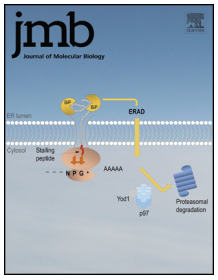


Allosteric Modulation of Binding Specificity by Alternative Packing of Protein Cores

Moshe Ben-David[1,2,†], Haiming Huang[1,2,†], Mark G.F. Sun[1,2,3], Carles Corbi-Verge[1,2,3], Evangelia Petsalaki[4,5,6], Ke Liu[7], David Gfeller[1,2], Pankaj Garg[1,2], Wolfram Tempel[7], Irina Sochirca[1,2], Julia M. Shifman[8], Alan Davidson[5], Jinrong Min[7], Philip M. Kim[1,2,3,5] and Sachdev S. Sidhu[1,2,5]

*1 - **The Donnelly Centre for Cellular and Biomolecular Research. University of Toronto,** Toronto, ON M5S 3E1, Canada*
*2 - **Banting and Best Department of Medical Research. University of Toronto,** Toronto, ON M5S 3E1, Canada*
*3 - **Department of Computer Science. University of Toronto,** Toronto, ON M5S 3E1, Canada*
*4 - **Lunenfeld Tanenbaum Research Institute. Mount Sinai Hospital,** Toronto, ON M5G 1X5, Canada*
*5 - **Department of Molecular Genetics. University of Toronto,** Toronto, ON M5S 3E1, Canada*
*6 - **European Molecular Biology Laboratory,** European Bioinformatics Institute (EMBL-EBI), CB10 1SD, Hinxton, UK*
*7 - **Structural Genomics Consortium,** University of Toronto, Toronto, ON M5S 3E1, Canada*
*8 - **Department of Biological Chemistry,** The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel*

***Correspondence to Sachdev S. Sidhu:**at: The Donnelly Centre for Cellular and Biomolecular Research. University of Toronto, Toronto, ON M5S 3E1, Canada sachdev.sidhu@utoronto.ca*
https://doi.org/10.1016/j.jmb.2018.11.018
Edited by A. Skerra

Abstract

Hydrophobic cores are often viewed as tightly packed and rigid, but they do show some plasticity and could thus be attractive targets for protein design. Here we explored the role of different functional pressures on the core packing and ligand recognition of the SH3 domain from human Fyn tyrosine kinase. We randomized the hydrophobic core and used phage display to select variants that bound to each of three distinct ligands. The three evolved groups showed remarkable differences in core composition, illustrating the effect of different selective pressures on the core. Changes in the core did not significantly alter protein stability, but were linked closely to changes in binding affinity and specificity. Structural analysis and molecular dynamics simulations revealed the structural basis for altered specificity. The evolved domains had significantly reduced core volumes, which in turn induced increased backbone flexibility. These motions were propagated from the core to the binding surface and induced significant conformational changes. These results show that alternative core packing and consequent allosteric modulation of binding interfaces could be used to engineer proteins with novel functions.

Crown Copyright © 2018 Published by Elsevier Ltd. All rights reserved.

Introduction

A common characteristic of most native proteins is a well-ordered and solvent-inaccessible core, composed mainly of hydrophobic side chains. Hydrophobic cores are often tightly packed and show densities resembling those of organic crystals [1]. They play critical roles in protein folding and in maintaining conformational stability [2–4]. The evolutionary rates of core residues are significantly slower than those of surface residues [5], and hence, hydrophobic cores are highly conserved and explore fewer combinations in sequence space [5–7]. However, as demonstrated by numerous mutagenesis studies, protein cores do show some plasticity [8–12]. Thus, proteins with new core compositions may preserve native function and biophysical properties [13–15], but they may also display new properties such as increased stability [16–19], altered binding specificity [20–22] and modulated conformational dynamics [23–25]. However, while the

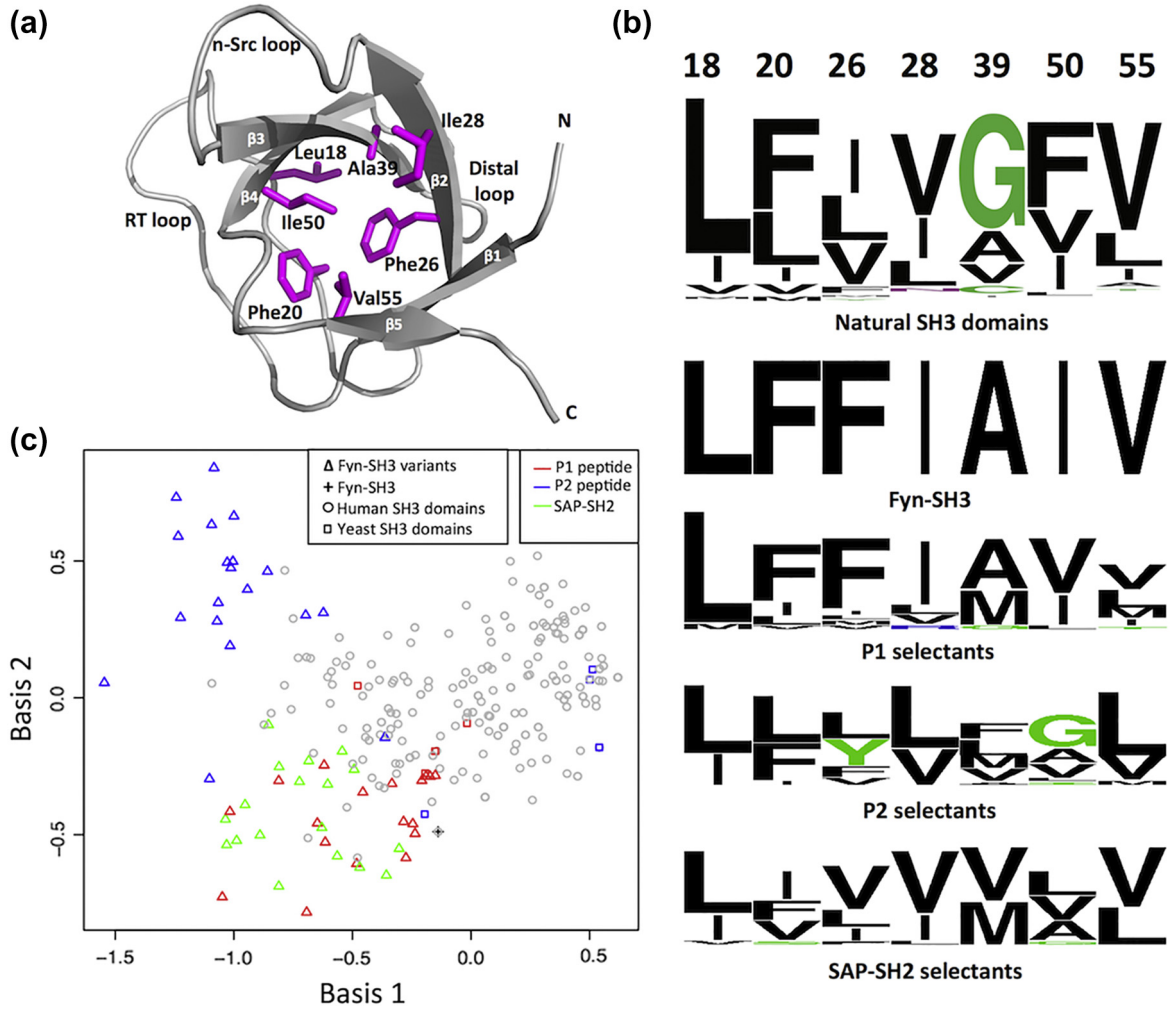


Fig. 1. Fyn-SH3 core variants evolved in response to different selection pressures. (a) Structure of human Fyn-SH3 (PDB entry 1SHF). The main chain is shown as a gray ribbon and the seven core residues that were randomized in the library are shown as magenta sticks. (b) Sequence logo representation (Weblogo [111]) of natural SH3 and Fyn-SH3 variant cores. Logos are shown for sequences of natural SH3 domains (297 human and 9 yeast SH3 domains), Fyn-SH3 and Fyn-SH3 variants selected for binding to class I peptide P1 (n = 18), class II peptide P2 (n = 19) or SAP-SH2 (n = 18). (c) 2D MCML projection of the sequence space defined by the amino acid composition of the 7 core residues for 9 yeast SH3 domains (squares), 297 human SH3 domains (circles) and 55 Fyn-SH3 variants (triangles). Symbols for Fyn-SH3 variants are colored red, blue or green for those selected for binding to the P1 peptide, the P2 peptide or SAP-SH2, respectively.

importance of core mutations for folding and structural stability is well described, the relationship between the protein core and function remains poorly understood.

Here, we used the SH3 domain from human Fyn tyrosine kinase (Fyn-SH3) as a model to explore the relationship between the protein core and ligand recognition. SH3 domains typically recognize short stretches of proline-rich sequences within proteins [26] and serve to mediate interactions within signaling pathways. These proline-rich sequences usually form a left-handed polyproline type II helix [26]. SH3 domains are classified according to the orientation of the peptide ligand, with class I or II domains recognizing peptides of the form [R/K]xxPxxP or PxxPx[R/K], respectively [27,28]. The orientation of the peptide backbone is reversed in the two classes [28,29], but in both, the Pro residues fit into pockets on the domain surface and the arginine/lysine side chain is bound in a negatively charged cleft [30] (Fig. S1A–B). Thus, a conserved peptide-binding site that is common to most SH3 domains mediates the canonical proline recognition modes. However, non-canonical binding modes have been reported [31], and some of these interactions are mediated through alternative binding sites [30,32]. Fyn-SH3 is particularly interesting because it can interact with both class I and class II ligands (Fig. S1A–C), and in addition, with the globular SH2 domain of SLAM-associated protein (SAP-SH2) through a

Table 1. Thermostabilities and affinities of Fyn-SH3 variants

Domain	Selection ligand	Sequence[a]							T_m[b] (°C)	K_D[c] (μM)		IC_{50}[d] (μM)
		18	20	26	28	39	50	55		P1	P2	SAP-SH2
wt		L	F	F	I	A	I	V	74	0.18 ± 0.02 (1)	1.70 ± 0.02 (1)	0.98 ± 0.12 (1)
A1	P1	–	–	–	–	–	V	L	74	0.16 ± 0.01 (↓ 1.1)	1.9 ± 0.2 (↑ 1.1)	1.1 ± 0.1 (↑ 1.1)
A7	P1	–	–	I	–	–	V	L	67	0.27 ± 0.01 (↑ 1.5)	0.70 ± 0.08 (↓ 2.4)	0.80 ± 0.04 (↓ 1.2)
A10	P1	–	–	–	V	M	V	–	65	0.50 ± 0.06 (↑ 2.8)	1.80 ± 0.3 (↑ 1.1)	5.4 ± 0.5 (↑ 5.5)
C10	P2	–	L	Y	L	F	S	L	63	8.7 ± 0.5 (↑ 48)	0.6 ± 0.1 (↓ 2.8)	2.70 ± 0.5 (↑ 2.8)
D10	P2	–	L	L	V	L	G	L	51	1.00 ± 0.06 (↑ 5.6)	4.0 ± 0.4 (↑ 2.4)	2.6 ± 0.5 (↑ 2.7)
E6	SAP-SH2	I	V	V	–	V	V	–	65	14.9 ± 1.2 (↑ 83)	11.2 ± 1.1 (↑ 6.6)	2.4 ± 0.5 (↑ 2.4)
E8	SAP-SH2	V	I	V	V	V	V	L	73	12.5 ± 1.1 (↑ 69)	9.5 ± 0.6 (↑ 5.6)	2.5 ± 0.4 (↑ 2.6)

[a] The amino acid sequence at each core position is shown, and dashes indicate identity with the wild type.
[b] The melting temperature (T_m) was determined by CD spectroscopy.
[c] K_D was determined by tryptophan florescence. The fold-change relative to wt is shown in parenthesis.
[d] IC_{50} was determined by competition ELISA. The fold-change relative to wt is shown in parenthesis.

distinct binding site that has only minimal overlap with the canonical binding site [33] (Fig. S1D).

While binding specificity is predominantly determined by surface residues [34,35], previous studies suggested that SH3 domains are plastic and binding specificity can be influenced by long-range effects [36–38]. Thus, the hydrophobic core may play a role in ligand recognition. To explore the role of the core in SH3 domain function, we randomized the Fyn-SH3 core and examined alternative packing compositions in response to different functional pressures. Our results indicate that Fyn-SH3 adopts different core compositions in response to different functional pressures, which alter binding specificity without any direct changes at the domain surface. Crystallographic and molecular dynamic (MD) simulation analyses of a Fyn-SH3 variant showed that mutations reduced the core volume, which in turn enhanced conformational flexibility and altered the binding surface. Our results highlight the importance of networks that mediate long-range interactions and connect core residues with surface residues.

Results

Divergent cores evolve under different functional pressures

We used *in vitro* evolution of Fyn-SH3 to explore the effects of different functional pressures on core composition and ligand specificity. We constructed a phage-displayed library to fully randomize the seven most buried residues in Fyn-SH3 (Fig. 1a), henceforth referred to as the "core". The library was subjected to three independent selections toward three distinct ligands: a class I peptide (P1, VSLARRPLPPLP) [39], a class II peptide (P2, PPLPPRNRPRL) [40], and SAP-SH2 [33,41]. A total of 55 variants with unique core sequences were selected from the three selections (Fig. S2). Analysis of these core sequences revealed remarkable diversity among the different selections, as well as relative to natural SH3 domains (Fig. 1b). On average, the variants selected for binding to P1 had the fewest mutations relative to wild-type Fyn-SH3, while those selected for binding to P2 or SAP-SH2 were more heavily mutated (Figs. S2 and S3). We also compared each variant group to natural SH3 domain sequences using a two-dimensional projection of the sequence space (Fig. 1c). This analysis showed that the core sequences of variants selected for binding to P1 or SAP-SH2 occupy sections that overlap with each other and with some of the natural SH3 sequences. In contrast, the variants selected for binding to P2 occupied a distinct region in sequence space, which lies far away from natural SH3 domains, including from those that bind class II peptides (Fig. 1c).

Close inspection of variant sequences shed light on the observed differences between the three groups. Although most of the sequences resemble naturally occurring variations (Fig. 1b), positions 39 and 50 differ greatly, especially for the P2 selectants. Most natural SH3 domains, including Fyn-SH3, contain a small residue (e.g., glycine or alanine) at position 39 and a large residue (e.g., phenylalanine

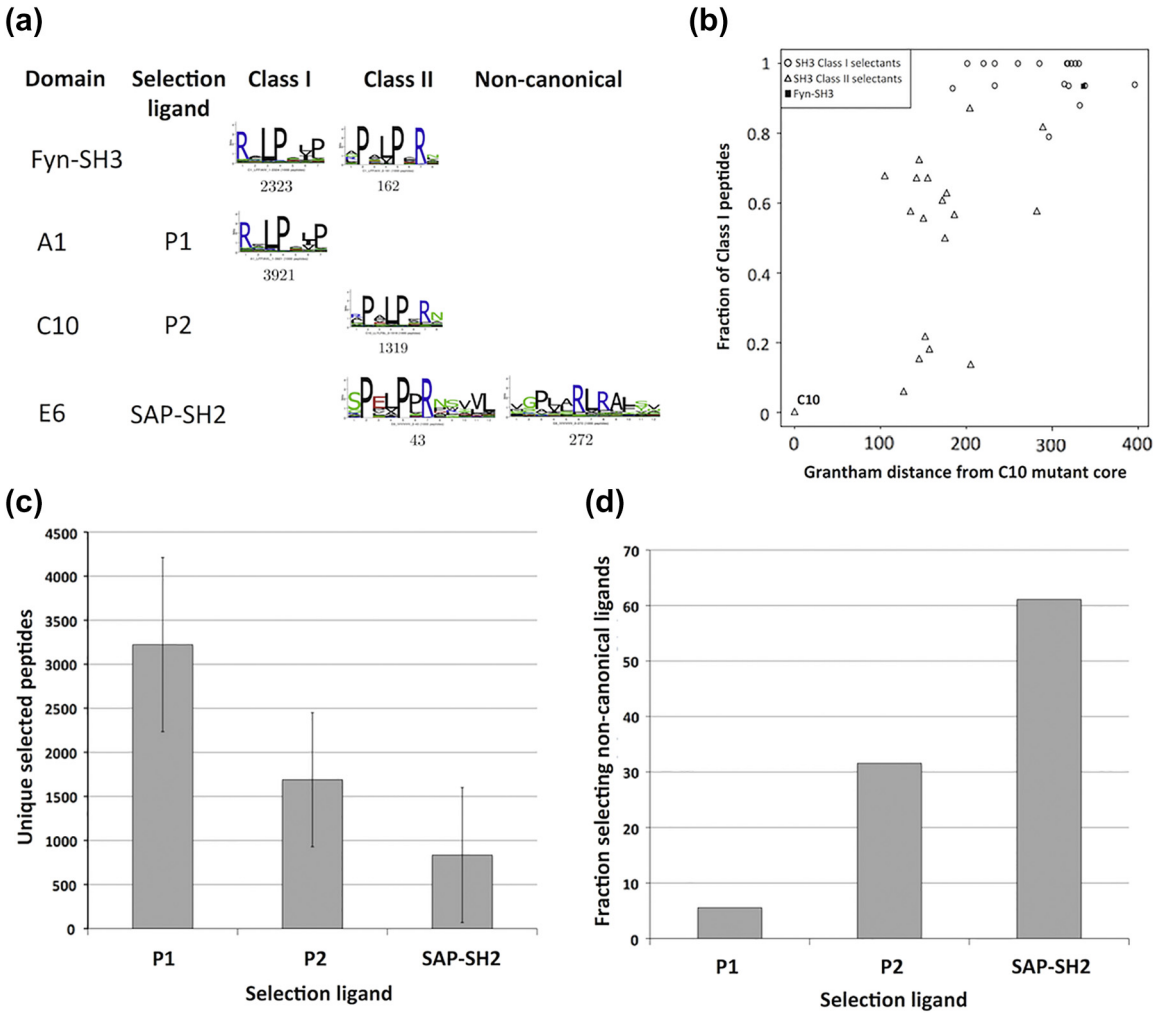


Fig. 2. Peptide-binding specificities of Fyn-SH3 core variants. (a) Binding specificities represented as logos (WebLogo server [111]) for Fyn-SH3 and representative variants selected for binding to peptide P1 (A1), peptide P2 (C10) or SAP-SH2 (E6). The number under each logo denotes the number of unique peptides used to generate the logo. (b) Divergent binding specificities correlate with divergent core compositions. The *x*-axis is a linear projection of sequence space of the cores. The *y*-axis shows the fraction of unique selected peptides corresponding to class I. The variant C10, which selected only class II peptides, was set as the origin of coordinates. The positions of other variants on the *x*-axis were determined by their Grantham distance [112] from C10. (c) The average peptide-binding capacity (i.e., number of unique selected peptides, *y*-axis) of Fyn-SH3 variants selected for binding to the indicated selection ligand (*x*-axis). (Data represent mean ± s.d.). (d) For Fyn-SH3 variants selected for binding to the indicated selection ligand (*x*-axis), the plot shows the fraction that selected at least some non-canonical peptides (*y*-axis). See Fig. S5 for details.

or isoleucine) at position 50. In contrast, most of the P2 selectants showed the opposite trend with large and small residues dominating at position 39 or 50, respectively (Fig. 1b). Furthermore, in both P1 and SAP-SH2 selectants, position 39 is frequently occupied by methionine (40%), which is rare in natural SH3 domains (<1%). Positions 39 and 50 are in close proximity (distance between side chain atoms is 3.7 Å), suggesting that coevolution of these positions may be driven by volume compensation effects. Indeed, it was shown previously that strong covariation exists between these positions [42]. The P2 selectant group is also unique in the frequent occurrence of tyrosine at position 26 (37%), which is rarely found in natural SH3 domains (<3%), and this difference is also likely coupled to differences at positions 39 and 50.

These results show that Fyn-SH3 variants adopt drastically different core compositions in response to different selective pressures. In turn, the composition of the hydrophobic core may affect the conformation of the binding surface and thus alter domain function. Therefore, to further characterize the SH3 core variants, we selected representatives from each selection group and performed stability and binding assays.

Fyn-SH3 variants are stable but show altered ligand affinities

We purified seven Fyn-SH3 variants and assessed protein stability by measuring melting temperatures (T_m) using circular dichroism (CD) spectroscopy. Wild-type Fyn-SH3 exhibited a T_m of 74 °C, which is in good agreement with a previous report [43]. The T_m values of the variants ranged from 51 to 74 °C, with a mean of 65 °C (Table 1, Fig. S4A). Thus, all tested variants were fairly stable, and any observed changes in binding affinities were likely due to changes in binding interactions (i.e., local effects rather than global effects).

We assessed the affinities of the Fyn-SH3 variants for the three ligands used in the selections, either by direct detection of binding for the peptides or assessment of competition for SAP-SH2 (Table 1, Fig. S4B–D). The three P1 selectants retained wild-type-like affinities for peptide P1. Notably, with the exception of SH3-A10, which exhibited 5-fold reduced affinity for SAP-SH2, these variants also exhibited wild-type-like affinities for peptide P2 and SAP-SH2, although there was no selective pressure to retain affinity for these ligands. The two P2 selectants showed varied binding properties. While SH3-D10 showed mild decreases in affinity for all three ligands, SH3-C10 showed an almost 50-fold decrease in affinity for peptide P1.

The two SAP-SH2 selectants showed mild decreases in affinity for SAP-SH2 (~2-fold) and peptide P2 (~5-fold), and drastic decreases in affinity for peptide P1 (>50-fold). This may be explained by the fact that the SAP-SH2 binding interface overlaps only partially with the canonical peptide-binding site [33,41] (Fig. S1C–D). Therefore, the absence of selective pressure to retain the appropriate conformation of the peptide-binding site may have led to its disruption, as indicated by the drastic decrease in affinity for peptide P1. However, this explanation is complicated by the fact that affinities for peptide P2 are only reduced slightly, although both class I and class II peptides bind to the same site, albeit in opposite orientations [28,29]. Thus, despite a common binding site, it seems that the different binding modes for class I and class II peptides resulted in different effects on binding due to core mutations evolved in response to SAP-SH2 binding.

Overall, these results suggest that alternative core compositions selected in response to different ligands may induce structural changes in the canonical peptide-binding site, which consequently have differential effects on binding affinities for class I and class II peptides. Thus, for all 55 selected variants, we further analyzed the effects of the diverse core compositions on peptide ligand specificity.

Peptide ligand specificity is linked to core composition

To comprehensively assess binding specificities of the 55 selected variants in an unbiased manner, we performed binding selections with a random phage-displayed dodecapeptide library. By deep sequencing of the phage pools, we obtained large numbers (several hundreds to thousands) of unique peptide ligands for wild-type Fyn-SH3 and for each of the 55 variants (Figs. 2a and S5). As described previously [44,45], large sets of peptide ligands can be clustered on the basis of sequence similarity to define different binding modes in cases where domains recognize more than one type of ligand. Moreover, the relative affinities of a domain for each type of ligand are roughly correlated with the proportion of total peptides assigned to each cluster (i.e., more peptides in a given cluster indicate that peptides of that type are preferred relative to peptides in other clusters).

The sequencing results largely confirmed the picture that emerged from the affinity measurements. Unlike variants from the class I selection, which resembled wild-type Fyn-SH3 with strong preferences for class I peptides, variants from the class II selection displayed diverse specificities. Interestingly, as the core composition diverged further from the wild-type, the class I binding mode became progressively weaker and the class II binding mode became stronger (Fig. 2b). Furthermore, the variants from the class II selection segregated into three sub-groups: (i) class-I-like binders, (ii) intermediate binders that bound almost equally well to class I and class II peptides, and (iii) class-II-like binders. These groups seem to represent points in the course of transition from class I domains to class II domains. In general, the peptide-binding capacity (i.e., the number of unique peptides selected by an SH3 variant) was reduced for class II domains relative to class I domains (Fig. 2c). Moreover, whereas only 1 of 18 variants from the class I selection recognized non-canonical peptides, almost one third of the variants from the class II selection were capable of recognizing non-canonical peptide motifs (Figs. 2d and S5), suggesting that core substitutions in variants from the class II selection not only altered specificity but also increased promiscuity.

Relative to both P1 and P2 selectants, the SAP-SH2 selectants showed the lowest peptide-binding capacity, (Figs. 2c and S5), presumably due to reduced binding to polyproline motifs. Furthermore, two variants (E10 and F9) did not select any peptides, suggesting that the peptide-binding surface was severely compromised. On the other hand, a majority (60%) of the SAP-SH2 selectants bound to non-canonical peptides, many of which contained a distinctive RLR sequence (Figs. 2d and S5). Notably, the surface of SAP-SH2 that interacts with Fyn-SH3 is positively charged [33], suggesting that these non-canonical peptides may bind to the SAP-SH2 binding surface rather than the polyproline-binding site. Indeed, other non-canonical peptides have been shown to bind to alternative binding sites [31].

Table 2. Summary of data collection and refinement statistics

Data collection	SH3-E6
Space group	$P2_1$
Unit cell: *a*, *b*, *c* (Å), *β* (°)	23.99, 36.23, 38.65, 99.52
Resolution range (outer shell, Å)[a]	36.23–1.40 (1.42–1.40)
Number of unique reflections[a]	12,781 (564)
R_{sym}[a]	0.053 (0.400)
Completeness (%)[a]	98.4 (87.7)
Redundancy[a]	3.7 (3.1)
Mean($I/\sigma(I)$)[a]	16.5 (3.3)
Refinement statistics of the current model	
R_{work}/R_{free}	0.143/0.172
Rmsd bonds (Å)/angles (°)	0.016/1.8
Ramachandran plot[b]	
Favored regions (%)	98.67
Outliers (%)	0.00
No. atoms	
Protein, water, others	612, 61, 20
B factors	
Protein, water, others	12.3, 20.1, 22.9
PDB code	6EDF

[a] Values in parentheses are for the highest-resolution shell.
[b] Molprobity (http://molprobity.biochem.duke.edu/).

The structure of variant SH3-E6 reveals an unusual ligand interaction and reduced core volume

To gain insights into the structural changes induced by core substitutions, we determined the crystal structure of SH3-E6, a Fyn-SH3 variant with significantly altered binding specificity. Variant SH3-E6 was selected for and retained binding to SAP-SH2, but it exhibited reduced affinity for peptides P1 and P2 (Table 1), and it bound to non-canonical peptides containing an RLR motif (Fig. 2a). We designed a fusion protein consisting of SH3-E6 with a non-canonical peptide ligand (^{64}MGPVL**RLR**AFYN^{75}, referred to as peptide RLR) fused to its C-terminus and solved the crystal structure at a resolution of 1.4 Å (Table 2). As hoped, the SH3 domain is bound to the fused peptide of a symmetry-related copy of the fusion protein inside the crystal lattice.

The binding site for peptide RLR, a groove between the RT and n-Src loops, overlapped mainly with the SAP-SH2 binding site rather than with the polyproline-binding site (Fig. 3a–c). The molecular

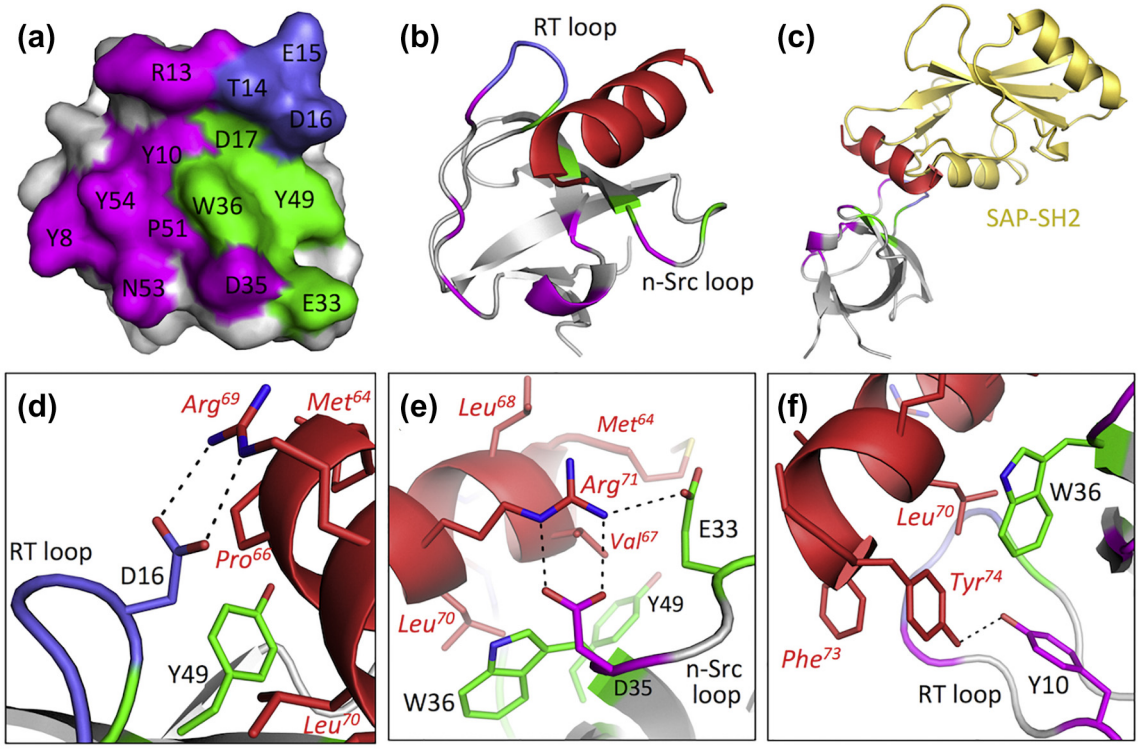


Fig. 3. The crystal structure of SH3-E6 in complex with peptide RLR. (a) Surface representation of SH3-E6 with residues that interact with polyproline peptides, SAP-SH2 or both, colored magenta, blue or green, respectively. (b) Ribbon representation of SH3-E6 (colored as in panel A) in complex with peptide RLR (red). (c) Superposition of SH3-E6 (colored as in panel A) in complex with peptide RLR (red) and Fyn-SH3 (not shown) in complex with SAP-SH2 (yellow, PDB entry 1M27). (d–f) Molecular interactions between SH3-E6 and peptide RLR residues (d) Arg^{69}, (e) Arg^{71} and (f) Tyr^{74}. Side chains are shown as sticks colored as in panel B, and salt bridges and hydrogen bonds are shown as dotted lines.

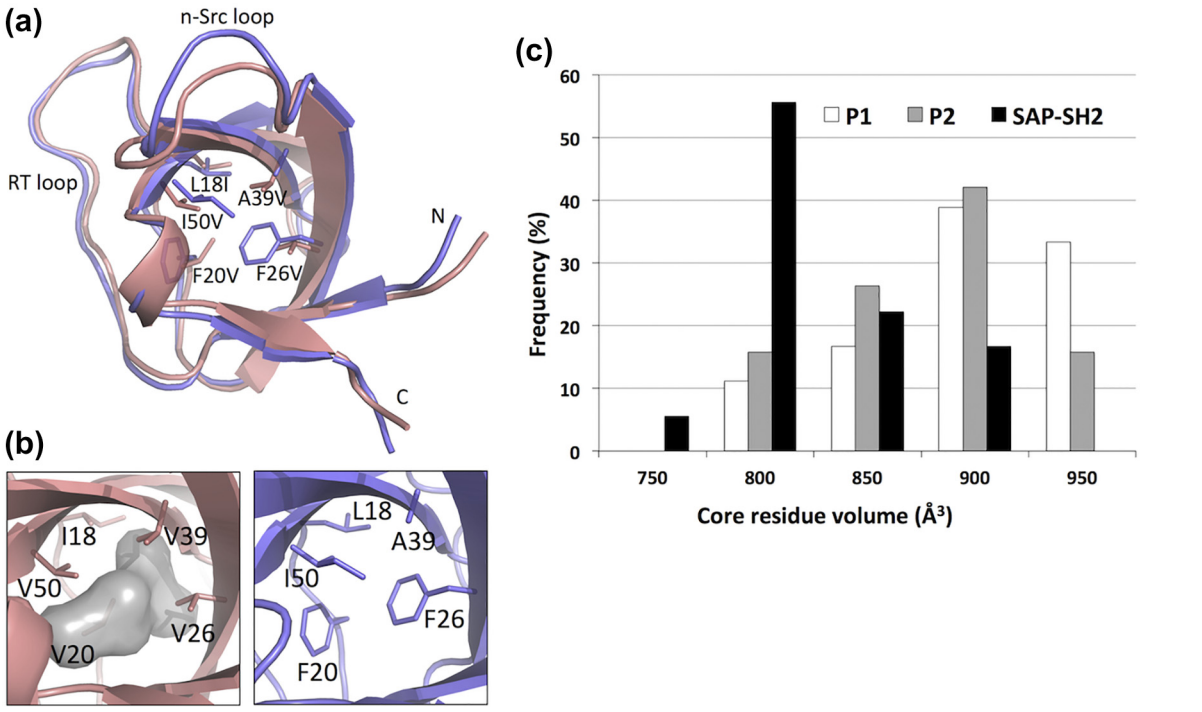
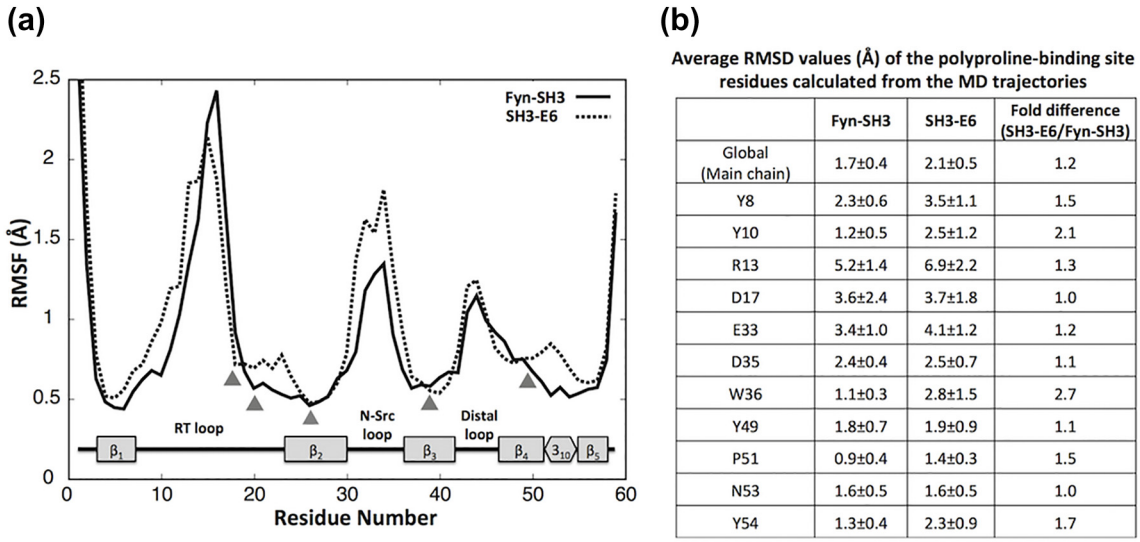


Fig. 4. Comparison of the cores of SH3-E6 and Fyn-SH3. (a) Superposition of SH3-E6 (salmon) and Fyn-SH3 (blue, PDB entry 1SHF). Side chains of core residues that differ are shown as sticks and labeled with the residue type in Fyn-SH3 or SH3-E6 to the left or right of the residue number, respectively. (b) Assessment of cavities (gray) in the cores of SH3-E6 (salmon) and Fyn-SH3 (blue). (c) Distributions of the computationally calculated core volumes of Fyn-SH3 variants selected for binding to peptide P1 (white), peptide P2 (gray) or SAP-SH2 (black). The variants were grouped according to their core residue volume in a range of 50 Å^3 (*x*-axis, e.g., "800 Å^3" indicates domains with core volume between 800 and 850 Å^3), and the frequencies were plotted (*y*-axis). The core volumes of Fyn-SH3 and SH3-E6 were 969 and 825 Å^3, respectively.

interactions between peptide RLR and SH3-E6 were consistent with the sequence conservation observed in the non-canonical peptides selected by SH3-E6 (Fig. 2a). In particular, the side chains of Arg^{69} and Arg^{71} in the RLR motif formed salt bridges with Asp^{16} in the RT loop (Fig. 3d) or with Glu^{33} and Asp^{35} in the n-Src loop (Fig. 3e), respectively, and the Tyr^{74} side chain in the peptide formed a hydrogen bond with the Tyr^{10} side chain in the RT loop (Fig. 3f). These polar interactions were reinforced by hydrophobic interactions between peptide residues Pro^{66}, Val^{67} and Leu^{70}, and SH3-E6 residues Trp^{36} and Tyr^{49} (Fig. 3d–f). Notably, several positively charged side chains in SAP-SH2 also interact with negatively charged side chains in Fyn-SH3 (Fig. S6) [33], and thus, peptide RLR partially mimics the interactions of SAP-SH2 with Fyn-SH3.

Superposition of SH3-E6 with Fyn-SH3 revealed that while the global folds were very similar (Fig. 4a), there were differences in the main chains of the RT and n-Src loops. These regions are known to be flexible and often show different conformations between bound and free SH3 domains [46–48], implying that the binding of peptide RLR contributed to the observed conformational changes. The hydrophobic core of SH3-E6 showed significantly reduced packing volume compared with that of Fyn-SH3, mainly due to valines substituting for phenylalanines at both positions 20 and 26 (Fig. 4a). Consequently, analysis with the HOLLOW program [49] revealed no detectable core cavities for Fyn-SH3, whereas SH3-E6 contained large core cavities comprising a total volume of 215 Å^3 (Fig. 4b). We also calculated the densities of the core packing for all variants by modeling the substituted side chains of each Fyn-SH3 variant based on existing crystal structures, including that of SH3-E6, and measuring the filled volume based on the van der Waals radii of the atoms in the core. All but 8 of the 55 variants were predicted to have less tightly packed cores than Fyn-SH3, and the SAP-SH2 selectants showed the smallest filled core volumes (Fig. 4c). Notably, core cavities have been shown to be positively correlated with increased conformational flexibility [50,51], suggesting that the main chains of most of the Fyn-SH3 variants are likely to be more flexible that of Fyn-SH3.

Taken together, our analysis suggested that variations in the core volumes of the Fyn-SH3 variants contributed to altered ligand specificities. The



	Fyn-SH3	SH3-E6	Fold difference (SH3-E6/Fyn-SH3)
Global (Main chain)	1.7±0.4	2.1±0.5	1.2
Y8	2.3±0.6	3.5±1.1	1.5
Y10	1.2±0.5	2.5±1.2	2.1
R13	5.2±1.4	6.9±2.2	1.3
D17	3.6±2.4	3.7±1.8	1.0
E33	3.4±1.0	4.1±1.2	1.2
D35	2.4±0.4	2.5±0.7	1.1
W36	1.1±0.3	2.8±1.5	2.7
Y49	1.8±0.7	1.9±0.9	1.1
P51	0.9±0.4	1.4±0.3	1.5
N53	1.6±0.5	1.6±0.5	1.0
Y54	1.3±0.4	2.3±0.9	1.7

Fig. 5. MD simulation analysis of Fyn-SH3 and SH3-E6. (a) Average RMSF curves of the main chain atoms of SH3-E6 (dashed line) and Fyn-SH3 (solid line). The secondary structure of Fyn-SH3 is shown below. The gray triangles indicate sequence differences between SH3-E6 and Fyn-SH3. (b) Comparison of the average RMSD values of the polyproline-binding site residues calculated from the MD trajectories of Fyn-SH3 and SH3-E6. (c) Analysis of representative MD simulation models of Fyn-SH3 (left) and SH3-E6 (right). Representative ensemble of 27 models was selected from each MD trajectory (i.e., each 6 ns, a model was sampled in the full trajectory of 160 ns of MD simulation). The peptide-binding residues are shown as sticks (colored differently for each model), and the Fyn-SH3 reference structure (PDB entry 1SHF) is shown in gray.

structure of SH3-E6 showed that the reduced core volume induced slight shifts in the main chain and side chains, presumably due to enhanced conformational flexibility (Fig. 4). This in turn may have affected the conformation of the polyproline-binding site and may have contributed to the reduced affinities for peptides P1 and P2.

MD simulations show enhanced conformational flexibility for SH3-E6

Crystal structures alone provide only a limited understanding of conformational flexibility [51], and thus, we used MD simulations to explore conformational dynamics. In four replicas, we calculated MD trajectories totaling 160 ns for Fyn-SH3 and SH3-E6 (the fused peptide RLR of SH3-E6 was omitted). To assess the effect of the new core composition on global conformational dynamics, we measured the root mean square fluctuation (RMSF) (Fig. 5a), which provided a measure of the average atomic mobility of the backbone atoms during the MD simulation. As expected, the highest values were obtained for the RT, n-Src and distal loops, with SH3-E6 showing overall increased RMSF values compared to Fyn-SH3. An increase in RMSF was also seen in the vicinity of the core substitutions, even in unexpected regions, such as the type II β-turn

(residues 20–27) and the 3_{10} helix (residues 51–54) [48,52] (Fig. 5a). These observations suggested that the conformational flexibility of SH3-E6 was enhanced relative to that of Fyn-SH3.

To gain insights into how the increased conformational flexibility of SH3-E6 perturbed the function of the polyproline-binding site, we used the MD trajectories to calculate the root mean square deviation (RMSD) of this region [42]. Compared to Fyn-SH3, almost all residues of the SH3-E6 polyproline-binding site showed higher RMSD values (Fig. 5b). Indeed, whereas Fyn-SH3 largely retained its conformation (Fig. 5c, left panel), residues in the polyproline-binding site of SH3-E6 showed significant conformational changes (Fig. 5c, right panel). These included Tyr^{8}, Tyr^{10}, Trp^{36}, Pro^{51} and Tyr^{54}, which are known to play critical roles in the formation of the polyproline-binding site and interactions with peptides (Fig. S1A–B) [28,30,42]. However, the most significant changes were observed for Trp^{36} (Fig. 5b–c), which is highly conserved in SH3 domains and is important for peptide recognition and specificity [53–55,36]. The crystal structure of SH3-E6 showed only moderate conformational changes for these residues, but this may have been due to crystal packing and interactions with peptide RLR, which were absent in the MD simulations. Indeed, consistent with our MD simulations, previously determined structures of unbound Fyn-SH3 core variants showed similar drastic conformational changes for Trp^{36} [56,57].

Altogether, these data suggest that the core composition of SH3-E6 affected the conformations of the RT and n-Src loops and thus altered the polyproline-binding site. These changes likely enabled SH3-E6 to explore a broader spectrum of conformations, which in turn affected binding specificity and enhanced promiscuity [58–60].

Discussion

We show that alternative core compositions provide a promising strategy to alter protein function. Selecting SH3 cores under different selective pressures lead to drastically different core compositions, despite the fact that the selective pressure only acted directly on the protein surface. Unlike the selection with a class I ligand that seemed to impose the tightest constraints and yielded cores closest to the wild type, the selections with a class II ligand or SAP-SH2 imposed weaker constraints that allowed for more diverse core compositions. Most importantly, alternative cores had strong and distinct influences on the affinities and specificities of SH3 variants for different ligands.

Specificity analysis of many SH3 core variants showed a prevalence of arginine residues in non-canonical peptide ligands. The crystal structure of one such ligand in complex with the variant SH3-E6 revealed the critical role that arginines played in the interaction (Fig. 3). Notably, the SAP-SH2 surface that interacts with Fyn-SH3 is also positively charged, and it seems that the arginine residues in the peptide mimic the SAP-SH2 surface and provide electrostatic complementarity at the interface (Figs. 3 and S6). This fits with previous studies, which showed that arginine residues are a key feature of other non-canonical peptide ligands [31,34], and that the arginine alone can bind to Fyn-SH3 [61].

Most SH3 variants, especially SH3-E6 and others from the SAP-SH2 selection, had reduced core volumes that could enhance conformational flexibility, as shown previously [50,51] (Fig. 5). Moreover, even relatively large volumes with altered core packing may have induced significant structural changes that may have been propagated to the surface [62]. Conformational changes induced by altered cores seemed to affect the polyproline-binding site and the RT and n-Src loops, which in turn altered binding preferences (Fig. 2c–d). Unlike changes caused by alterations in surface loops that often show conformational flexibility, the changes in the polyproline-binding surface caused by changes in the core were not necessarily expected (Fig. 5).

The evolved cores, especially those selected for binding to the class II peptide or SAP-SH2, seemed to attenuate binding to the class I peptide more than to the class II peptide (Table 1). Structures of Fyn and c-Src SH3 domains in complex with class I and class II peptides [39,63] show that the most significant differences are in interactions with the specificity pocket, which comprises Trp^{36}, Tyr^{49}, and the RT and n-Src loops, and binds residues flanking the polyproline motif of the ligand [30,64,65] (Fig. S1A). Class I peptides interact with the specificity pocket more extensively than class II peptides, and consequently, Fyn and c-Src SH3 domains bind more tightly to class I ligands than to class II ligands (Table 1) [39,40]. This suggests that the conformational changes in Trp^{36} and the surface loops perturbed the configuration of the specificity pocket, and this in turn affected binding of the class I peptide more than that of the class II peptide. This may also explain the more drastic (Fig. 1c) and extensive (Fig. S2) substitutions observed at positions 39 and 50 in the variants selected with the class II ligand and SAP-SH2 compared with those selected with the class I ligand, as these positions interact with the specificity pocket and thus may have been more constrained by the class I ligand [42,66].

The effects of different selection pressures on ligand specificity may have implications for the evolution of SH3 domains. Selections for binding to the class II peptide and SAP-SH2 seem to have pushed Fyn-SH3 into a less specialized state. The resulting variants showed greater core flexibility, which appears to be associated with more promiscuous specificity. It is

thus intriguing to speculate that class II specificity may be a more ancestral function than class I specificity. Indeed, previous studies have shown that class II specificity is more frequent than class I specificity in natural SH3 domains [67–69] and in phage-displayed peptide analysis [31]. Furthermore, peptide array screening of multiple SH3 domains showed a general bias toward class II ligands [69]. Moreover, binding of class I peptides requires a more precise positioning of Trp^{36} [36] and more extensive interactions with the specificity pocket [39,63]. Presumably, the ability to sample a broader spectrum of conformations in the less specialized state could provide promising starting points [70] for *in vitro* evolution of SH3 domains with tailored binding specificities.

These results raise an interesting question: how do structural changes in the SH3 core propagate to the surface to alter function? The adjacent antiparallel β3 and β4 strands may play a role, as they contain the key binding-site residues Trp^{36} and Tyr^{49} and the heavily substituted core positions 39 and 50 (Fig. 1a). Acting as connectors between the core and surface, these strands could transduce structural changes in the core into functional changes at the surface. Specifically, substitutions at core positions 39 and 50 may affect the conformation of the critical binding site residue Trp^{36}, and this may in turn affect ligand recognition. Moreover, these strands were shown to be part of highly connected main chain hydrogen bond networks that extend from the surface, through the hydrophobic core, to the most distant surface loops [38,48]. These networks may enable the transmission of structural changes from substitutions in the core to residues on the surface, and previous studies of SH3 domains have shown transmission of structural motions induced by either peptide binding [38,48] or by mutations [36,50,71,72]. Thus, intramolecular interaction networks mediate allosteric effects that play critical roles in protein function [73–75], and deeper understanding of these networks may facilitate the design of better binding interfaces.

Materials and Methods

Phage-displayed library construction and selection

The phage-displayed library of Fyn-SH3 core variants was constructed as described [76,77]. The codons encoding the seven most buried positions in the core (18, 20, 26, 28, 39, 50, 55) were substituted with NNK (N = A/G/C/T, K = G/T) degenerate codons that encode for all 20 genetically encoded amino acids, using oligonucleotide-directed mutagenesis with appropriately designed mutagenic oligonucleotides [78]. The constructed library contained ~10^{10} unique members, which exceeded the maximum combinatorial amino acid diversity at the seven positions (1.3×10^9).

The SH3 ligands were purified as C-terminal GST–ligand fusion proteins, including the class I peptide P1 (VSLARRPLPPLP), the class II peptide P2 (PPLPPRNRPRL), and SAP-SH2. Each GST fusion protein was immobilized on 96-well Maxisorp Immunoplates (Nunc, Rochester, NY) by incubating 100 μl protein (5 μg/ml) overnight at 4 °C. Five rounds of binding selections with phage pools representing the Fyn-SH3 library were performed against the immobilized protein, as described [76,79]. Individual phage clones from each selection round were assayed for binding to the ligand by phage ELISA, as described [80,81], and positive clones were identified as those that exhibited at least 10-fold higher signals for binding to the GST–ligand fusion protein compared with GST alone. The sequences of positive Fyn-SH3 variants were determined by sequencing the encoding DNA.

Comparison of SH3 domain core sequences

Manual alignment was performed for the 55 Fyn-SH3 variants, 9 yeast SH3 domains and 297 human SH3 domains. The sequences of the seven mutated core positions of each protein were compared against all other sequences using percent identity as a difference metric to generate a 361 × 361 difference matrix. Visualization of the clusters was performed by using a supervised dimensionality reduction method, Maximal Collapsing Metric Learning (MCML) [82], to maximize the separation between the three Fyn-SH3 selection groups from which a two-dimensional representation can be attained. In a supervised setting, each Fyn-SH3 core sequence is associated with one of three labels (class I, class II, or SAP-SH2). The MCML objective is to learn a distance function: $d(x_i, x_j | A) = (x_i - x_j)^T A(x_i - x_j)$ (Mahalanobis distance), where x_i and x_j are core sequences, such that matrix A is optimized so that sequences belonging to the same label are close together. For Fyn-SH3 core variants the selection groups were depicted by different colors (class I, red; class II, blue; SAP-SH2, green), whereas different specificities for human and yeast SH3 domains were not depicted.

Thermostability analysis

Each Fyn-SH3 variant was expressed as a 6xHis-tagged fusion protein and was purified by Ni-NTA affinity chromatography (Qiagen) followed by size exclusion chromatography (Superdex-75, GE Healthcare). Measurements of thermal stability were performed with an Aviv 62A DS CD spectrometer (Aviv Associates, Lakewood, NJ). Temperature-induced melting of the protein was monitored by changes in the CD signal (ellipticity) at 233 nm, and samples were heated from 25 to 109 °C in 2 °C increments with a 1-min equilibration time and a 5-s averaging

time. Melt profiles were fit to obtain the T_m values, as described [83].

Affinity measurements

Peptide-binding assays were performed as described previously [83]. Briefly, peptide ligands were purified as 6xHis-tagged fusion proteins to the N-terminal domain of λ repressor. Each Fyn-SH3 core variant was titrated with the peptide fusion protein, binding was monitored by measuring the tryptophan florescence (excitation at 295 nM and emission at 326 nM), and the dissociation constant was calculated, as described [83].

SAP-SH2 binding assays were performed by competitive ELISA [76,79] with human SAP-SH2 (residues 1–104) purified as described [84,85]. Briefly, a subsaturating concentration of Fyn-SH3 variant was incubated with various concentrations of SAP-SH2 in solution and unbound Fyn-SH3 variant was captured with immobilized SAP-SH2 and measured, as described [76,79]. The IC_{50} value was determined as the concentration of solution-phase SAP-SH2 that prevented 50% of Fyn-SH3 variant binding to immobilized SAP-SH2.

SH3 domain specificity profiling by peptide-phage display

Fyn-SH3 variants were purified in a high-throughput manner as C-terminal GST fusion proteins, as described [86]. Peptide-phage selections were performed using a library of 10^{10} random dodecapeptides fused to the N-terminus of the M13 bacteriophage major coat protein (p8), as described [86]. Briefly, binding selections were performed in a 96-well format with one well dedicated to each Fyn-SH3 variant. Phage pools representing the peptide library were cycled through five rounds of binding selections against the immobilized Fyn-SH3 variants, the output from round 5 was amplified by a PCR that added unique barcodes for each Fyn-SH3 variant, and the pooled samples were subjected to deep sequencing, as described [86,87]. Briefly, phage pools were amplified with unique combinations of barcoded primers for 12–24 cycles using Phusion DNA polymerase (NEB). Amplified DNA products were analyzed by gel electrophoresis and quantified using Picogreen (Thermo Fisher Scientific), and normalized pools were subjected to deep sequencing analysis (Illumina Solexa, paired end 100 base reads). Peptide-binding specificity profiling was carried out as described [31]. The deep sequencing output was analyzed and filtered using the PHRED quality score, as well as for absence of cysteine residues and stop codons, to provide high-quality reads. This analysis resulted in ~10^5 peptide sequences that were used to generate multiple binding specificity profile logos using the MUSI software [44].

Calculation of core volumes

The sequences of natural human SH3 domains with structures in the Protein Data Bank (PDB) were extracted [88], and using CD-HIT [89], redundant domains with sequence identity greater than 90% were removed to produce a set of 123 SH3 domains. The STAMP package [90] was used to superpose these structures and to perform a structure-based multiple alignment, which was used to define the residues for each domain that corresponded to the seven core residues of Fyn-SH3 that were diversified in the library. A “core box” was defined as the volume that contained all side chain atoms of the defined core residues, and for each SH3 domain structure from the PDB, the core volume was calculated as the volume occupied by atoms of the core residues by adding the volumes of their van der Waals spheres. To obtain an estimate of the core volume of each Fyn-SH3 variant, MODELLER [91] was used to create a homology model that was used to calculate the core volume, as described above.

Crystallization, diffraction, structure determination and refinement

The variant SH3-E6 was purified as a fusion protein with a C-terminal extension consisting of a four-residue linker followed by peptide RLR (GSAAMGPVLRLRAFYN), as described [31]. The protein sample was concentrated to 15 mg/ml and subjected to crystallization trials. Crystals formed at 18 °C under sitting drop vapor diffusion with the precipitant buffer containing 20% PEG 3350 and 0.2 M magnesium acetate.

Diffraction data were collected at the Canadian Light Source, beamline 08ID [92], integrated with XDS [93] and merged with the program AIMLESS [94] from CCP4 [95]. The structure was solved by molecular replacement with PHASER [96] using coordinates of mouse Fyn-SH3 (PDB entry 3UF4, Joint Center for Structural Genomics & Partnership for T-Cell Biology) as a model. Residues of the C-terminal extension were manually built using COOT [97] into a map from ARP/WARP [98]. The current model was obtained after several rounds of manual rebuilding in COOT, refinement in REFMAC [99] and validation with MOLPROBITY [100] within the PHENIX suite [101]. PDB [102] deposition was prepared with PDB_EXTRACT [103].

MD simulations

The crystal structure coordinates of human Fyn-SH3 (PBD entry 1SHF) and SH3-E6 were used as starting points for MD simulation analysis. Each structure model was explicitly solvated by 3370 TIP3P water molecules [104] in truncated octahedral periodic

boundary conditions using TLEAP [105]. Sodium counter ions were added for overall charge neutrality and periodic boundary conditions were applied. MDs were performed using the SANDER module in the AMBER 10 package with FF03 force field [106]. Bonds to hydrogen were constrained using SHAKE [107] to permit a 2-fs time step, and the particle mesh Ewald [108] algorithm was used to treat long-range electrostatic interactions. The non-bonded cutoff was set at 12.0 Å. Systems were energy minimized using a combination of steepest descent and conjugate gradient methods. Each system was equilibrated and heated over 100 ps to 300 K, and positional restraints were gradually removed. The Langevin thermostat was used to maintain the temperature of the systems at 300 K with a collision frequency of 1 ps^{-1} [109]. Production MD runs of four replicates of 40 ns for each system were obtained from randomized starting velocities, and a total of 160 ns of conformational space exploration was obtained for each system. Structural alignments were performed fitting the atoms of the main chain of each domain to the crystal structure of Fyn-SH3. The RMSF and the RMSD of all the replicas were obtained using the analysis tool VMD [110]. Hydrogen bonds were identified and quantified using CPPTRAJ in AMBER with a length cutoff of 3.5 Å.

Accession numbers

Coordinates and structure factors of the Fyn-SH3 variant SH3-E6 have been deposited in the PDB under the accession number 6EDF.

Acknowledgments

This work was supported by grants from the Canadian Institute for Health Research, the Natural Science and Engineering Research Council, the Ontario Research Fund and the Canada Foundation for Innovation. D.G. acknowledges the financial support of EMBO (ALTF 241-2010). We thank Robert Lam for conducting diffraction experiments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2018.11.018.

Received 22 September 2018;
Received in revised form 4 November 2018;
Accepted 14 November 2018
Available online 22 November 2018

Keywords:
SH3 domain;
hydrophobic core;
phage display;
conformational flexibility;
binding specificity

†These authors contributed equally to this work.

Abbreviations used:
Fyn-SH3, Fyn tyrosine kinase; SAP-SH2, SH2 domain of SLAM-associated protein; MD, molecular dynamic; CD, circular dichroism; RMSF, root mean square fluctuation; MCML, Maximal Collapsing Metric Learning; PDB, Protein Data Bank.

References

[1] F.M. Richards, Areas, volumes, packing and protein structure, Annu. Rev. Biophys. Bioeng. 6 (1977) 151–176.
[2] W.A. Lim, D.C. Farruggio, R.T. Sauer, Structural and energetic consequences of disruptive mutations in a protein core, Biochemistry 31 (1992) 4324–4333.
[3] M. Levitt, M. Gerstein, E. Huang, S. Subbiah, J. Tsai, Protein folding: the endgame, Annu. Rev. Biochem. 66 (1997) 549–579.
[4] W.S. Sandberg, T.C. Terwilliger, Energetics of repacking a protein interior, Proc. Natl. Acad. Sci. U. S. A. 88 (1991) 1706–1710.
[5] A. Toth-Petroczy, D.S. Tawfik, Slow protein evolutionary rates are dictated by surface-core association, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 11151–11156.
[6] P. Koehl, M. Levitt, De novo protein design. II. Plasticity in sequence space, J. Mol. Biol. 293 (1999) 1183–1193.
[7] J. Chen, W.E. Stites, Packing is a key selection factor in the evolution of protein hydrophobic cores, Biochemistry 40 (2001) 15280–15289.
[8] R.K. Jain, R. Ranganathan, Local complexity of amino acid interactions in a protein core, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 111–116.
[9] E.P. Baldwin, B.W. Matthews, Core-packing constraints, hydrophobicity and protein design, Curr. Opin. Biotechnol. 5 (1994) 396–402.
[10] J. Chen, Z. Lu, J. Sakon, W.E. Stites, Proteins with simplified hydrophobic cores compared to other packing mutants, Biophys. Chem. 110 (2004) 239–248.
[11] G.A. Lazar, T.M. Handel, Hydrophobic core packing and protein design, Curr. Opin. Chem. Biol. 2 (1998) 675–679.
[12] W.A. Lim, R.T. Sauer, Alternative packing arrangements in the hydrophobic core of lambda repressor, Nature 339 (1989) 31–36.
[13] S.Y. Lee, L. Pullen, D.J. Virgil, C.A. Castaneda, D. Abeykoon, D.N. Bolon, et al., Alanine scan of core positions in ubiquitin reveals links between dynamics, stability, and function, J. Mol. Biol. 426 (2014) 1377–1389.
[14] J.R. Lai, J.D. Fisk, B. Weisblum, S.H. Gellman, Hydrophobic core repacking in a coiled-coil dimer via phage display: insights into plasticity and specificity at a protein–protein interface, J. Am. Chem. Soc. 126 (2004) 10514–10515.
[15] M.D. Finucane, D.N. Woolfson, Core-directed protein design. II. Rescue of a multiply mutated and destabilized variant of ubiquitin, Biochemistry 38 (1999) 11613–11623.

[16] G.S. Murphy, J.L. Mills, M.J. Miley, M. Machius, T. Szyperski, B. Kuhlman, Increasing sequence diversity with flexible backbone protein design: the complete redesign of a protein hydrophobic core, Structure 20 (2012) 1086–1096.
[17] S. Ventura, M.C. Vega, E. Lacroix, I. Angrand, L. Spagnolo, L. Serrano, Conformational strain in the hydrophobic core and its implications for protein folding and design, Nat. Struct. Biol. 9 (2002) 485–493.
[18] W.A. Lim, A. Hodel, R.T. Sauer, F.M. Richards, The crystal structure of a mutant protein with altered but improved hydrophobic core packing, Proc. Natl. Acad. Sci. U. S. A. 91 (1994) 423–427.
[19] W.A. Lim, R.T. Sauer, The role of internal packing interactions in determining the structure and stability of a protein, J. Mol. Biol. 219 (1991) 359–376.
[20] A. Haririnia, R. Verma, N. Purohit, M.Z. Twarog, R.J. Deshaies, D. Bolon, et al., Mutations in the hydrophobic core of ubiquitin differentially affect its recognition by receptor proteins, J. Mol. Biol. 375 (2008) 979–996.
[21] J.A. Hunt, M. Ahmed, C.A. Fierke, Metal binding specificity in carbonic anhydrase is influenced by conserved hydrophobic core residues, Biochemistry 38 (1999) 9054–9062.
[22] M. Shimaoka, J.M. Shifman, H. Jing, J. Takagi, S.L. Mayo, T.A. Springer, Computational design of an integrin I domain stabilized in the open high affinity conformation, Nat. Struct. Biol. 7 (2000) 674–678.
[23] S. Michielssens, J.H. Peters, D. Ban, S. Pratihar, D. Seeliger, M. Sharma, et al., A designed conformational shift to control protein binding specificity, Angew. Chem. Int. Ed. Engl. 53 (2014) 10367–10371.
[24] Y. Zhang, L. Zhou, L. Rouge, A.H. Phillips, C. Lam, P. Liu, et al., Conformational stabilization of ubiquitin yields potent and selective inhibitors of USP7, Nat. Chem. Biol. 9 (2013) 51–58.
[25] E.C. Johnson, G.A. Lazar, J.R. Desjarlais, T.M. Handel, Solution structure and dynamics of a designed hydrophobic core variant of ubiquitin, Structure 7 (1999) 967–976.
[26] R. Ren, B.J. Mayer, P. Cicchetti, D. Baltimore, Identification of a ten-amino acid proline-rich SH3 binding site, Science 259 (1993) 1157–1161.
[27] H. Yu, J.K. Chen, S. Feng, D.C. Dalgarno, A.W. Brauer, S.L. Schreiber, Structural basis for the binding of proline-rich peptides to SH3 domains, Cell 76 (1994) 933–945.
[28] S. Feng, J.K. Chen, H. Yu, J.A. Simon, S.L. Schreiber, Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3–ligand interactions, Science 266 (1994) 1241–1247.
[29] W.A. Lim, F.M. Richards, R.O. Fox, Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains, Nature 372 (1994) 375–379.
[30] S.S. Li, Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction, Biochem. J. 390 (2005) 641–653.
[31] J. Teyra, H. Huang, S. Jain, X. Guan, A. Dong, Y. Liu, et al., Comprehensive analysis of the human SH3 domain family reveals a wide variety of non-canonical specificities, Structure 25 (2017) 1598–1610 (e3).
[32] R. Tonikian, X. Xin, C.P. Toret, D. Gfeller, C. Landgraf, S. Panni, et al., Bayesian modeling of the yeast SH3 domain interactome predicts spatiotemporal dynamics of endocytosis proteins, PLoS Biol. 7 (2009), e1000218.
[33] B. Chan, A. Lanyi, H.K. Song, J. Griesbach, M. Simarro-Grande, F. Poy, et al., SAP couples Fyn to SLAM immune receptors, Nat. Cell Biol. 5 (2003) 155–160.
[34] J. Kim, C.D. Lee, A. Rath, A.R. Davidson, Recognition of non-canonical peptides by the yeast Fus1p SH3 domain: elucidation of a common mechanism for diverse SH3 domain specificities, J. Mol. Biol. 377 (2008) 889–901.
[35] S. Panni, L. Dente, G. Cesareni, In vitro evolution of recognition specificity mediated by SH3 domains reveals target recognition rules, J. Biol. Chem. 277 (2002) 21666–21674.
[36] G. Fernandez-Ballester, C. Blanes-Mira, L. Serrano, The tryptophan switch: changing ligand-binding specificity from type I to type II in SH3 domains, J. Mol. Biol. 335 (2004) 619–629.
[37] G. Cesareni, S. Panni, G. Nardelli, L. Castagnoli, Can we infer peptide recognition specificity mediated by SH3 domains? FEBS Lett. 513 (2002) 38–44.
[38] F. Cordier, C. Wang, S. Grzesiek, L.K. Nicholson, Ligand-induced strain in hydrogen bonds of the c-Src SH3 domain detected by NMR, J. Mol. Biol. 304 (2000) 497–505.
[39] S. Feng, C. Kasahara, R.J. Rickles, S.L. Schreiber, Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 12408–12415.
[40] R.J. Rickles, M.C. Botfield, X.M. Zhou, P.A. Henry, J.S. Brugge, M.J. Zoller, Phage display selection of ligand residues important for Src homology 3 domain binding specificity, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 10909–10913.
[41] S. Latour, R. Roncagalli, R. Chen, M. Bakinowski, X. Shi, P.L. Schwartzberg, et al., Binding of SAP SH2 domain to FynT SH3 domain reveals a novel mechanism of receptor signalling in immune regulation, Nat. Cell Biol. 5 (2003) 149–154.
[42] S.M. Larson, A.A. Di Nardo, A.R. Davidson, Analysis of covariation in an SH3 domain sequence alignment: applications in tertiary contact prediction and the design of compensating hydrophobic core substitutions, J. Mol. Biol. 303 (2000) 433–446.
[43] A.A. Di Nardo, S.M. Larson, A.R. Davidson, The relationship between conservation, thermodynamic stability, and function in the SH3 domain hydrophobic core, J. Mol. Biol. 333 (2003) 641–655.
[44] T. Kim, M.S. Tyndel, H. Huang, S.S. Sidhu, G.D. Bader, D. Gfeller, et al., MUSI: an integrated system for identifying multiple specificity from very large peptide or nucleic acid data sets, Nucleic Acids Res. 40 (2012), e47.
[45] D. Gfeller, F. Butty, M. Wierzbicka, E. Verschueren, P. Vanhee, H. Huang, et al., The multiple-specificity landscape of modular peptide recognition domains, Mol. Syst. Biol. 7 (2011) 484.
[46] C.J. Morton, D.J. Pugh, E.L. Brown, J.D. Kahmann, D.A. Renzoni, I.D. Campbell, Solution structure and peptide binding of the SH3 domain from human Fyn, Structure 4 (1996) 705–714.
[47] S. Arold, R. O'Brien, P. Franken, M.P. Strub, F. Hoh, C. Dumas, et al., RT loop flexibility enhances the specificity of Src family SH3 domains for HIV-1 Nef, Biochemistry 37 (1998) 14683–14691.
[48] C. Wang, N.H. Pawley, L.K. Nicholson, The role of backbone motions in ligand binding to the c-Src SH3 domain, J. Mol. Biol. 313 (2001) 873–887.
[49] B.K. Ho, F. Gruswitz, HOLLOW: generating accurate representations of channel and interior surfaces in molecular structures, BMC Struct. Biol. 8 (2008) 49.
[50] A. Mittermaier, L.E. Kay, The response of internal dynamics to hydrophobic core mutations in the SH3 domain from the Fyn tyrosine kinase, Protein Sci. 13 (2004) 1088–1099.

[51] C.J. Lopez, Z. Yang, C. Altenbach, W.L. Hubbell, Conformational selection and adaptation to ligand binding in T4 lysozyme cavity mutants, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) E4306–E4315.
[52] W.T. Chu, J.L. Zhang, Q.C. Zheng, L. Chen, H.X. Zhang, Insights into the folding and unfolding processes of wild-type and mutated SH3 domain by molecular dynamics and replica exchange molecular dynamics simulations, PLoS One 8 (2013), e64886.
[53] B. Fazi, M.J. Cope, A. Douangamath, S. Ferracuti, K. Schirwitz, A. Zucconi, et al., Unusual binding properties of the SH3 domain of the yeast actin-binding protein Abp1: structural and functional analysis, J. Biol. Chem. 277 (2002) 5290–5298.
[54] G. Bottger, P. Barnett, A.T. Klein, A. Kragt, H.F. Tabak, B. Distel, *Saccharomyces cerevisiae* PTS1 receptor Pex5p interacts with the SH3 domain of the peroxisomal membrane protein Pex13p in an unconventional, non-PXXP-related manner, Mol. Biol. Cell 11 (2000) 3963–3976.
[55] W. Yang, S.N. Malek, S. Desiderio, An SH3-binding site conserved in Bruton's tyrosine kinase and related tyrosine kinases mediates specific protein interactions in vitro and in vivo, J. Biol. Chem. 270 (1995) 20832–20840.
[56] P. Neudecker, P. Robustelli, A. Cavalli, P. Walsh, P. Lundstrom, A. Zarrine-Afsar, et al., Structure of an intermediate state in protein folding and aggregation, Science 336 (2012) 362–366.
[57] A. Zarrine-Afsar, S. Wallin, A.M. Neculai, P. Neudecker, P.L. Howell, A.R. Davidson, et al., Theoretical and experimental demonstration of the importance of specific nonnative interactions in protein folding, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 9999–10004.
[58] L. Hou, M.T. Honaker, L.M. Shireman, L.M. Balogh, A.G. Roberts, K.C. Ng, et al., Functional promiscuity correlates with conformational heterogeneity in A-class glutathione *S*-transferases, J. Biol. Chem. 282 (2007) 23264–23274.
[59] M. Munz, J. Hein, P.C. Biggin, The role of flexibility and conformational selection in the binding promiscuity of PDZ domains, PLoS Comput. Biol. 8 (2012), e1002749.
[60] G. Schreiber, A.E. Keating, Protein binding specificity versus promiscuity, Curr. Opin. Struct. Biol. 21 (2011) 50–61.
[61] A. Zarrine-Afsar, A. Mittermaier, L.E. Kay, A.R. Davidson, Protein stabilization by specific binding of guanidinium to a functional arginine-binding surface on an SH3 domain, Protein Sci. 15 (2006) 162–170.
[62] E.C. Johnson, T.M. Handel, Effect of hydrophobic core packing on sidechain dynamics, J. Biomol. NMR 15 (1999) 135–143.
[63] J. Bacarizo, A. Camara-Artigas, Atomic resolution structures of the c-Src SH3 domain in complex with two high-affinity peptides from classes I and II, Acta Crystallogr. D Biol. Crystallogr. 69 (2013) 756–766.
[64] T.M. Kapoor, A.H. Andreotti, S.L. Schreiber, Exploring the specificity pockets of two homologous SH3 domains using structure-based, split-pool synthesis and affinity-based selection, J. Am. Chem. Soc. 120 (1998) 23–29.
[65] K. Saksela, P. Permi, SH3 domain ligand binding: what's the consensus and where's the specificity? FEBS Lett. 586 (2012) 2609–2614.
[66] E.J. Stollar, H. Lin, A.R. Davidson, J.D. Forman-Kay, Differential dynamic engagement within 24 SH3 domain: peptide complexes revealed by co-linear chemical shift perturbation analysis, PLoS One 7 (2012), e51282.
[67] T. Kaneko, L. Li, S.S. Li, The SH3 domain—a family of versatile peptide- and protein-recognition module, Front. Biosci. 13 (2008) 4938–4952.
[68] S. Karkkainen, M. Hiipakka, J.H. Wang, I. Kleino, M. Vaha-Jaakkola, G.H. Renkema, et al., Identification of preferred protein interactions by phage-display of the human Src homology-3 proteome, EMBO Rep. 7 (2006) 186–191.
[69] C. Wu, M.H. Ma, K.R. Brown, M. Geisler, L. Li, E. Tzeng, et al., Systematic identification of SH3 domain-mediated human protein–protein interactions by peptide array target screening, Proteomics 7 (2007) 1775–1785.
[70] N. Tokuriki, D.S. Tawfik, Protein dynamism and evolvability, Science 324 (2009) 203–207.
[71] A.A. Di Nardo, D.M. Korzhnev, P.J. Stogios, A. Zarrine-Afsar, L.E. Kay, A.R. Davidson, Dramatic acceleration of protein folding by stabilization of a nonnative backbone conformation, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 7954–7959.
[72] X. Espanel, M. Sudol, Yes-associated protein and p53-binding protein-2 interact through their WW and SH3 domains, J. Biol. Chem. 276 (2001) 14514–14523.
[73] M. Ben-David, J.L. Sussman, C.I. Maxwell, K. Szeler, S.C.L. Kamerlin, D.S. Tawfik, Catalytic stimulation by restrained active-site floppiness—the case of high density lipoprotein-bound serum paraoxonase-1, J. Mol. Biol. 427 (2015) 1359–1374.
[74] E.C. Campbell, G.J. Correy, P.D. Mabbitt, A.M. Buckle, N. Tokuriki, C.J. Jackson, Laboratory evolution of protein conformational dynamics, Curr. Opin. Struct. Biol. 50 (2017) 49–57.
[75] B.K. Kobilka, X. Deupi, Conformational complexity of G-protein-coupled receptors, Trends Pharmacol. Sci. 28 (2007) 397–406.
[76] A. Ernst, G. Avvakumov, J. Tong, Y. Fan, Y. Zhao, P. Alberts, et al., A strategy for modulation of enzymes in the ubiquitin system, Science 339 (2013) 590–595.
[77] R. Tonikian, Y. Zhang, C. Boone, S.S. Sidhu, Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries, Nat. Protoc. 2 (2007) 1368–1386.
[78] T.A. Kunkel, J.D. Roberts, R.A. Zakour, Rapid and efficient site-specific mutagenesis without phenotypic selection, Methods Enzymol. 154 (1987) 367–382.
[79] M. Gorelik, S. Orlicky, M.A. Sartori, X. Tang, E. Marcon, I. Kurinov, et al., Inhibition of SCF ubiquitin ligases by engineered ubiquitin variants that target the Cul1 binding site on the Skp1–F-box interface, Proc. Natl. Acad. Sci. U. S. A. 113 (2016) 3527–3532.
[80] S. Miersch, Z. Li, R. Hanna, M.E. McLaughlin, M. Hornsby, T. Matsuguchi, et al., Scalable high throughput selection from phage-displayed synthetic antibody libraries, J. Vis. Exp. 51492 (2015).
[81] J. Nilvebrant, S.S. Sidhu, Construction of synthetic antibody phage-display libraries, Methods Mol. Biol. 2018 (1701) 45–60.
[82] A. Globerson, S. Roweis, Metric learning by collapsing classes, Advances of Neural Information Processing Systems, 18, 2006, pp. 451–458.
[83] K.L. Maxwell, A.R. Davidson, Mutagenesis of a buried polar interaction in an SH3 domain: sequence conservation provides the best prediction of stability effects, Biochemistry 37 (1998) 16172–16182.
[84] C. Li, C. Iosef, C.Y. Jia, V.K. Han, S.S. Li, Dual functional roles for the X-linked lymphoproliferative syndrome gene product SAP/SH2D1A in signaling through the signaling lymphocyte activation molecule (SLAM) family of immune receptors, J. Biol. Chem. 278 (2003) 3852–3859.

[85] F. Poy, M.B. Yaffe, J. Sayos, K. Saxena, M. Morra, J. Sumegi, et al., Crystal structures of the XLP protein SAP reveal a class of SH2 domains with extended, phosphotyrosine-independent sequence recognition, Mol. Cell 4 (1999) 555–561.
[86] H. Huang, S.S. Sidhu, Studying binding specificities of peptide recognition modules by high-throughput phage display selections, Methods Mol. Biol. 781 (2011) 87–97.
[87] M.E. McLaughlin, S.S. Sidhu, Engineering and analysis of peptide-recognition domain specificities by phage display and deep sequencing, Methods Enzymol. 523 (2013) 327–349.
[88] P.W. Rose, B. Beran, C. Bi, W.F. Bluhm, D. Dimitropoulos, D.S. Goodsell, et al., The RCSB Protein Data Bank: redesigned web site and web services, Nucleic Acids Res. 39 (2011) D392–D401.
[89] W. Li, L. Jaroszewski, A. Godzik, Clustering of highly homologous sequences to reduce the size of large protein databases, Bioinformatics 17 (2001) 282–283.
[90] R.B. Russell, G.J. Barton, Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels, Proteins 14 (1992) 309–323.
[91] N. Eswar, B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M.Y. Shen, et al., Comparative protein structure modeling using Modeller, Curr. Protoc. Bioinformatics 15 (2006) 5.6.
[92] P. Grochulski, M.N. Fodje, J. Gorin, S.L. Labiuk, R. Berg, Beamline 08ID-1, the prime beamline of the Canadian Macromolecular Crystallography Facility, J. Synchrotron Radiat. 18 (2011) 681–684.
[93] W. Kabsch, Xds, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 125–132.
[94] P.R. Evans, G.N. Murshudov, How good are my data and what is the resolution? Acta Crystallogr. D Biol. Crystallogr. 69 (2013) 1204–1214.
[95] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, et al., Overview of the CCP4 suite and current developments, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 235–242.
[96] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, J. Appl. Crystallogr. 40 (2007) 658–674.
[97] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 486–501.
[98] G. Langer, S.X. Cohen, V.S. Lamzin, A. Perrakis, Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7, Nat. Protoc. 3 (2008) 1171–1179.
[99] G.N. Murshudov, P. Skubak, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, et al., REFMAC5 for the refinement of macromolecular crystal structures, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 355–367.
[100] V.B. Chen, W.B. Arendall III, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, et al., MolProbity: all-atom structure validation for macromolecular crystallography, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 12–21.
[101] P.D. Adams, P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, et al., PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 213–221.
[102] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, et al., The Protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242.
[103] H. Yang, V. Guranovic, S. Dutta, Z. Feng, H.M. Berman, J.D. Westbrook, Automated and accurate deposition of structures solved by X-ray diffraction to the Protein Data Bank, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 1833–1839.
[104] Y. Sun, R.A. Latour, Comparison of implicit solvent models for the simulation of protein-surface interactions, J. Comput. Chem. 27 (2006) 1908–1922.
[105] D.A. Case, T.A. Darden, T.E. Cheatham Iii, C.L. Simmerling, J. Wang, R.E. Duke, et al., AMBER 12, University of California, San Francisco, 2012 142.
[106] D.A. Case, T.A. Darden, T.E. Cheatham, C.L. Simmerling, J. Wang, R.E. Duke, et al., Amber 10, University of California, 2008.
[107] J.-P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes, J. Comput. Phys. 23 (1977) 327–341.
[108] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an Nṡlog(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[109] D.S. Cerutti, R. Duke, P.L. Freddolino, H. Fan, T.P. Lybrand, Vulnerability in popular molecular dynamics packages concerning Langevin and Andersen dynamics, J. Chem. Theory Comput. 4 (2008) 1669–1680.
[110] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38 (27-8).
[111] G.E. Crooks, G. Hon, J.M. Chandonia, S.E. Brenner, WebLogo: a sequence logo generator, Genome Res. 14 (2004) 1188–1190.
[112] R. Grantham, Amino acid difference formula to help explain protein evolution, Science 185 (1974) 862–864.