The EMBO Journal vol.13 no.23 pp.5598–5604, 1994

Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries

Richard J.Rickles[2], Martyn C.Botfield, Zhigang Weng[1], Jennifer A.Taylor, O.M.Green, Joan S.Brugge and Mark J.Zoller

ARIAD Pharmaceuticals, 26 Landsdowne Street, Cambridge, MA 02139 and [1]Graduate Group in Molecular Biology, Biomedical Graduate Studies, the University of Pennsylvania, School of Medicine, Philadelphia, PA 19104, USA
[2]Corresponding author

Communicated by S.A.Courtneidge

Many proteins involved in intracellular signal transduction contain a small, 50–60 amino acid domain, termed the Src homology 3 (SH3) domain. This domain appears to mediate critical protein–protein interactions that are involved in responses to extracellular signals. Previous studies have shown that the SH3 domains from several proteins recognize short, contiguous amino acid sequences that are rich in proline residues. While all SH3 recognition sequences identified to date share a conserved P-X-X-P motif, the sequence recognition specificity of individual SH3 domains is poorly understood. We have employed a novel modification of phage display involving biased libraries to identify peptide ligands of the Src, Fyn, Lyn, PI3K and Abl SH3 domains. With biased libraries, we probed SH3 recognition over a 12 amino acid window. The Src SH3 domain prefers the sequence XXX-RPLPPLPXP, Fyn prefers XXXRPLPP(I/L)PXX, Lyn prefers RXXRPLPPLPXP, PI3K prefers RXXRPL-PPLPPP while the Abl SH3 domain selects phage containing the sequence PPPYPPPP(I/V)PXX. We have also analysed the binding properties of Abl and Src SH3 ligands. We find that although the phage-displayed Abl and Src SH3 ligands are proline rich, they are distinct. In surface plasmon resonance binding assays, these SH3 domains displayed highly selective binding to their cognate ligands when the sequences were displayed on the surface of the phage or as synthetic peptides. The selection of these high affinity SH3 peptide ligands provides valuable information on the recognition motifs of SH3 domains, serve as new tools to interfere with the cellular functions of SH3 domain-mediated processes and form the basis for the design of SH3-specific inhibitors of disease pathways.

Key words: peptide ligands/phage display library/recognition motif/SH3 domains

Introduction

Numerous proteins involved in the regulation of cellular events that occur in response to external stimuli contain domain(s) of conserved sequence. Two such domains, termed SH2 and SH3, serve as non-catalytic modules that regulate cellular processes by the formation and dissociation of protein–protein interactions (Pawson, 1988; Koch *et al.*, 1991; Pawson and Gish, 1992; Mayer and Baltimore, 1993). SH2 domains associate with specific proteins via sequences containing phosphotyrosine (Mayer *et al.*, 1991; Songyang *et al.*, 1993), while SH3 domains appear to have preference for proline-rich motifs (Chen *et al.*, 1993; Ren *et al.*, 1993; Yu *et al.*, 1994).

Inroads toward an understanding of SH3 ligand specificity were made with the screen of a cDNA expression library using the c-Abl SH3 domain as a probe. Two proteins were identified, termed 3BP-1 and 3BP-2 (Cicchetti *et al.*, 1992). Short proline-rich stretches within these proteins, 10 amino acids in length, were defined as the minimal sequence requirement for c-Abl SH3 binding (Ren *et al.*, 1993). The proline-rich binding sites in the Abl SH3-selected proteins 3BP-1 and 3BP-2 provided the first information on SH3 domain ligand recognition. The Src SH3 domain binds to both 3BP-1 and 3BP-2; however, the Src SH3 domain binds less efficiently than the Abl SH3 domain to these proteins, suggesting that the Src SH3 domain may have a different sequence preference (Cicchetti *et al.*, 1992; Ren *et al.*, 1993). Other proteins have been identified that bind to SH3 domains and all contain proline-rich motifs (Egan *et al.*, 1993; Flynn *et al.*, 1993; Gout *et al.*, 1993; Liu *et al.*, 1993; Olivier *et al.*, 1993; Prasad *et al.*, 1993; Rozakas-Adcock *et al.*, 1993; Vogel and Fujita, 1993; Weng *et al.*, 1993, 1994; Fumagalli *et al.*, 1994; Kapeller *et al.*, 1994; Pleiman *et al.*, 1994; Taylor and Shalloway, 1994). However, the residues in these proteins that are important for SH3 recognition have not been clearly defined.

To define SH3 ligand specificity, we have employed phage display selection using bacteriophage M13 (Smith, 1985; Parmley and Smith, 1988). Peptides were displayed on the surface of the bacteriophage, inserted into phage gene III sequences. Phage display libraries were prepared and used to identify peptides that bind to the SH3 domains of the proteins c-Src, c-Fyn, Lyn, PI3K and c-Abl.

Results

The first phage display library contained an insert consisting of six residues of random amino acid sequence, flanked on the N and C terminus by glycine (-GX_6G-). GST and GST–Src SH3 were each immobilized on polystyrene and binding selection performed. After six cycles of selection, GST-Src SH3-dependent phage enrichment was observed; the phage titer was 100-fold higher with GST–Src SH3 than with GST. After eight cycles (~1000-fold enrichment), 10 bacteriophage were isolated and sequenced. Each phage had the same nucleotide sequence, encoding a proline-rich peptide insert. In

Table I. Binding of phage-displayed peptides to GST–hSrc SH3

Phage-displayed sequences	cRU
SIARPLPPIP	133
KSERPLPPIP	129
SHRHPLPPIP	<1
AKRAPLPPIP	<1

Binding of phage-displayed peptides was measured as described under Materials and methods. The amount of bound phage is given as arbitrary resonance units that have been corrected for bulk refractive index contributions of the unbound phage (cRU).

addition, a point mutation in a flanking codon (GGG→A-GG) had occurred, converting glycine to arginine:

--- G X X X X X X G --- starting library
--- **R** S L P P I P G --- GST–Src SH3 selected

The isolation of only one Src SH3 ligand, displayed by phage containing a point mutation, suggests that Src SH3 ligands were not present in the original library and that one was generated during binding selection. The point mutation probably occurred during the amplification of phage and was necessary for Src SH3 recognition. This result suggested that the Src SH3 recognition sequence is proline rich, at least seven residues in length and perhaps prefers an arginine residue as part of the recognition sequence.

The Src SH3 phage-selected peptide has some similarity to the proline-rich SH3 recognition sequences in 3BP-1, 3BP-2 and other proteins (see below). Therefore, with the assumption that proline-rich sequences are important for Src SH3 recognition, a second *biased* library was prepared, consisting of six random amino acid residues flanking the tetrapeptide proline-rich sequence identified in the first screen (-X_6PPIPG-). After only one cycle of selection, GST–Src SH3 retained 10-fold more phage than GST suggesting the enrichment of certain phage. After three cycles of binding selection (~1000-fold enrichment) phage were isolated and sequenced. The results are shown in Figure 1A.

No amino acid preferences were observed at positions 1, 2 and 3 of the six randomized residues. In contrast, at position 4, arginine was preferred; at position 5, proline; and at position 6, leucine. That arginine is preferred at the fourth position strengthens the notion that the point mutation observed in the isolate from the first library (changing glycine to arginine) was critical for Src SH3 selection. In addition, we noted that the inclusion of the amino acids PPIP was important for Src SH3 recognition as phage enrichment was observed after only one round of panning.

The X_6PPIPG library was next screened using Fyn, Lyn, PI3K and Abl GST–SH3 fusion proteins. Each of these proteins enriched phage from the X_6PPIPG library. The results from sequencing selected phage are shown in Figure 1A. Like the Src SH3 domain, the Fyn, Lyn and PI3K SH3 domains prefer ligands that contain arginine at position 4, proline at position 5 and leucine at position 6. Again, no preferences were observed at positions 2 and 3. However, unlike Src and Fyn, the Lyn and PI3K SH3 domains show a preference for arginine at the first position of the six residues randomized. The Abl SH3 domain, in contrast, selected a distinct group of ligands. The Abl SH3 domain showed preferences for peptides containing prolines at the first three positions of the six randomized residues. Furthermore, tyrosine and proline were selected at the 4th and 6th position respectively, while the Src, Fyn, Lyn and PI3K SH3 domains select arginine at the 4th and leucine at the 6th position.

To investigate the sequence preferences in the 'PPIP' region of the SH3 ligands, two new biased libraries were generated. The library RSLRPLX_6 was prepared to identify additional Src, Fyn, Lyn and PI3K SH3 ligands. The preferences described above determined the placement of residues at the 1st, 4th, 5th and 6th positions. The library PPPYPPX_6 was prepared and screened to further define Abl SH3 binding preferences.

Bacteriophage were selected from these new libraries and sequenced. As shown in Figure 1B, the proline motif found originally by screening the GX_6G library using the Src SH3 domain was again selected, this time by all five SH3 domains. The C-terminal consensus sequences derived from the RSLPRLX_6 library are PPLPXP (Src), PP(I/L)PXX (Fyn), PPLPXP (Lyn), PPLPPP (PI3K). When using the PPPYPPX_6 library, the Abl SH3 domain prefers peptides that contain the C-terminal residues PP(I/V)PXX.

To further chacterize the peptide sequences selected from the phage display libraries, we have analyzed the binding properties of peptides selected by Src and Abl, which represent two different 'classes' of peptides. Figure 1C compares the consensus for Src- and Abl-selected peptides with the proline-rich sequences from 3BP-1 and 3BP-2. The Abl SH3 ligand consensus sequence identified by phage display is similar to the 3BP-1 core sequence and almost identical to those of 3BP-2. In contrast, the Src SH3-selected ligands share only the PXPPψP proline motif (ψ=hydrophobic residue). Several Src SH3-binding proteins have been identified recently (Flynn *et al.*, 1993; Liu *et al.*, 1993; Weng *et al.*, 1993, 1994; Fumagalli *et al.*, 1994; Taylor and Shalloway, 1994). These proteins contain proline-rich motifs, however the residues that actually constitute the SH3 recognition sites have not been defined precisely.

To evaluate rapidly the binding properties of the phage-displayed ligands, and to gain insight into SH3 domain selectivity, the relative binding affinity of phage isolates was measured by a novel application of surface plasmon resonance. The binding of phage-displayed ligands to immobilized GST–SH3 proteins was analysed using a BIAcore Biosensor®. Initially, the relative binding of four phage to GST–Src SH3 was assessed. Binding was observed using two different bacteriophage that contained the Src SH3 ligand consensus sequence RPLPPIP (Table I, samples 1 and 2). However, two variants that lacked arginine at the first position of the consensus failed to bind (Table I, samples 3 and 4), consistent with arginine being important for Src SH3 domain recognition. To obtain preliminary evidence as to whether the Abl- and Src-selected phage showed selective binding for each of these SH3 domains, relative binding affinity was measured. As shown in Figure 2, binding of phage containing the Src SH3 selected consensus sequence was detected readily using GST–Src SH3 but not using GST–Abl SH3. Conversely, phage displaying the Abl SH3-selected sequence PPPYPPPPIP bound to GST–Abl SH3 but

A

ala glu GLY **xxx xxx xxx xxx xxx xxx** GLY *thr val* ---

	arg ser **leu** pro pro ile pro	1st library isolate
xxx xxx xxx xxx xxx xxx	**pro pro ile pro**	X_6PPIP library
ser leu ala **arg pro leu**		
thr ser met **arg pro leu**		
lys ser glu **arg pro leu**		
thr leu gly **arg pro leu**		
ser ile ala **arg pro leu**		
ala pro arg ile **pro leu**		
leu his arg **arg** ala **leu**		
tyr asn his **arg** ser **leu**		
leu arg gln **arg pro leu**		SRC
pro ala ser **arg pro leu**		
ala arg asp **arg pro leu**		
pro arg ser **arg pro leu**		
phe val ser **arg pro leu**		
asn lys gly **arg** ser **leu**		
asp arg leu **arg pro leu**		
leu ala asn **arg** glu **leu**		
pro thr arg **arg pro leu**		
ile gln his **arg** leu **leu**		
arg leu leu **lys pro leu**		
gln ser arg **arg** ser **leu**		FYN
ala lys arg ala **pro leu**		
arg leu leu **lys pro leu**		
thr gly gly **arg pro leu**		
his his ile **arg pro leu**		
ser tyr pro **arg pro leu**		
his ser thr **arg** ala **leu**		
pro thr his **arg** ile **leu**		
arg asn phe leu ile **leu**		
arg val ser tyr asn **leu**		
pro val leu his ser **leu**		
arg tyr leu **arg pro leu**		
arg ser ser **arg pro leu**		LYN
val asn lys **arg pro leu**		
arg ser asn **arg pro leu**		
arg ala lys **arg pro leu**		
ser his arg his **pro leu**		
ala ser thr **arg pro leu**		
ala ser thr **arg pro leu**		
arg leu phe **arg pro leu**		
arg pro gln **arg pro leu**		PI3K
arg phe lys **arg pro leu**		
arg ala lys **arg pro leu**		
arg pro tyr **arg pro leu**		
arg ile pro **arg pro leu**		
arg ser leu **arg pro leu**		
pro pro pro tyr pro pro		
pro pro pro tyr pro pro	(4x)	
ala **pro pro tyr pro pro**	(2x)	
ala **pro** his **tyr pro pro**		
pro pro pro tyr his **pro**		
pro pro ser **tyr pro pro**		ABL
pro pro ala **tyr pro pro**		
pro pro ala **tyr pro pro**		
ala **pro** asn **tyr pro pro**		
ala **pro** ser **tyr** ser **pro**		
ala **pro** ser **tyr pro pro**	(2x)	
pro pro his **tyr pro pro**		

B

ala glu	**arg ser leu arg pro leu**	**xxx xxx xxx xxx xxx xxx** ---
		pro pro leu pro gly ser
		pro pro leu pro arg pro
		pro pro leu pro pro pro
		pro pro leu pro gly ser
	SRC	**pro pro leu pro pro pro**
		pro pro leu pro pro pro
		pro pro leu pro ser pro
		pro pro val pro pro pro
		pro pro leu pro asn gly
		pro pro leu pro phe asp
		pro pro pro pro phe pro
		pro pro leu pro val pro
		pro pro leu pro trp --- thr*
		pro pro leu pro pro leu
	FYN	**pro pro ile pro gly arg**
		pro pro ile pro pro pro
		pro pro leu pro ala pro
		pro pro ile pro leu gly
		pro pro ile pro asp ser
		pro pro leu pro ala pro
		pro pro leu pro ser pro
		pro pro leu pro leu pro
		pro pro leu pro leu arg
		pro pro leu pro leu arg
	LYN	**pro ser leu pro phe pro**
		pro pro leu pro pro pro
		pro pro leu pro leu leu
		pro pro leu pro gln leu
		pro pro leu pro ser ala
		met val ser leu val **pro**
		leu pro phe gly **pro pro**
		pro pro thr **pro** val-glu ser**
	PI3K	**pro pro leu pro pro pro**
		pro pro leu pro pro pro
		pro pro leu pro pro pro
		pro pro leu pro pro pro
		pro pro leu pro pro pro
ala glu	**pro pro pro tyr pro pro**	**xxx xxx xxx xxx xxx xxx** ---
		pro pro ile pro val arg
		pro pro val pro tyr ala
		pro pro ile pro ile leu
		pro pro val pro tyr pro
	ABL	**pro pro val pro arg ser**
		pro pro val pro gly ala
		pro pro ile pro arg phe
		pro asp ile pro leu leu

C

Src SH3	xxx xxx xxx arg PRO leu PRO PRO LEU PRO xxx pro
Abl SH3	pro pro pro tyr PRO pro PRO PRO ILE/VAL PRO xxx xxx
3BP-1	ala **pro** thr met **pro pro pro leu pro** pro
3BP-2	**pro pro** ala **tyr pro pro pro pro val pro**

D

Src SH3 Selected Peptides

Phage Display	xxx xxx xxx ARG pro LEU PRO PRO LEU PRO xxx pro
Peptides on Beads	ARG xxx LEU PRO PRO LEU PRO arg Ψ

PI3K SH3 Selected Peptides

Phage Display	arg xxx xxx ARG pro LEU PRO PRO leu PRO pro pro
Peptides on Beads	ARG xxx LEU PRO PRO arg PRO xxx xxx

Fig. 1. Amino acid sequence of peptide inserts selected from phage libraries using GST−SH3 fusion proteins. **(A)** Sequences selected using GST−Src SH3, GST−Fyn SH3, GST−Lyn SH3, GST−PI3K SH3 and GST−Abl SH3. The N-terminal sequences of gene III with the peptide insert of the GX_6G library is shown. The isolate from this library containing the glycine to arginine change is also shown. Amino acids that are in italics are gene III residues. Isolates selected from the X_6PPIPG are listed. The amino acids that are underlined were inserted into gene III adjacent to the six randomized residues. Amino acid preferences are highlighted in bold type. Numbers in parentheses denote isolates that have the same nucleotide sequence. **(B)** Isolates selected from the $RSLRPLX_6$ and $PPPYPPX_6$ libraries are shown. Amino acid preferences are highlighted in bold type. *Isolate has a three nucleotide deletion in the insert. **Isolate has a nine nucleotide deletion (three insert and six gIII nucleotides). **(C)** The Src and Abl SH3 domain ligand consensus sequences identified by phage display and the sequences of 3BP-1 and 3BP-2 recognized by the Abl SH3 domain (Ren *et al.*, 1993). Amino acids in capital letters represent conserved residues in the Src and Abl SH3 binding sequence, identified by phage display. Residues in 3BP-1 and 3BP-2 that match the Abl SH3 phage display sequence are in bold type. **(D)** The Src and PI3K SH3 domain ligand consensus sequences as identified by phage display and using a biased peptide bead library, XXXPPXPXX (Yu *et al.*, 1994). Amino acids in upper case represent shared consensus sequence residues identified by both methods of ligand identification. Ψ represents hydrophobic amino acids. Binding measurements (tryptophan fluorescence, Chen *et al.*, 1993) using the Src SH3 domain and a peptide identified by phage display (SIARPLPPIP) or one identified using a biased peptide bead library (RALPPLPRY, Yu *et al.*, 1994) are similar, K_d = 8.8 and 7.8 μm respectively (Chen and Schreiber, personal communication).

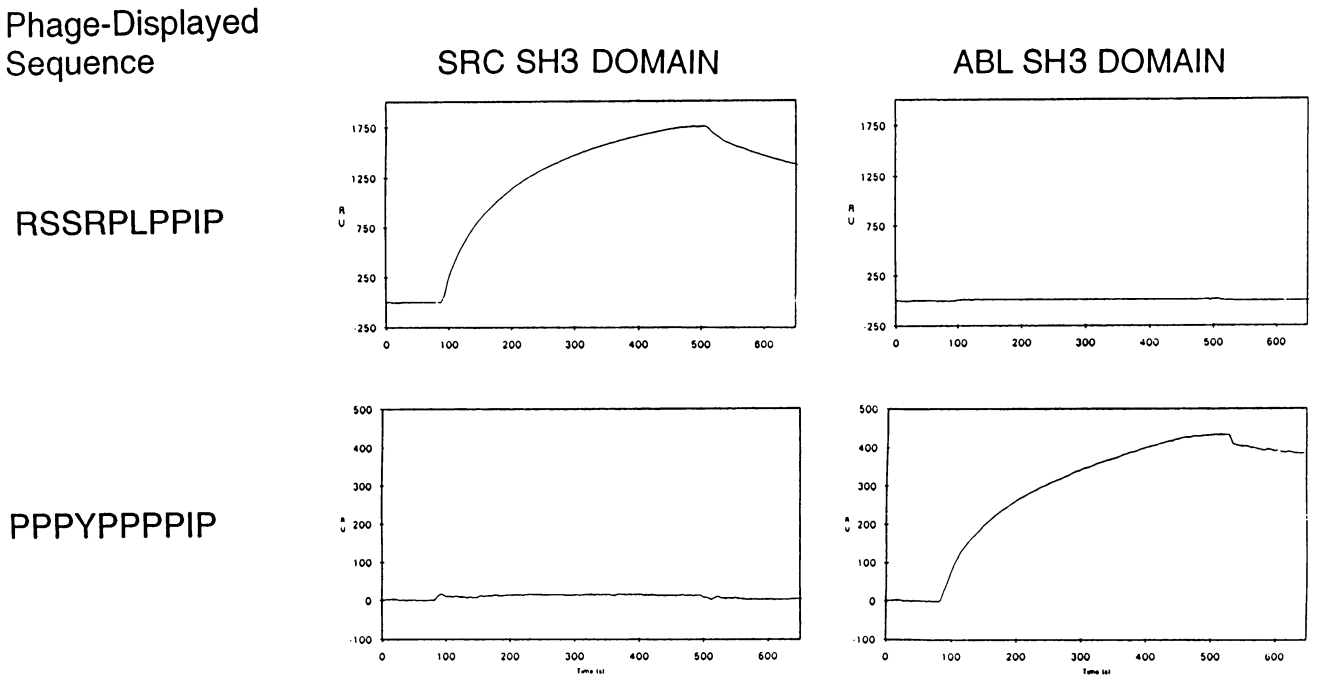

Fig. 2. Binding of phage-displayed peptides to Src and Abl SH3 domains. Binding of phage-displayed peptides was measured using BIAcore® technology as described under Materials and methods. Interaction between the receptor (immobilized SH3 domain) and ligand (phage-displayed peptide) is indicated by an increase in RU value. Traces have been corrected for bulk refractive index contributions of the unbound phage. **Upper right panel**: binding of 'RSSLRPLPPIP phage' to immobilized GST−Src SH3. **Upper left panel**: binding of 'RSSLRPLPPIP phage' to immobilized GST−Abl SH3. **Lower left panel**: binding of 'PPPYPPPPIP phage' to immobilized GST−Src SH3. **Lower right panel**: binding of 'PPPYPPPPIP phage' to immobilized GST−Abl SH3.

Table II. Binding affinities of Src and Abl SH3 domains for peptides

	Abl (X ± SEM)	hSrc (X ± SEM)	aSrc (X ± SEM)
aePPPYPPPIPggk	2.0 ± 0.2	NDB	NDB
aeRSSRPLPPIPggk	NDB	19.5 ± 1.1	58.5 ± 10
kggAPTMPPPLPPVPPg	66.3 ± 19.3	NDB	NDB

Binding of human Src (hSrc), avian Src (aSrc) and Abl (abl) SH3 domains to Src-selected (aeRSSRPLPPIPggk), Abl-selected (aePPPYPPPPIPggk) and 3BP1 (kggAPTMPPPLPPVPPg) sequences. kD values are micromolar and represent the Mean ± SEM (n = 3). NDB: no detectable binding at a protein concentration of 50 μM.

binding to GST−Src SH3 was not observed. Peptides corresponding to phage displayed sequences were synthesized and used to quantify Src and Abl SH3 binding more accurately. Using the 'Src' and 'Abl' peptides, binding was studied *in vitro* by surface plasmon resonance. Src and Abl SH3 binding was selective. The Abl SH3 domain bound with high affinity to the Abl- and not the Src-selected peptide. Conversely, the human and avian Src SH3 domains recognized the Src- but not the Abl-selected sequence. The Abl SH3 domain, but not the human or avian Src SH3 domain, bound a peptide containing residues from the protein 3BP-1. The results are summarized in Table II.

The SH3-specific peptide ligands selected here can be used as tools to dissect cellular processes mediated by SH3 domains. We previously characterized a set of proteins that associate with Src SH3 from cellular extracts (Weng *et al.*, 1993, 1994). We examined the ability of Src- and Abl-selected peptides to block the association of SH3-binding proteins from cellular extracts (Figure 3). In this experiment, tyrosine-phosphorylated proteins that bind to the Src SH3 domain were visualized by immunoblotting. The Src SH3 selected peptide blocked binding of these proteins to the Src SH3 domain with 50% inhibition detected at ~3.5 μM. The Abl SH3-selected peptide competed at a significantly higher concentration with an ~30-fold difference in the binding affinities of these two peptides observed. An unrelated peptide (KGELRL-RNYYYDVV) failed to compete even at 112 μM (data not shown).

Discussion

Our studies confirm that the SH3 domains of Src, Fyn, Lyn, PI3K and Abl SH3 prefer proline-rich peptides. Through the repeated use of biased libraries we were able to probe SH3 recognition over a 12 amino acid window. The Src SH3 domain prefers the sequence XXXRPLPPLPXP, Fyn prefers XXXRPLPP(I/L)PXX, Lyn prefers RXXRPLPPLPXP, PI3K prefers RXX-RPLPPLPPP while the Abl SH3 domain selects phage containing the sequence PPPYPPPP(I/V)PXX. Although phage panning was performed using SH3 domains as GST fusions, we do not believe that GST is contributing to sequence selectivity for two reasons. First, GST by itself failed to enrich any specific phage. Second, a truncated form of the Lyn tyrosine kinase containing the unique SH3 and SH2 domains selects sequences similar to those isolated when using GST−Lyn SH3 (data not shown). The peptides identified in this report probably define the minimal sequence requirements for efficient recognition by the SH3 domains studied. For example, using GST-

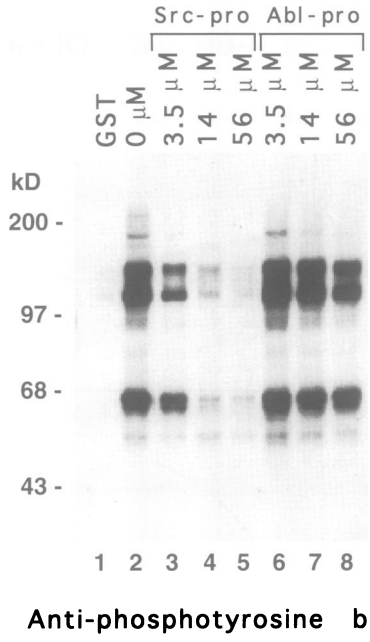

Fig. 3. Binding of tyrosine phosphorylated proteins to the Src SH3 domain and peptide competition assay. SRD 3T3 cells were lysed and incubated with glutathione agarose bound GST (lane 1) or GST-Src SH3 (lanes 2) as described (Weng *et al.*, 1993). The bound proteins were eluted with SDS sample buffer, fractionated on a 10% polyacrylamide gel, transferred to nitrocellulose and probed with the phosphotyrosine monoclonal antibody 4G10. For lanes 3–8, binding assays were performed with increasing amounts of 'src-pro' peptide (aeRSSRPLPPIPggk, lanes 3–5) or 'abl-pro' peptide (aePPPYPPPPIPggk, lanes 6–8).

Src SH3 to screen phage display libraries that contain random residues either upstream or downstream of the Src SH3 preferred sequence (RPLPPLPXP), we observe additional amino acid preferences (data not shown). Whether these flanking sequences make contact with the SH3 domain or are important for ligand presentation are currently being investigated.

While this manuscript was in preparation, Yu *et al.*, (1994) described studies where they identified RXLPPLPRψ as the Src SH3 and RXLPPRPXX as the PI3K SH3 ligand consensus sequence using a library of peptides on beads (Figure 1D). A single combinatorial library was synthesized, biased so that the prolines were invariant (XXXPPXPXX). Using peptide beads or phage display, the Src SH3 domain selected ligands with arginine at position 1 and leucine at positions 3 and 6 while the PI3K SH3 domain selected arginine at position 1 and leucine at position 3. All of the ligands that we have identified by phage display also contain prolines at the positions that were biased in the peptide library, confirming the importance of the PPXP motif for SH3 domain recognition. The differences in the Src and PI3K SH3 consensus sequences selected by the phage display and peptide bead libraries probably represent differences in the presentation of the peptides using these two distinct approaches.

The selection of specific, high affinity peptide ligands has aided our understanding of the molecular details of SH3 ligand recognition. Yu and co-workers have derived the solution structure of the SH3 domain of p85 PI3K bound to a high affinity peptide selected from bead libraries. This structure predicts that the peptide ligand forms a left-handed type II polyproline helix when bound to the PI3K SH3 domain. The N-terminal arginine from this peptide contacts Asp21 of the PI3K SH3 domain through a salt bridge (Yu *et al.*, 1994). Aspartic or glutamic acid is present at the analogous position in most SH3 domains, including those found in the members of the Src family of tyrosine kinases, PI3K, Grb-2, Nck, Crk and Csk (Musacchio *et al.*, 1992). In contrast, the Abl SH3 domain contains threonine at this position. This difference in Abl may play a significant role in ligand selection as Src, Fyn, Lyn and PI3K selected arginine while the Abl SH3 domain selected tyrosine at the 4th position of the peptides selected by phage display.

Proline-rich motifs have been identified in proteins affinity isolated using SH3 domains (Egan *et al.*, 1993; Flynn *et al.*, 1993; Gout *et al.*, 1993; Liu *et al.*, 1993; Olivier *et al.*, 1993; Prasad *et al.*, 1993; Ren *et al.*, 1993; Rozakas-Adcock *et al.*, 1993; Vogel and Fujita, 1993; Weng *et al.*, 1993, 1994; Fumagalli *et al.*, 1994; Kapeller *et al.*, 1994; Pleiman *et al.*, 1994; Taylor and Shalloway, 1994). These candidate natural SH3 ligands resemble, but are not identical to, the sequences selected by phage display. Natural SH3 binding determinants may include additional discontinuous sequences that are only observed in the folded structure. In addition, SH3 domains may recognize proteins that contain poor binding sites. This could occur if two SH3 domains co-ordinated binding intramolecularly (e.g. Grb-2) or intermolecularly, when part of a signalling complex, so as to bind efficiently to proteins with low affinity sites. In addition, other modules, such as the SH2 or pleckstrin homology domains could help to stabilize weak affinity SH3 interactions (Pawson, 1988; Koch *et al.*, 1991; Pawson and Gish, 1992; Mayer and Baltimore, 1993; Mayer *et al.*, 1993; Musacchio *et al.*, 1993).

We have tried to obtain peptide ligands for Csk and the neuronal Src (nSrc) SH3 domains using the hexapeptide libraries described in this report as well as a phage library (10^9 recombinants) displaying 10 random amino acids. No phage enrichment was observed for either SH3 domain using these libraries. It may be that the appropriate binding sequences were not present in these libraries. Alternatively, these two SH3 domains may not be folded properly after purification from bacteria. These issues are being investigated currently.

In this report we have shown that the use of phage display can provide important information on molecular recognition by SH3 domains. For example, we have shown that although the sequences recognized by the Abl and Src SH3 domains are proline rich, ligands can be identified that are distinct. Phage display can be used to identify other SH3 domain binding motifs. Furthermore, the selected peptides can be used to study interactions between natural ligands and SH3 domain-containing proteins. Lastly, structures of these peptides complexed with an SH3 domain may help in the design of drugs that block SH3-mediated disease states.

Materials and methods

Reagents

Oligonucleotides were synthesized with an Applied Biosystems 394 DNA–RNA synthesizer and purified on OPC columns. Peptides were

synthesized with an Applied Biosystems 433A Peptide synthesizer and purified by HPLC. Sequencing reactions were carried out using the Sequenase kit (U.S. Biochemicals) as described by the manufacturer.

Plasmids

The Abl SH3 domain [amino acids 84–138 of the murine type IV c-Abl cloned into pGEX-2T (Cicchetti *et al.*, 1992)] was kindly provided by David Baltimore. The Fyn, Lyn and avian Src SH3 expression vectors have been described (Weng *et al.*, 1993, 1994). The Src SH3 domain [amino acids 84–145 of the human c-Src (Tanaka *et al.*, 1987)] was amplified by PCR and cloned into the *Bam*HI site of pGEX-2TK (Pharmacia). PI3K SH3 domain sequences [amino acids 2–86 (Skolnik *et al.*, 1991)] were cloned into the *Bam*HI site of pGEX-2T. Fidelity of the PCR and cloning steps was confirmed by direct DNA sequencing of the final construction.

Protein purification

All proteins were produced as GST fusions in *Escherichia coli* strain BL21 (Novagen, Inc.) and purified as described (Smith and Johnson, 1988). Where indicated, free SH3 domains were produced from the GST protein following digestion with thrombin and removal of the GST domain by glutathione affinity chromatography (Smith and Johnson, 1988). Both fusion and non-fusion proteins were >95% pure as judged by SDS–PAGE.

Construction of phage display libraries

Four phage display libraries were prepared in bacteriophage M13. Each library explored a sequence space of six amino acids. The random sequences were inserted into the M13 gene III sequence to produce chimeric proteins (after post-translational processing) of the following forms:

Library 1 'GX_6G' NH_2-AE GXXXXXXG TVESCL

Library 2 'X_6PPIP' NH_2-AE XXXXXXPPIPG TVESCL

Library 3 '$RSLRPLX_6$' NH_2-AE RSLRPLXXXXXXG TVESCL

Library 4 '$PPPYPPX_6$' NH_2-AE PPPYPPXXXXXXG TVESCL

Underlined amino acids belong to the native gene III protein, bold amino acids represent library positions, and X represents the randomized residues. The number of recombinants in each unamplified library was: 2.5×10^9 (Library 1); 2.0×10^9 (Library 2); 5.8×10^8 (Library 3); and 3.1×10^8 (Library 4).

The libraries were synthesized by site-directed mutagenesis (Zoller and Smith, 1987) using M13 bacteriophage containing an amber codon (TAG) at position 20 of the gene III sequence. Mutagenic oligonucleotides were designed to delete or revert the amber codon to glutamic acid and insert the library sequence between position 20 and 21 of the gene III DNA. Oligonucleotide annealing and synthesis of complementary DNA were performed using reagents from the Muta-Gene M13 Mutagenesis kit (Biorad). Oligonucleotides sequences were: 5′-taaacaactttcaacagtccc XNN XNN XNN XNN XNN XNN cccagcggag tgagaata-3′ (Library 1); 5′-gttttgctaaacaactttcaacagtcccggg tattggtgg XNN XNN XNN XNN XNN XNN ttcagcggagtgagaatagaaagg-3′ (Library 2); 5′-gttttgctaaacaactttcaacagt YNN YNN YNN YNN YNN YNN caaggggcggagtgagcgttcagcggagtgagaatagaaagg-3′ (Library 3); and 5′-gttttgctaaacaactttcaacagt YNN YNN YNN YNN YNN YNN cggtgg atatggtggtggttcagcgg agtgagaatagaaagg-3′ (Library 4). 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, X and Y (N = A, G, C+T; X = G+T; Y = C+A). Recombinants were selected by electroporation into the *E.coli supE*$^-$ strain JS5 and bacteriophage propagated in 500 ml of 2× YT media (Sambrook *et al.*, 1989) containing 10 ml of stationary phase JS5 cells at 37°C for 6 h. The supernatant was cleared twice by centrifugation (4100 *g*) to remove bacteria and phage precipitated by adjusting the supernatant to 0.5 M NaCl/4% polyethylene glycol and recentrifuging as above. Phage pellets were resuspended and stored in Tris-buffered saline (TBS, 50 mM Tris–HCl pH 7.5/150 mM NaCl) at 4°C.

Phage enrichment

Polystyrene tissue culture wells (2.0 cm^2) were coated overnight at 4°C with 10 μg of GST or GST–SH3 protein in 100 mM $NaHCO_3$ (pH 8.5) followed by bovine serum albumin (BSA) blocking buffer (30 mg BSA/ml in 100 mM $NaHCO_3$, pH 8.5) for 30 min at room temperature. Phage (1.5×10^{11} infectious particles) were panned overnight in 250 μl of 50 mM Tris–HCl, pH 7.5/150 mM NaCl/0.5% BSA/0.05% Tween 20 at 4°C. Wells were washed three times for 10 min with 0.5 ml TBS/0.5% BSA/0.05% Tween 20 at room temperature. Bound phage were eluted using 200 μl of 0.2 M glycine, pH 2.1/0.5% BSA/0.05% Tween 20 (30 min at room temperature) and neutralized with 50 μl 1 M Tris base. Eluted phage were amplified in 20 ml 2×YT media containing 0.4 ml of stationary phase JS5 cells and purified/concentrated using 0.5 M NaCl/4% polyethylene glycol as described above. Panning was repeated as above with the amplified phage until the desired degree of enrichment was achieved (see Results).

Binding assays

Binding assays were performed using a BIAcore Biosensor® (Pharmacia Biosensor, Piscataway, NJ). This instrument measures the interaction between two or more macromolecules in real-time through the generation of a quantum-mechanical surface plasmon. For reviews of the theory, the instrument and its application see references Jonnson *et al.* (1991), Panayoyou *et al.* (1993) and Malmqvist (1993).

Binding of phage-displayed peptides. To rank the selected sequences in order of binding affinity, homogeneous phage stocks were prepared in phosphate buffered saline (PBS, 10^{13} pfu/ml) and scored for the ability to bind to immobilized GST–Src and GST–Abl proteins using BIAcore. A Biosensor Chip CM5 was derivatized with a lysine-reactive coupling group (*N*-hydroxy-succinimide) and 400–450 RU of GST–Src or GST–Abl were covalently linked to the chip through solvent-exposed lysines. Specifically, the chip was treated with 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride/0.05 M *N*-hydroxy-succinimide for 5 min to create the reactive linker; 10 μg/ml of GST–Src or GST–Abl in 10 mM potassium acetate (pH 5.0) for 3 min; and 1.0 M ethanolamine hydrochloride for 5 min to block unreacted *N*-hydroxy-succinimide esters. The chip was then washed for 2 min with 0.1 N NaOH and 5 min with PBS (1 mM KH_2PO_4, 10 mM Na_2HPO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4) to remove non-covalently bound protein. Identical wash cycles were used to regenerate the chip between assays. No loss of bound SH3 or decrease in phage-specific binding capacity was detected after >100 NaOH/PBS regeneration cycles. The immobilization and assays were performed at 25°C with a flow rate of 5 μl/min in a PBS running buffer. All measurements were corrected for contributions of the unbound phage to the bulk refractive index of the buffer. Contrary to normal BIAcore conditions, Surfactant P20 (Pharmacia Biosensor) was not included in the running buffer as it negatively affected phage integrity.

Determination of binding constants. 250–300 RU of peptide (Ac-AEPPPYPPIPGGK-NH_2, Ac-AERSSRPLPPIPGGK-NH_2, or Ac-KGG-APTMPPPIPPVPPG-NH_2) was covalently linked to a Biosensor Chip CM5 through the N- or C-terminal lysine residue using the chemistry described above. Binding constants were determined from equilibrium binding values at five or more protein concentrations at 25°C in PBS at a flow rate of 5 μl/min using BIAcore. To remove the confounding influence of GST-based dimerization, binding constants were determined using non-fusion Src or Abl SH3 domains as described above. Removal of non-covalently bound peptide and regeneration of the chip between assays was accomplished by treatment with 6 M guanidine hydrochloride (pH 7.0) for 1 min and PBS for 5 min. No loss of bound peptide or decrease in SH3 binding capacity was detected after >400 regeneration cycles. All measurements were corrected for contributions of the unbound protein to the bulk refractive index of the buffer.

Acknowledgements

We would like to thank our colleagues at ARIAD for helpful discussions and D.Baltimore for kindly providing the GST–Abl SH3 expression vector. This work was funded in part by grant CA27951 (JSB) from the National Institutes of Health.

References

Chen,J.K., Lane,W.S., Brauer,A.W., Tanaka,A. and Schreiber,S.L. (1993) *J. Am. Chem. Soc.*, **115**, 12591–12592.

Cicchetti,P., Mayer,B.J., Thiel,G. and Baltimore,D. (1992) *Science*, **257**, 803–806.

Egan,S.E., Giddings,B.W., Brooks,M.W., Buday,L., Sizeland,A.M. and Weinberg,R.A. (1993) *Nature*, **363**, 45–51.

Flynn,D.C., Leu,T.-H., Reynolds,A.B. and Parsons,J.T. (1993) *Mol. Cell. Biol.*, **13**, 7892–7900.

Fumagalli,S., Totty,N.F., Hsuan,J.J. and Courtneidge,S.A. (1994) *Nature*, **368**, 871–874.
Gout,I. *et al.* (1993) *Cell*, **75**, 25–36.
Jonsson,U., Fagerstam,B., Ivarsson,B., Johnsson,B. and Karlsson,R. (1991) *BioTechniques*, **11**, 620–627.
Kapeller,R., Prasad,K.V.S., Jansson,O., Hou,W., Schaffhausen,B.S., Rudd,C.E. and Cantley,L.C. (1994) *J. Biol. Chem.*, **269**, 1927–1933.
Koch,C.A., Anderson,D., Moran,M.F., Ellis,C. and Pawson,T. (1991) *Science*, **252**, 668–674.
Liu,X., Marengere,L.E.M., Koch,C.A. and Pawson,T. (1993) *Mol. Cell. Biol.*, **13**, 5225–5232.
Malmqvist,M. (1993) *Nature*, **361**, 186–187.
Mayer,B.J. and Baltimore,D. (1993) *Trends Cell Biol.*, **4**, 8–13.
Mayer,B.J., Jackson,P.K. and Baltimore,D. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 627–631.
Mayer,B.J., Ren,R., Clark,K.L. and Baltimore,D. (1993) *Cell*, **73**, 629–630.
Musacchio,A., Gibson,T., Lehto,V.-P. and Saraste,M. (1992) *FEBS Lett.*, **307**, 55–61.
Musacchio,A., Gibson,T., Rice,P., Thompson,J. and Saraste,M. (1993) *Trends Biol. Sci.*, **18**, 343–345.
Olivier,J.P., Raabe,T., Henkemeyer,M., Dickson,B., Mbabalu,G., Margolis,B., Schlessinger,J., Hafen,E. and Pawson,T. (1993) *Cell*, **73**, 191–197.
Panayotou,G., Waterfield,M.D. and End,P. (1993) *Curr. Biol.*, **3**, 913–915.
Parmley,S.F. and Smith,G.P. (1988) *Gene*, **73**, 305–318.
Pawson,T. (1988) *Oncogene*, **3**, 491–495.
Pawson,T. and Gish,G. (1992) *Cell*, **71**, 359–362.
Pleiman,C.M., Hertz,W.M. and Cambier,J.C. (1994) *Science*, **263**, 1609–1612.
Prasad,K.V., Janssen,O., Kapeller,R., Raab,M., Cantley,L.C. and Rudd,C.E. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 7366–7370.
Ren,R., Mayer,B.J., Cicchetti,P. and Baltimore,D. (1993) *Science*, **259**, 1157–1161.
Rozakas-Adcock,M., Fernley,R., Wade,J., Pawson,T. and Bowtell,D. (1993) *Nature*, **363**, 83–85.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Skolnik,E.Y., Margolis,B., Mohammadi,M., Lowenstein,E., Fischer,R., Drepps,A., Ullrich,A. and Schlessinger,J. (1991) *Cell*, **65**, 83–90.
Smith,D.B. and Johnson,K.S. (1988) *Gene*, **67**, 31–40.
Smith,G.P. (1985) *Science*, **228**, 1315–1317.
Songyang,Z. *et al.* (1993) *Cell*, **72**, 767–778.
Tanaka,A., Gibbs,C.P., Arthur,R.B., Anderson,S.K., Kung,H.-J. and Fujita,D.J. (1987) *Mol. Cell. Biol.*, **7**, 1978–1983.
Taylor,S.J. and Shalloway,D. (1994) *Nature*, **368**, 867–871.
Vogel,L.B. and Fujita,D.J. (1993) *Mol. Cell. Biol.*, **13**, 7408–7417.
Weng,Z., Taylor,J.A., Turner,C.E., Brugge,J.S. and Seidel-Dugan,C. (1993) *J. Biol. Chem.*, **268**, 14956–14963.
Weng,Z., Thomas,S.M., Rickles,R.J., Taylor,J.A., Brauer,A., Seidel-Dugan,C., Michael,W.M., Dreyfuss,G. and Brugge,J.S. (1994) *Mol. Cell. Biol.*, **14**, 4509–4521.
Yu,H., Chen,J.K., Feng,S., Dalgarno,D.C., Brauer,A.W. and Schreiber,S.L. (1994) *Cell*, **76**, 933–945.
Zoller,M.J. and Smith,M. (1987) *Methods Enzymol.*, **154**, 329–350.

Received on July 20, 1994, revised on September 6, 1994

Note added in proof

After the acceptance of this manuscript for publication, two papers were published that describe the identification of Src SH3 ligands from phage display libraries [Sparks *et al.* (1994) *J. Biol. Chem.*, **269**, 23853–23856; Cheadle *et al.* (1994) *J. Biol. Chem.*, **269**, 24034–24039].