Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 10909–10913, November 1995
Biochemistry

Phage display selection of ligand residues important for Src homology 3 domain binding specificity

RICHARD J. RICKLES*, MARTYN C. BOTFIELD, XIAO-MAI ZHOU, PAMELA A. HENRY, JOAN S. BRUGGE, AND MARK J. ZOLLER

ARIAD Pharmaceuticals, 26 Landsdowne Street, Cambridge, MA 02139

Communicated by Ed Harlow, Massachusetts General Hospital Cancer Center, Charlestown, MA, August 14, 1995

ABSTRACT **The Src homology 3 (SH3) domain is a 50-aa modular unit present in many cellular proteins involved in intracellular signal transduction. It functions to direct protein–protein interactions through the recognition of proline-rich motifs on associated proteins. SH3 domains are important regulatory elements that have been demonstrated to specify distinct regulatory pathways important for cell growth, migration, differentiation, and responses to the external milieu. By the use of synthetic peptides, ligands have been shown to consist of a minimum core sequence and to bind to SH3 domains in one of two pseudosymmetrical orientations, class I and class II. The class I sites have the consensus sequence ZP(L/P)PPΨP whereas the class II consensus is PPΨPPZ (where Ψ is a hydrophobic residue and Z is a SH3 domain-specific residue). We previously showed by M13 phage display that the Src, Fyn, Lyn, and phosphatidylinositol 3-kinase (PI3K) SH3 domains preferred the same class I-type core binding sequence, RPLPPΨP. These results failed to explain the specificity for cellular proteins displayed by SH3 domains in cells. In the current study, class I and class II core ligand sequences were displayed on the surface of bacteriophage M13 with five random residues placed either N- or C-terminal of core ligand residues. These libraries were screened for binding to the Src, Fyn, Lyn, Yes, and PI3K SH3 domains. By this approach, additional ligand residue preferences were identified that can increase the affinity of SH3 peptide ligands at least 20-fold compared with core peptides. The amino acids selected in the flanking sequences were similar for Src, Fyn, and Yes SH3 domains; however, Lyn and PI3K SH3 domains showed distinct binding specificities. These results indicate that residues that flank the core binding sequences shared by many SH3 domains are important determinants of SH3 binding affinity and selectivity.**

One feature common to many proteins involved in signal transduction is the capacity to associate specifically with other molecules, forming a complex that can alter cell activity. Such multiprotein associations are facilitated by domain(s) of conserved sequence which serve as noncatalytic modules that direct molecular interactions during signal transduction (1, 2). One such domain, termed the Src homology 2 (SH2) domain, binds to phosphotyrosine-containing proteins in a sequence-specific manner (3). Another module, termed the SH3 domain, has a preference for proline-rich binding sites (4–9). In a number of cell types, both SH2 and SH3 domains function to regulate cellular events such as protein localization, enzyme activity, and substrate recruitment (1, 2).

SH2 and SH3 domains help determine the routes of intracellular communication (1, 2). The sequences that these domains recognize are important determinants of specific signal transduction pathways. Targeting these modules with pharmaceutical agents that mimic binding sites may provide a means to block specific events in the cell. To investigate SH3 binding specificity, SH3 domains have been used to screen cDNA libraries, phage display libraries, and biased peptide libraries (5–10). With these approaches, the minimal sequence requirement for SH3 domain recognition has been determined. SH3 binding sites contain a conserved Pro-Xaa-Xaa-Pro (PXXP) motif, and structural analysis indicates that the ligands adopt a left-handed polyproline type II helix conformation (5, 11).

To identify SH3 domain binding preferences, we previously screened bacteriophage that displayed the peptides XXXXX-XPPIP or RSLRPLXXXXXX (X, randomized residue) (6). These biased libraries were used to reveal SH3 domain binding preferences for a "window of recognition" that totaled 12 residues. The consensus binding sequence identified from the Src, Fyn, Lyn, and phosphatidylinositol 3-kinase (PI3K) SH3 domains is

X	X	X	**R**	**P**	**L**	**P**	**P**	Ψ	**P**	X	X
−6	−5	−4	−3	−2	−1	0	1	2	3	4	5

[with numbering as described by Yu *et al.* (5); Ψ, hydrophobic residue].

The SH3 domains all showed a preference for the core sequence RPLPPΨP. In addition, preferences for certain residues flanking the core were observed. For example, both the Lyn and PI3K SH3 domain selected arginine at position −6, and the PI3K SH3 domain preferred prolines at positions 4 and 5. These preferences suggest that there may be additional binding determinants which have not been identified.

The Src SH3 domain has also been used to screen a biased peptide library, XXXPPXPXX (5). Two types of ligands have been identified and have been termed class I and class II ligands. The class I ligands are similar to those identified by phage display selection and have the consensus sequence RXLPPLP. The class II ligands have the consensus sequence ΨPPLPXR. Since both types of ligands are proline-rich and contain a conserved arginine at the N or C terminus, Yu *et al.* (5) suggested that these two classes of peptides bind to the SH3 domain in reverse orientations. Structural analysis of SH3–ligand complexes has confirmed this hypothesis (12). SH3 ligands are pseudosymmetrical and can bind in two opposite orientations, termed class I ($NH_2 \rightarrow COOH$) or class II ($COOH \rightarrow NH_2$) binding sites, depending on the orientation in which they bind (5, 13).

To further explore the ligand residues important for specific SH3 domain recognition, we have prepared and screened new biased phage display libraries containing five random amino acid residues N- or C-terminal to a class I recognition sequence. In addition, we have made a library of bacteriophage which display five random amino acids C-terminal to a class II recognition sequence. We demonstrate that residues which flank the core SH3 binding sequences are important determinants of SH3 binding affinity and selectivity.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: SH3, Src homology 3; GST, glutathione *S*-transferase; PI3K, phosphatidylinositol 3-kinase.

*To whom reprint requests should be addressed.

MATERIALS AND METHODS

Reagents. HPLC-purified peptides were purchased from Quality Control Biochemicals (Hopkington, MA). Oligonucleotides were synthesized with an Applied Biosystems DNA/RNA synthesizer. The Src, Fyn, and Lyn glutathione *S*-transferase (GST)-SH3 expression vectors have been described (6). PI3K SH3 domain coding sequences (aa 3–83; ref. 14) were amplified by PCR and cloned into *Bam*HI/*Eco*RI-digested pGEX-2TK (Pharmacia). The chicken Yes SH3 domain (aa 96–177; ref. 15) in pGEX-3X was kindly provided by Marius Sudol. Fusion proteins were produced and purified as described (6).

Construction of Phage Display Libraries. The synthesis and purification of phage display libraries were as described (6). Three libraries were prepared. Peptides were inserted into M13 gene III sequences to produce chimeric proteins of the following forms:

N+5 class I NH_2-AEXXXXXRPLPPLPPPTVESCL

C+5 class I NH_2-AERSLRPLPPLPXXXXXTVESCL

C+5 class II NH_2-AEGAAPPLPPRXXXXXTVESCL

Underlined amino acids represent native gene III residues and X represents the randomized residues. The number of recombinants in each unamplified library was 7.5×10^7 (N+5 class I), 5.4×10^7 (C+5 class I), and 2.2×10^7 (C+5 class II).

The oligonucleotides used for library constructions were 5′-GTTTTGCTAAACAACTTTCAACAGTTGGGGGTGGCAATGGTGGCAAGGGGCG(MNN)$_5$TTCAGCGGAGTGAGAATAGAAAGG-3′ (N+5 class I library), 5′-GTTTTGCTAAACAACTTTCAACAGT(MNN)$_5$GGGCAATGGTGGCAAGGGGCGGAGTGAGCGTTCAGCGGAGTGAGAATAGAAAGG-3′ (C+5 class I library), and 5′-GTTTTGCTAAACAACTTTCAACAGT(MNN)$_5$GCGTGGTGGCAATGGTGGAGCTGCTCCTTCAGCGGAGTGAGAATAGAAAGG-3′ (C+5 class II library). Underlined nucleotides represent complementary M13 sequences; sequences inserted into gene III are not underlined. Nucleotides used to create random amino acid sequences are denoted N (A, G, C, T) and M (C, A). Phage enrichment was as described (6).

Binding Assays. BIAcore binding assays using phage-displayed ligands and immobilized GST-SH3 domains were as described (6). Tryptophan fluorescence measurements were performed to determine the affinity of peptides for the Src, Fyn, Lyn, Yes, and PI3K SH3 domains. Measurements were made with a Perkin–Elmer LS50B luminescence spectrophotometer. The excitation wavelength was 280 nm (5-nm slit width) and the emission wavelength was 335 nm (5-nm slit width). Aliquots of peptide were added to 3 ml of 0.5 μM GST-SH3 solution in phosphate-buffered saline (GIBCO) and the mixture was stirred in a cuvette for 5 min prior to analysis. Peptide was added to a final concentration of either 64 or 128 μM. The equilibrium dissociation constant (K_d) was calculated as described (16). Four peptides were used for tryptophan fluorescence measurements: class I core (Ac-RPLPPLPGGK-NH_2), class I core plus Src-selected flanking sequences (Ac-VSLARRPLPPLPGGK-NH_2), class II core (Ac-KGGGAAPPLPPR-NH_2), and class II core plus Src-selected flanking sequence (Ac-KGGGAAPPLPPRNRPRL-NH_2). GGK (class I) or KGG (class II) was added to the end of the peptides to allow immobilization of peptides to solid phase for BIAcore analysis or for biotinylation. The pair of glycines provides a spacer to minimize steric interference from a solid support. Similar K_d values were obtained when using GST-Src SH3 or free Src SH3 domain purified after cleaving the GST fusion protein with thrombin (data not shown).

RESULTS

Initially, two biased libraries were prepared to study additional SH3 binding preferences. Bacteriophage libraries were constructed that displayed a class I core and five adjacent N-terminal random residues (library N+5: XXXXXRPLPPLPPP; core residues are underlined or a class I core with five C-terminal random residues (library C+5: RSLRPLPPLPXXXXX). GST-SH3 fusion proteins were immobilized on polystyrene and binding selections were performed (6).

The N+5 library was screened with Src, Fyn, Lyn, Yes, and PI3K GST-SH3 fusion proteins. After four rounds of selection, bacteriophage isolates were sequenced. All five SH3 domains displayed preferences for specific residues amino terminal to the class I core (Fig. 1*A*). For example, the Src, Fyn, Lyn, and Yes SH3 domains selected a hydrophobic residue, most commonly leucine at the third position of the five positions randomized. There were also some preferences specific for particular SH3 domains. For example, the Src SH3 domain exhibited a preference for serine at the second position. In contrast, the PI3K SH3 domain selected a leucine or valine whereas the Lyn SH3 domain enriched for phage containing a tryptophan at the second position. These preferences were not observed when only three positions N-terminal to the core ligand sequence (RPLPPLP) were randomized. We infer that inclusion of two additional randomized positions in our N+5 library permits identification of the observed preferences.

Src, Fyn, Lyn, and PI3K GST-SH3 fusion proteins were also used to screen the C+5 library, which consisted of phage displaying five random residues C-terminal to the class I core. Some of the SH3 domains preferred specific residues C-terminal to the class I core (Fig. 1*B*). For example, most of the phage selected by the Lyn SH3 domain contained the residues LPPRH C-terminal to the class I core. In contrast, the PI3K SH3 domain preferred phage containing the C-terminal residues PRPXX. The Fyn SH3 domain had a preference for a threonine or serine immediately downstream of the class I core. GST-Src SH3 mainly selected phage from the RSLRPLPPLPXXXXX library that had nucleotide changes or deletions, resulting in mutations in the invariant portion (RSLRPLPPLP) of the displayed peptide.

Using surface plasmon resonance binding assays, we previously showed that SH3 binding could be studied when the ligands are displayed on the surface of the bacteriophage (6). As a preliminary analysis to determine whether the preferred sequences flanking the class I core contribute to SH3 domain recognition, the relative binding affinity of some of the selected phage was determined with a BIAcore biosensor (17). Src, Fyn, Lyn, or PI3K GST-SH3 fusion proteins were immobilized on a biosensor chip and the binding of homogeneous phage stocks displaying various SH3 ligands was assessed (Fig. 2). Initially, we analyzed phage previously isolated (6) which each display the class I core consensus sequence RPLPPLP with three distinct N-terminal amino acids (samples 1–3). All three phage bound with similar affinity to the Src, Fyn, Lyn, and PI3K SH3 domains. In contrast, variants that did not have a perfect class I consensus sequence (sample 4, LILPPLP; sample 5, RPLPPPP and sample 6, RPLPSLP, nonconsensus residues are underlined) bound with reduced or undetectable affinity. Bacteriophage containing SH3 domain-specific N-terminal residues adjacent to core ligand sequences bound selectively, with higher affinity for the SH3 domain that was used to select that particular sequence. For example, phage displaying the Src SH3-selected sequence **VPLARR**PLPPLPPP (sample 7, the N-terminal residues flanking the core are in boldface type) bound with highest affinity to the Src SH3 domain. Phage displaying the Lyn SH3-selected sequence **PWWLD**RPLPPLPPP (sample 8) bound with the highest affinity to the Lyn SH3 domain. The PI3K SH3-selected se-

A

xxx xxx xxx xxx xxx **arg pro leu pro pro leu pro pro pro**

GST-Src SH3	GST-Fyn SH3	GST-Lyn SH3	GST-Yes SH3	GST-PI3K SH3
trp **ser leu** phe his	ser ala pro **ser** arg	ser **trp leu** ile ala	pro asn **leu** arg thr	**cys leu** asn cys leu(4x)
ser trp **leu** ser arg(2x)	leu arg trp **ser** ala	arg ala **leu** pro pro	leu pro **leu** ser **arg**	phe lys arg pro pro
cys leu **leu** cys ser	lys ala **leu** ala ser	trp ser **ile** arg ser	**ser** trp trp ser **arg**	leu **leu** ser gly pro(3x)
ile his **leu** arg thr	thr thr **leu ser** ala	tyr pro phe his gly	**thr** phe **leu** phe **arg**	leu **leu** ser gly pro
pro his pro ala arg	ser tyr **leu ser** arg	pro **trp leu** leu arg	**thr** his pro **ala arg**	phe lys arg pro phe
val pro **leu** ala arg(2x)	leu ser pro met ser	ser **trp leu** ser val	phe pro **val ala** gln	leu **val** leu pro arg
val **ser leu** ala gln	leu pro met ala arg	ser **trp leu** gly ser	pro ser pro arg ser	phe his arg pro asn
ile **ser leu** ala arg	ala ser phe ala ser	pro **trp** trp leu asp	arg ser **leu** ser met	arg arg val pro pro
ser trp **leu** ala asn	val gly **leu ser** asn	leu thr met ala gly	val leu **leu ala** leu	trp ser arg thr lys
leu **ser leu** gln met	**ile** pro **ile** gly his	asp **phe** met ala ser	**ser** val **leu ala arg**	arg **val** arg pro ser
ser **ser leu** ser arg	ile ala **leu ser** thr	asp **trp** tyr his ser	asn arg ala **ala lys**	arg tyr leu pro trp
leu **ser leu** ala gln	phe phe **leu** ala ser	leu ser phe ser asn	**ser** ala **leu ala arg**	**cys** cys leu thr thr
leu **ser leu** phe arg	leu asp arg **ser** arg	met asp **leu** leu arg	**ser** ser **leu ala** gln	arg pro phe pro arg
ser **ser leu** arg phe(2x)	arg trp **ile** thr ser	ala ala met val gly	**thr** pro **leu** gly **arg**	arg **val** pro ile arg
ser trp **leu** ser his	gly leu gly ala arg	arg leu **leu** pro pro	pro asn pro ser tyr	ala arg leu pro pro
ser **ser leu** cys gly	asp asp **leu ser** gly	ser **trp val** ala cys	**ser** ala ser **ala arg**	**cys leu** asn cys phe(6x)
	asn pro **leu** leu arg	pro **trp** pro val gly	**ser** trp met **ala arg**	arg tyr leu ser trp
			ser ala **leu ala arg**	**cys leu** asn cys ser
			ser pro **leu ala arg**	**cys leu** leu thr thr(7x)
				cys leu asn cys ser(4x)

B

arg ser leu arg pro leu pro pro leu pro xxx xxx xxx xxx xxx

LYN	**leu pro pro arg** leu
	leu pro pro arg thr
	leu pro pro arg his(9x)
	leu pro pro pro **his**
PI3K	leu pro **pro** arg leu
	phe pro arg tyr leu
	pro ser ser ala pro
	pro arg ala pro phe
	pro arg pro pro phe
	ser **arg pro** gly pro
	pro arg pro ala tyr
FYN	leu pro pro leu thr
	ser ala pro arg val
	val leu ser glu pro
	thr ser thr gln pro
	his leu pro asp ser
	ser tyr thr ser his
	val ala thr his pro
	ser ser leu ser arg
	ala phe his ala arg
arg ser phe arg pro leu pro pro leu pro	**thr** val ala ser pro
SRC	
arg ser leu arg pro leu pro --- leu pro	pro val pro ser leu
arg ser leu arg pro leu pro pro --- pro	pro val pro ser leu(12x)
arg ser leu --- pro leu pro pro leu pro	ala arg pro his pro
arg ser phe arg pro leu pro pro leu pro	thr leu leu pro ser
arg ser leu arg pro leu pro pro leu pro	thr pro pro leu his

FIG. 1. Amino acid sequences of peptide inserts selected from biased phage libraries by GST-SH3 fusion proteins. (*A*) Sequences selected from the N+5 class I library by GST-Src SH3, GST-Fyn SH3, GST-Yes SH3, GST-Lyn SH3, and GST-PI3K SH3. The amino acids that are underlined were inserted into gene III adjacent to the five random residues. Amino acid preferences (at least 50% of the isolates) are highlighted in boldface type. Numbers in parentheses denote isolates that have the same nucleotide sequence. (*B*) Isolates selected from the C+5 class I library by GST-Src SH3, GST-Fyn SH3, GST-Yes SH3, GST-Lyn SH3, and GST-PI3K SH3. Dashes denote a 3-nt (1-aa) deletion of invariant residues.

quence **CLNCF**RPLPPLPPP bound with the highest affinity to the PI3K SH3 domain (sample 9).

Enhanced binding was also observed with phage displaying the class I core sequence and selected downstream residues. For example, the Lyn SH3-selected phage displaying RSLRPLPPLP**LPPRT** (sample 10) and RSLRPLPPLP**LPPRH** (sample 11) bound selectively and with the highest affinity to Lyn SH3 domain. Phage displaying the PI3K SH3-selected sequence RSLRPLPPLP**PRPPF** (sample 12) selectively bound to the PI3K SH3 domain. In contrast, two Src SH3-selected phage, both displaying peptides with deletions in the core consensus sequence, RSLRPLPP(−)**PPVPSL** (sample 13) and RSLRPLP(−)LP**PVPSL** (sample 14), did not bind much better to the Src SH3 domain than phage that displayed the class I core ligand sequence (samples 1–3).

Overall, the SH3 domain-specific phage that were selected from the N+5 class I library bound with higher affinity than SH3 domain-specific phage selected from the C+5 library. These results indicate that the sequences which flank the SH3 ligand core are important for binding and that the sequences adjacent to the class I core arginine residue may be the more important determinants of high-affinity binding. To determine whether similar preferences would be observed when the peptide ligand was bound in the opposite orientation, we prepared a biased library where residues adjacent to the arginine of a class II core were randomized. This library, GAAPPLPPRXXXXX, was screened using Src, Fyn, Lyn, Yes, and PI3K GST-SH3 domain fusion proteins. Each SH3 domain had preferences for specific downstream residues (Fig. 3). The C-terminal consensus sequences derived from phage selected from the GAAPPLPPRXXXXX library were N(R/K)-XR(L/V) (Src), NPXRΨ (Fyn), NPPRΨ (Yes), PX(W/F)MX (Lyn), and PPRPX (PI3K).

Peptide ligands consisting of core sequences or core plus Src-selected flanking sequences were synthesized and analyzed to evaluate more quantitatively the effects of flanking sequences on SH3 binding affinity and specificity. Tryptophan fluorescence binding assays were performed with the Src, Fyn, Lyn, Yes, and PI3K SH3 domains and the peptide ligands. A peptide containing class I core residues and the Src SH3-selected flanking sequence (**VSLAR**RPLPPLPPP) bound at least 20-fold better to the Src and Fyn SH3 domain than did

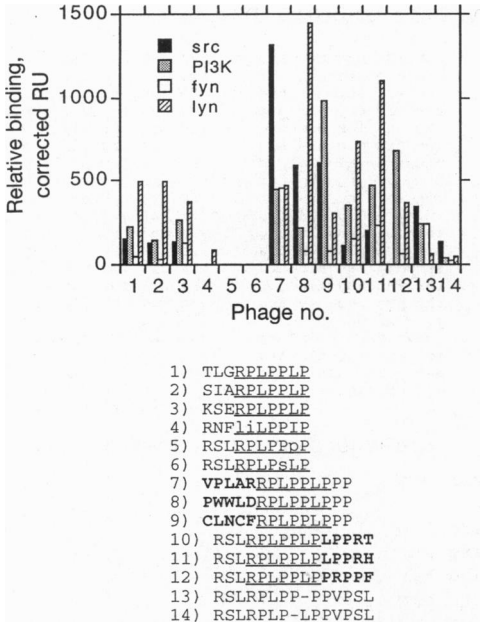

FIG. 2. Binding of phage-displayed peptides to Src, Fyn, Lyn, and PI3K GST-SH3 fusion proteins. Binding of phage-displayed peptides was measured by BIAcore technology (6). The amount of phage binding to immobilized SH3 domain is given as arbitrary resonance units (RU) and has been corrected for bulk refractive-index contributions of the unbound phage. Phage-displayed sequences are shown. Core ligand sequences are underlined. Core ligand residues that do not match the RPLPPΨP consensus are not capitalized. Selected flanking sequences are in boldface type.

a peptide containing just the core recognition sequence (Table 1). The flanking sequences had an even larger effect (≈75-fold increase) on Yes SH3 domain binding affinity. In contrast the Src SH3-selected flanking sequences only marginally affected (2- to 3-fold) the affinity of the interaction of the core ligand residues with the PI3K or the Lyn SH3 domain.

Results similar to those obtained with class I peptides were observed when a class II core peptide was compared with one that had the additional C-terminal residues (NRPRL) identified by using the Src SH3 domain (Table 1). A peptide consisting of class II core residues and the Src SH3-selected flanking sequences bound ≈80-fold better to the Src SH3 domain than did a peptide containing just the class II core recognition sequence. The peptide containing the C-terminal selected sequences also bound with much higher affinity to the Fyn and Yes SH3 domains. In contrast, binding to the Lyn SH3 domain was only marginally (≈3-fold) increased.

DISCUSSION

Phage display libraries were previously utilized to define the minimal ligand recognition sequences of the Src, Fyn, Lyn, PI3K, and Abl SH3 domains (6). Src, Fyn, Lyn, and PI3K SH3 domains selectively enriched phage displaying the core sequence RPLPPΨP. In contrast, the Abl SH3 domain preferred a distinct core motif, PPPYPPPPΨP. Now, using biased phage display libraries to search for expanded motifs that improve both affinity and selectivity, we have shown that residues which flank the core recognition sequence are important binding determinants. These affinity/specificity determinants can be N- or C-terminal to a SH3 domain core recognition sequence and be SH3 domain-specific. Further, the flanking sequences may selectively affect SH3 binding.

In our studies, SH3 domains were not presented to ligands in an environment of their natural flanking sequences, but rather as GST fusion proteins. The observation that residues flanking class I or class II core sequences are important for SH3 domain recognition does not appear to be due to the creation of novel artificial binding surfaces by preparing SH3 domains as GST fusion proteins, since a protein consisting of the Lyn unique region-SH3-SH2 showed binding preferences similar to those observed for Lyn GST-SH3 (data not shown).

The selection of SH3 domain-specific sequences both N-terminal and C-terminal to a class I core ligand suggests that SH3 domains may have additional binding surfaces for ligand interaction. Alternatively, the SH3-selected flanking residues could be important for phage growth and/or core ligand presentation as displayed on the surface of the bacteriophage. By BIAcore binding analysis, we have shown that the flanking sequences affect SH3 binding, exerting their effect independent of phage growth. Furthermore, using peptides, we have demonstrated that residues N-terminal (class I) and C-terminal (class II) to core sequences are important for high-affinity interaction with SH3 domains, strongly suggesting a direct interaction. The significance of the residues selected C-terminal to class I cores is not known. These sequences may make contact with the SH3 domain or perhaps be important for ligand presentation. We also do not know why the Src and Yes SH3 domains selected phage displaying peptides with deletions in the core consensus sequence. However, we do note that the PI3K SH3 domain may have selected phage displaying

gly ala ala pro pro leu pro pro arg xxx xxx xxx xxx xxx

GST-Src SH3	GST-Fyn SH3	GST-Lyn SH3	GST-Yes SH3	GST-PI3K SH3
xxx xxx xxx xxx xxx	xxx xxx xxx xxx xxx	xxx xxx xxx xxx xxx	xxx xxx xxx xxx xxx	xxx xxx xxx xxx xxx
asn pro pro **arg leu**	**asn pro** pro **arg val**	thr tyr gly thr thr	asn arg **pro arg val**	**pro pro** leu arg ser
asn lys val **arg** gly	**asn pro** arg pro **ile**	**pro** pro **trp met** met	**asn pro pro arg ile**(2x)	**pro pro** ile arg lys
asn gln thr phe arg	**asn** arg val **arg leu**	thr arg tyr ser leu	**asn pro** lys his arg	**pro pro arg pro** leu(2x)
asn pro pro **arg** thr	**asn** ala val **arg leu**	**pro** pro **trp met** thr	**asn pro pro arg** ser	**pro pro** leu **pro** phe
asn arg leu his arg	**asn** arg pro his phe	thr tyr leu leu pro	**asn pro** lys his trp	**pro pro lys pro** ala
asn val pro **arg leu**	**asn pro** lys his arg	thr tyr **trp met** pro	**asn pro** val **arg val**	**pro pro arg pro** pro
asn arg leu his ala	**asn** arg val **arg leu**	**pro** thr **phe met** ala	**asn** lys ala **arg leu**	**pro pro arg pro** ala(2x)
asn arg val **lys leu**	**asn pro** pro **arg** phe	**pro** thr **trp met** thr	**asn** lys his pro pro	**pro pro** leu arg gln
asn arg val **arg leu**	**asn** ser arg pro **leu**	thr tyr leu **met** pro	**asn pro pro arg** lys	arg ala leu arg leu
asn arg val val **leu**	**asn pro** ala **arg leu**	**pro** ala **trp met** ala	**asn** lys arg his **leu**	arg leu **arg** val gly
asn lys pro **arg leu**	**asn** thr pro **arg leu**	**ser** tyr leu leu pro	lys his leu ser ser	**pro pro arg** lys pro
asn ile pro **arg val**	**asn pro** arg pro **leu**	thr tyr leu leu pro	**asn pro pro arg** thr(2x)	**pro pro arg** thr pro
asn pro pro **arg** lys	**asn pro** pro **arg** thr	thr thr pro tyr thr	**asn pro pro arg val**	**pro pro arg pro** ser
asn lys pro **arg leu**	**asn pro** thr pro **val**	**pro** ser **trp met** ser	**asn** lys leu **arg val**	**pro pro arg pro** phe
asn pro pro **arg** thr		**pro** val thr trp ile		
asn lys ala **arg leu**		**pro** ser **phe met** ser		
asn his leu ser thr		**pro** ala **phe** phe ser		
		pro ser **phe** leu gln		

FIG. 3. Amino acid sequence of peptide inserts selected from the C+5 class II biased phage library by GST-SH3 fusion proteins. Phage selections used GST-Src SH3, GST-Fyn SH3, GST-Yes SH3, GST-Lyn SH3, and GST-PI3K SH3. Amino acids underlined were inserted into gene III adjacent to the five random residues. Amino acid preferences are in bold type. Numbers in parentheses denote isolates of the same nucleotide sequence.

Table 1. Binding of class I and class II peptides to Src, Fyn, Lyn, Yes, and PI3K GST-SH3 fusion proteins

SH3 domain	Dissociation constant (K_d), μM			
	Class I core (RPLPPLPggk)	Class I core + "Src flank" (**VSLAR**RPLPPLPggk)	Class II core (kgggaaPPLPPR)	Class II core + "Src flank" (kgggaaPPLPPR**NRPRL**)
Src	17.7 ± 3.3	0.86 ± 0.04	20.3 ± 0.8	0.24 ± 0.04
Fyn	25.5 ± 5.5	0.60 ± 0.10	144.4 ± 19.0	1.7 ± 0.15
Lyn	6.2 ± 0.5	2.8 ± 0.66	26.5 ± 3.3	8.1 ± 1.50
Yes	52.3 ± 10.1	0.67 ± 0.03	NDB	3.8 ± 0.03
PI3K	38.8 ± 4.5	14.9 ± 0.64	NDB	NDB

The affinity of each peptide for the SH3 domains was determined by tryptophan fluorescence. Shown are the mean ± SEM for the class I binding assays (performed in triplicate) and the class II binding assays (performed in duplicate). Ligand binding measurements were performed with three (class I ligands) or two (class II ligands) different preparations of GST fusion protein. Core ligand sequences are in uppercase lightface type; those in boldface type are the "flanking sequences" selected by the Src SH3 domain. NDB, no detectable change in fluorescence at 128 μM peptide.

the sequence RSLRPLPPLPPRPXX because such ligands contain both class I (underlined) and class II (boldface type) consensus sequences and might potentially bind in either orientation. This feature may be important for their selection.

The solution structures of both class I and class II core recognition sequences complexed with the Src SH3 domain have been determined by NMR spectroscopy (5, 12). Core peptide residues fit into three binding pockets on the SH3 domain. The ligand arginine residue forms a salt bridge with residue D99 of Src SH3 (pocket 1) while the rest of the peptide contacts two hydrophobic binding pockets formed primarily by Src SH3 residues Y90, Y92, W118, Y131, P133, and Y136. These residues are conserved in the Fyn, Lyn, Yes, and PI3K SH3 domains and thus are not the source of the observed binding differences. Our data suggest that there is an additional contact site(s) for ligand binding. Given the preferred flanking sequences of the SH3 domains analyzed, we predict that the residues which constitute this new binding surface will be mostly conserved among the Src, Fyn, and Yes SH3 domains and divergent with the Lyn and PI3K SH3 domains.

The residues that influence SH3 core ligand binding are distinct and depend on the ligand binding orientation. For example, the Src SH3 domain selects the residues XSLXX N-terminal to class I core sequences (**XSLXX**RPLPPLP) and the residues N(R/K)XR(L/V) C-terminal to a class II core [PPLPPRN**(R/K)XR(L/V)**]. The same binding surface is reportedly utilized by the Src SH3 domain for both class I and class II core ligand residues (12). In each case, the arginine of the core ligand (underlined) forms a salt bridge with D99 of the Src SH3. Presumably, additional SH3 residues participate in ligand binding. The observation that the class I and class II flanking sequences are distinct suggests either the use of two different binding surfaces or an alternative mode of binding to the same surface. Mutational analysis of ligand flanking sequences coupled with the determination of extended class I and class II SH3 structures will resolve these issues. Preliminary NMR studies of the extended peptide ligands (class I and class II) bound to the Src SH3 domain identify specific contacts between the flanking sequences and the SH3 domain (S. Feng, R.J.R., and S. Schreiber, unpublished work).

SH3 domains have been used to affinity isolate SH3-binding proteins from cellular extracts. We previously observed that while the Src, Fyn, and Lyn SH3 domains recognized many of the same cellular proteins, numerous distinct binding preferences were observed (figure 2 of ref. 18). Our results implied that the Src, Fyn, and Lyn SH3 domains all recognize a common motif and, in addition, that there might be SH3 domain-specific binding sites. The results presented here give a molecular basis for the SH3 domain selectivity observed for these Src tyrosine kinase family members. A protein that contains a poor SH3 domain core consensus sequence would bind with low affinity to the class of SH3 domains which recognize that particular core sequence. If SH3 domain-specific preferred residues flank a weak binding site, binding affinity would be selectively enhanced, depending on which residues flank the core binding site.

By using biased phage display libraries, we have identified SH3 domain ligand residues that are important for both SH3 domain binding affinity and selectivity. Some SH3 domains, such as the Src, Fyn, Lyn, Yes, and PI3K, have similar core binding motifs. The identification of ligand residues that are SH3 domain-specific provides valuable information for the molecular understanding of specificity, which can be used for the design of SH3 domain-specific inhibitory molecules. Such compounds may provide a means to modulate both physiologic and pathological processes.

We thank Stuart Schreiber, Sibo Feng, Tony Pawson, and our colleagues at ARIAD for helpful suggestions during the course of this work. This work was funded in part by Grant CA27951 from the National Cancer Institute (J.S.B.) and Small Business Innovation Research Grant 2 R44 AR42112-02 from the National Institutes of Health (M.J.Z.).

1. Cohen, G. B., Ren, R. & Baltimore, D. (1995) *Cell* **80,** 237–248.
2. Pawson, T. (1995) *Nature (London)* **373,** 573–580.
3. Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J., Neel, B. G., Birge, R. B., Fajardo, J. E., Chou, M. M., Hanafusa, H., Schaffhausen, B. & Cantley, L. C. (1993) *Cell* **72,** 767–778.
4. Ren, R., Mayer, B. J., Cicchetti, P. & Baltimore, D. (1993) *Science* **259,** 1157–1161.
5. Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W. & Schreiber, S. L. (1994) *Cell* **76,** 933–945.
6. Rickles, R. J., Botfield, M. C., Weng, Z., Taylor, J. A., Green, O. M., Brugge, J. S. & Zoller, M. J. (1994) *EMBO J.* **13,** 5598–5604.
7. Sparks, A. B., Quilliam, L. A., Thorn, J. M., Der, C. J. & Kay, B. K. (1994) *J. Biol. Chem.* **269,** 23853–23856.
8. Cheadle, C., Ivashchenko, Y., South, V., Searfoss, G. H., French, S., Howk, R., Ricca, G. A. & Jaye, M. (1994) *J. Biol. Chem.* **269,** 24034–24039.
9. Alexandropoulos, K., Cheng, G. & Baltimore, D. (1995) *Proc. Natl. Acad. Sci.* **92,** 3110–3114.
10. Cicchetti, P., Mayer, B. J., Thiel, G. & Baltimore, D. (1992) *Science* **257,** 803–806.
11. Lim, W. A. & Richards, F. M. (1994) *Struct. Biol.* **1,** 221–225.
12. Feng, S., Chen, J. K., Yu, H., Simon, J. A. & Schreiber, S. L. (1994) *Science* **266,** 1241–1247.
13. Lim, W. A., Richards, F. M. & Fox, R. O. (1994) *Nature (London)* **372,** 375–379.
14. Skolnik, E. Y., Margolis, B., Mohammadi, M., Lowenstein, E., Fischer, R., Drepps, A., Ullrich, A. & Schlessinger, J. (1991) *Cell* **65,** 83–90.
15. Sudol, M., Kieswetter, C., Zhao, Y.-H., Dorai, T., Wang, L.-H. & Hanafusa, H. (1988) *Nucleic Acids Res.* **16,** 9876–9883.
16. Chen, J. K., Lane, W. S., Brauer, A., Tanaka, A. & Schreiber, S. L. (1993) *J. Am. Chem. Soc.* **115,** 12591–12592.
17. Malmqvist, M. (1993) *Nature (London)* **361,** 186–187.
18. Weng, Z., Thomas, S. M., Rickles, R. J., Taylor, J. A., Brauer, A. W., Seidel, D. C., Michael, W. M., Dreyfuss, G. & Brugge, J. S. (1994) *Mol. Cell. Biol.* **14,** 4509–4521.