. Huculeci R, Garcia-Pino A, Buts L, Lenaerts T, Van NN (2015) Structural insights into the intertwined dimer of fyn SH2. Protein Sci 24:1964–1978.
42. Sidhu SS, Lowman HB, Cunningham BC, Wells JA (2000) Phage display for selection of novel binding peptides. Methods Enzymol 328:333–363.
43. Kunkel TA, Roberts JD, Zakour RA (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol 154:367–382.
44. Fellouse FA, Esaki K, Birtalan S, Raptis D, Cancasci VJ, Koide A, Jhurani P, Vasser M, Wiesmann C, Kossiakoff AA, Koide S, Sidhu SS (2007) High-throughput generation of synthetic antibodies from highly functional minimalist phage-displayed libraries. J Mol Biol 373:924–940.
45. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF (1999) Protein identification and analysis tools in the ExPASy server. Methods Mol Biol 112:531–552.
46. Tyanova S, Temu T, Cox J (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc 11:2301–2319.
47. Crooks GE, Hon G, Chandonia J-M, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Res 14: 1188–1190.