J. Mol. Biol. (1997) **269**, 694–703

JMB

Modified Phage Peptide Libraries as a Tool to Study Specificity of Phosphorylation and Recognition of Tyrosine Containing Peptides

Luciana Dente[1], Costantino Vetriani[1], Adriana Zucconi[1] Giuliana Pelicci[2], Luisa Lanfrancone[2], Pier Giuseppe Pelicci[2] and Gianni Cesareni[1]*

[1]*Dipartimento di Biologia Università di Roma Tor Vergata, Via Carnevale 00173, Rome, Italy*

[2]*Department of Experimental Oncology, European Institute of Oncology, Via Ripamonti 435 20141, Milano, Italy*

**Corresponding author*

Tyrosine phosphorylation and protein recognition, mediated by phosphotyrosine containing peptides, play an important role in determining the specific response of a cell, when stimulated by external signals. We have used peptide repertoires displayed by filamentous phage as a tool to study the substrate specificity of the protein tyrosine kinase (PTK) p55^{fyn} (Fyn). Peptide libraries were incubated for a short time in the presence of Fyn and phages displaying efficiently phosphorylated peptides were selected by panning over anti-phosphotyrosine antibodies. The characterization of the peptides enriched after three phosphorylation/selection rounds allowed us to define a canonical substrate sequence for the kinase Fyn, E–(φ/T)**Y**Gxφ, where φ represents any hydrophobic residue. A peptide conforming to this sequence is a better substrate than a second peptide designed to be in accord with the consensus sequence recognised by the Fyn SH2 domain. When the library phosphorylation reaction is carried out in saturation conditions, practically all the tyrosine containing peptides are phosphorylated, irrespective of their context. These "fully modified" peptide libraries are a valuable tool to study the specificity of phosphotyrosine mediated protein recognition. We have used this new tool to identify a family of peptides that bind the PTB domain of the adapter protein Shc. Comparison of the peptide sequences permits us to confirm the essential role of N at position −3, while P often found at position −2 in natural targets is not absolutely required. Furthermore, our approach permits us to reveal an "extended" consensus indicating that residues that do not seem to influence binding in natural peptides can make productive contacts, at least in linear peptides.

© 1997 Academic Press Limited

Keywords: Fyn kinase; molecular repertoires; phosphotyrosine binding domains; phage display; anti-phosphotyrosine mAbs

Introduction

Transduction of external signals in cells is mediated by an intricate pattern of protein interactions which contribute to the activation of enzymatic reactions, mostly protein phosphorylation, and which, in turn, can be modulated by post translational modifications of one of the two partners (Pawson, 1995). Describing and understanding the rules underlying substrate preference and binding specificity is of primary importance. Furthermore, the ability to isolate a peptide which competes with a protein function involved in the transmission of stimuli that lead to cell proliferation, often represents an essential step in the development of small molecules (peptidomimetics) with anti-tumour activity (Kohl *et al.*, 1993). Among the experimental approaches that have been recently developed and exploited to study molecular recognition, the screening of synthetic peptide libraries has raised considerable interest.

Abbreviations used: PTK, protein tyrosine kinase; anti-pY mAb, anti-phosphotyrosine monoclonal antibody.

0022–2836/97/250694–10 $25.00/0/mb971073
© 1997 Academic Press Limited

The success of this approach derives from its speed and precision in providing information about either the peptides that may bind a specific domain involved in protein-protein interaction (Songyang *et al.*, 1993) or the optimal substrate for a protein kinase (Songyang *et al.*, 1995a). The synthetic library approach, however, has some limitations which are inherent to the experimental protocol which relies on the analysis of pools rather than single sequences. For instance, whenever more than one sequence family binds the same target protein, the synthetic peptide approach may reveal only the more numerous and less specific one. Furthermore a rare, strongly binding peptide, would not be spotted by this type of experiment, since its sequence would be hidden by families of sequences which bind with lower affinity but are more numerous. Although these may not represent strong limitations, as demonstrated by the success of this method, it would still be desirable to investigate alternative approaches.

Peptide libraries displayed by filamentous phage by fusion either to the amino terminus of the receptor protein pIII (Scott & Smith, 1990) or to the major coat protein pVIII (Felici *et al.*, 1991) may represent a valid alternative. According to this approach, the peptides of random sequence are not free in solution but linked to a filamentous phage capsid protein. This is achieved by genetic fusion of oligonucleotides of random sequence to a capsid protein gene. Similarly to synthetic peptide libraries, phage libraries may be exploited as molecular repertoires from where to select, essentially by affinity purification (panning: Parmley & Smith, 1988), those phages that display peptides which may interact with a given protein of interest. The fundamental difference with respect to the synthetic library approach is that peptides are linked, indirectly *via* the fusion to a capsid protein, to their coding sequences. This allows ''cloning'' any binding peptide and facilitates its characterization after amplification of the corresponding phage clone. Only afterwards may these clones be grouped into families of similar sequences thus allowing the identification of one or more consensus sequences. Moreover, the phage library approach permits the selection of rare peptides that have a very high affinity for the target protein by applying several cycles of selection-amplification in stringent conditions.

Here we have extended the scope of phage display technology by showing how peptide libraries can be used to investigate the substrate specificity of Fyn, a protein kinase of the Src family. Furthermore we will describe the construction and exploitation of fully modified peptide libraries to determine the phosphotyrosine sequence context which modulates the recognition specificity of the PTB domain of the protein Shc.

Results

Tyrosine phosphorylation of peptides displayed by filamentous phage

All the peptide libraries, used in the selections described below, were constructed by inserting oligonucleotides of random sequence into the gene encoding the filamentous major capsid protein pVIII (Felici *et al.*, 1991; Iannolo *et al.*, 1995). Expression of these recombinant genes yields hybrid pVIII proteins which accommodate random peptides inserted after the amino-terminal pentapeptide AEGEF.

We reasoned that such libraries could be exploited to investigate substrate specificity in the protein kinase reaction by incubating a mixture of phage displaying random peptides in the presence of a specific protein tyrosine kinase (PTK). Phage that display peptides that are efficiently phosphorylated could then be selected with anti-phosphotyrosine monoclonal antibodies (anti-pY mAb). Furthermore, phosphorylated phage libraries would be a useful tool to approach recognition specificity of protein modules (SH2, PTB) that bind phosphotyrosine peptides.

In order to explore the feasibility of such an approach, we have incubated various phage suspensions in the presence of the Fyn PTK (Cheng *et al.*, 1992). The two tyrosine residues in the pVIII capsid protein (Y21 and Y24) are not phosphorylated by the Fyn PTK, since the labelling pattern of a mutant carrying a double substitution F21, S24 is indistinguishable from wild-type (Figure 1, lanes 2 and 3). By contrast the pVIII protein of hybrid phages displaying Y containing peptides is labelled by Fyn. In Figure 1, lane 1 is shown, as an example, the labelling pattern of phage pM48.23 displaying the peptide THYQGRAGA. Any labelling of tyrosine in the minor capsid proteins will not be seen in the gel autoradiography in Figure 1 because of the low copy number (5 to be compared with approximately 2700 of the major coat protein).

To test whether any residual labelling of wild-type pVIII, or of any minor coat protein, could interfere with the selection procedure that we have conceived, we have measured the fraction of phage that is retained by a polystyrene bound anti-pY mAb after incubation in the presence of Fyn PTK. Wild-type and F21-S24 mutant hardly bind the antibody, while a large fraction ($\sim 6 \times 10^{-2}$) of pM48.23 is specifically retained (Figure 1b). Collectively these results prove that labelling of wild-type capsid proteins, if any, cannot be detected, over background levels, by affinity purification with anti-pY mAb.

PTKs characterized so far display a certain degree of substrate specificity (Songyang *et al.*, 1995a). After a long incubation time or at high enzyme concentration, however, most tyrosine containing peptides are eventually phosphorylated. Kinetic studies performed on different phages displaying peptides containing Y (not shown) have

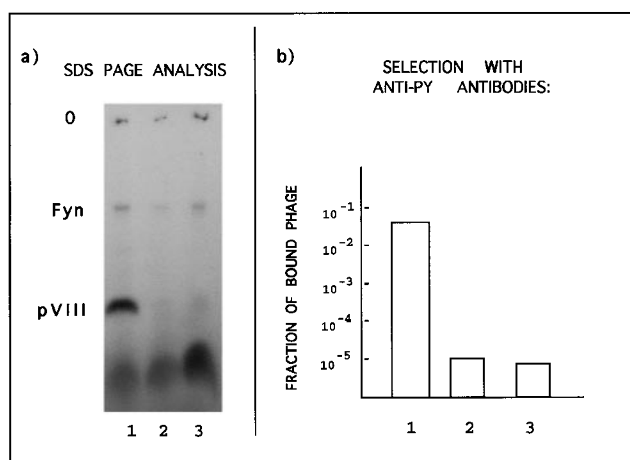

Figure 1. Selective phosphorylation of filamentous phage displaying tyrosine containing peptides. (a) Phage supernatants were phosphorylated, as described in Materials and Methods, and analysed by SDS-PAGE and autoradiography. Lane 1, PM48.23, a phage displaying the peptide THYQGRAGA (Iannolo *et al.*, 1995); lane 2, the wild-type phage f1; lane 3, PM50, a f1 phage mutant where Y21 and Y24 of PVIII are substituted by F and S respectively. O represents the electrophoresis origin. (b) Approximately 10^{10} phosphorylated phages were affinity selected with anti-pY mAb (alpha-1) and eluted by incubating in glycine-HCl, pH 2. Bars represent the ratio between the bound and the input phage titre.

permitted us to define the fraction of displayed peptides that are phosphorylated by the Fyn PTK at any time, in our conditions. In the experiments aimed at determining the substrate specificity of the Fyn PTK we have utilized short reaction times (three minutes) to limit substrate phosphorylation. In these conditions most phages displaying tyrosine containing peptides are not labelled at all, while only phage displaying a "good peptide substrate" will be phosphorylated and eventually selected by panning over an anti-pY affinity substrate.

The anti-pY mAbs that we have utilized, aside from binding phosphotyrosine containing peptides, also recognise peptides that do not contain phosphotyrosine. When we panned a nonapeptide library over the alpha 1 or the pY20 anti-pY mAbs, after two selection cycles, clones displaying the peptides in Table 1 were selected. These peptide families identify the recogniton consensus PWXGTT and WLDAR for alpha 1 and pY20, respectively, where X stands for any amino acid. These represent a second specificity that prevails over the common pY specificity when the library is poorly phosphorylated and interferes with our selection procedure. Thus, we developed a protocol involving sequential selection cycles, with the two antibodies, to favour selection of the few phages that display peptides that are phosphorylated in our conditions.

After three sequential cycles involving selection with mAb pY20, followed by mAb alpha 1 and finally again pY20, we have identified few phages that bind the antibodies in a phosphotyrosine dependent manner.

The relative enrichment of kinased over untreated phage was monitored after each cycle and is shown in Table 2. At the end of the selection procedure the pool of selected phages was retained at least two orders of magnitude better than the background by any of the immobilized anti-pY mAbs. Isolated clones were amplified and the DNA sequences of the pVIII genes were determined.

Binding of the anti-phosphotyrosine antibodies to the selected phages was proven to be phosphorylation-dependent by micropanning analysis. These phages were retained at least two orders of magnitude better when phosphorylated, in agreement with the results obtained with the pool of selected phages (Table 2).

Inspection of the amino acid sequence of the peptides displayed by the selected phages reveals some specific features that shed light on the substrate specificity of the Fyn kinase (Table 3). In position −1, with respect to the phosphorylated tyrosine, we observe either a T or a hydrophobic residue (φ). Position +1 is preferentially occupied by a G, while position +2 does not seem to affect substrate specificity. Finally, hydrophobic sidechains are strongly preferred at position +3.

It has been observed that PTKs preferentially phosphorylate Y residues preceded by acidic residues (House *et al.*, 1984; Hunter, 1982; Casnellie *et al.*, 1982). In our library, the negative charge is always provided by the amino terminus of the

Table 1. Amino acid sequences of peptides displayed by phages that bind mAb alpha1 and pY20 in a phosphotyrosine independent manner

alpha1									pY20								
		R	Y	**G**	**T**	**T**	F	P	S	F	**L**	**D**	S	**R**	L	P	S
A	**P**	**W**	N	**G**	**T**	**T**			S	Q	**L**	**D**	S	**R**	T	V	G
	P	**W**	M	**G**	**T**	G	S		E	**W**	**L**	**D**	**A**	**R**	T	N	M
	P	**W**	S	T	**T**	G	A		N	**W**	**L**	**D**	**A**	H	Q	S	M
R	**P**	**W**	N	**G**	S	**T**			A	**W**	**L**	**D**	**A**	A	A	T	L
P	**P**	**W**	G	**G**	G	**T**			T	**W**	**I**	**D**	**A**	**R**	V	T	A
									Y	L	**L**	**D**	R	**R**	T	G	L
Consensus									Consensus								
	P	**W**	×	**G**	**T**	**T**				**W**	**L**	**D**	**A**	**R**			

Table 2. Enrichment of phage displaying Fyn kinase substrates by panning over anti-pY mAbs

Cycle	Fyn PTK	Input titre	Output titre	Inp/Out	mAb utilized for panning
I	+	8×10^9	1×10^7	8×10^2	pY20
	−	1×10^{10}	1×10^7	1×10^3	
II	+	6×10^9	2×10^5	3×10^4	alpha1
	−	6×10^9	8.5×10^3	7×10^5	
III	+	3×10^9	3×10^8	1×10	pY20
	−	4×10^9	3.3×10^6	1.2×10^3	

pVIII molecule (the vector sequence preceding the insert is: AEGEF). This might explain why the Y in the selected peptides cluster near the amino terminus. Taking into consideration the E in the neighbouring pVIII sequences, all the selected peptides have an E residue in a position that varies from −5 to −2 with respect to the Y that is phosphorylated.

We conclude that E–(φ/T)**Y**GXφ represents the consensus sequence among the peptides that are phosphorylated by the Fyn PTK.

Synthetic consensus peptides as substrate for phosphorylation reaction

The consensus sequence that we have determined is in good agreement with the sequence of a p34cdc2 derived peptide (KVEKIGEGTYGVVYK), that has been found to discriminate Src like PTK activity from other PTKs, including EGFr and v-Abl (Cheng *et al.*, 1992). On the other hand, the Fyn consensus is similar, but not identical to the preferred substrate peptides identified for some Src-like kinases (c-Src, v-Src, Lck) on the basis of experiments that utilize libraries of synthetic random peptides (Songyang *et al.*, 1995a). In particular, these authors suggested that cytosolic PTKs preferentially phosphorylate peptides recognized by their own, or closely related, SH2 domains.

While no independent direct analysis of the catalytic specificity of Fyn is yet available, the recognition specificity of its SH2 domain has been determined (Songyang *et al.*, 1993). The corresponding consensus recognition peptide, pYEEI, differs from our canonical PTK substrate for the presence of negative residues in positions +1 and +2, that are not preferentially enriched in substrate peptides selection.

To confirm that the consensus peptide sequence deduced from our phage library analysis is a good substrate for Fyn, we have synthesized two peptides designed respectively on the PTK substrate consensus, as deduced from our approach (pF4) and on the Fyn SH2 recognition consensus (pS22) according to Songyang *et al.* (1993). The amino acid sequence of peptide pF4, EFGTYGTLS, corresponds to the peptide displayed by the phage that was most frequently found among the selected clones (Table 3). By contrast, the amino acid sequence of peptide pS22, PRDLYEEL, contains the consensus sequence of the Fyn SH2, as defined by Songyang *et al.* (1993). A third tyrosine containing peptide, GEFYQNTS, conforming to neither consensus was also utilized in the experiments.

The experiment in Figure 2(a) confirms that the peptide pF4 is a better substrate than pS22, with an initial reaction rate that is approximately three times higher. pG30, whose sequence GEFYQNTS does not contain the preferred residues +1 and +3, is a very poor substrate for the enzyme. The K_m of pF4 is estimated to be 70 μM, while it is higher than 200 μM in the case of pS22 (Figure 2(b)). We conclude that, at least in the case of Fyn, the best kinase substrate differs from the corresponding SH2 consensus binding peptide.

Table 3. Amino acid sequence of peptides displayed by phage that are efficiently phosphorylated by Fyn

	E	F	G	**T**	**Y**	**G**	T	**L**	S			4
E	F	Y	G	**T**	**Y**	**G**	L	**P**				2
	E	F	G	**T**	**Y**	**G**	H	**L**	S			2
			E	**F**	**Y**	**G**	R	**I**	S	G	T	
E	F	S	S	**I**	**Y**	A	S	**M**				
	E	F	H	**P**	**Y**	Y	A	G	P			
	E	F	N	**V**	**Y**	S	M	**M**	T			
			E	**F**	**Y**	R	F	N	G	E	T	

Consensus:
E– φ/**T** **G** X φ

Amino acids that belong to the pVIII insertion context are underlined. Figures on the right are the number of times that the corresponding peptide sequence was obtained by sequencing gVIII of different selected phage clones. φ in the consensus sequence represents residues with hydrophobic side-chains. In bold are residues that are consistent with the deduced consensus sequence.

Processive phosphorylation of phages mediated by the Fyn PTK SH2 domain

The phosphorylation experiments performed on synthetic peptides confirmed that, in the case of Fyn kinase, a peptide matching the consensus SH2 binding peptide does not coincide with the ideal substrate for the phosphorylation reaction. Recently, however, evidence has been accumulated that, *in vivo*, the SH2 domains of cytosolic kinases modify the spectrum of preferred substrates by binding to primary phosphorylation sites and promoting, by physically linking the substrate, the phosphorylation of secondary sites (Mayer *et al.*, 1995).

To test whether this effect could be reproduced *in vitro* we have incubated two phage preparations in the presence of Fyn kinase. Phage F4 and S22

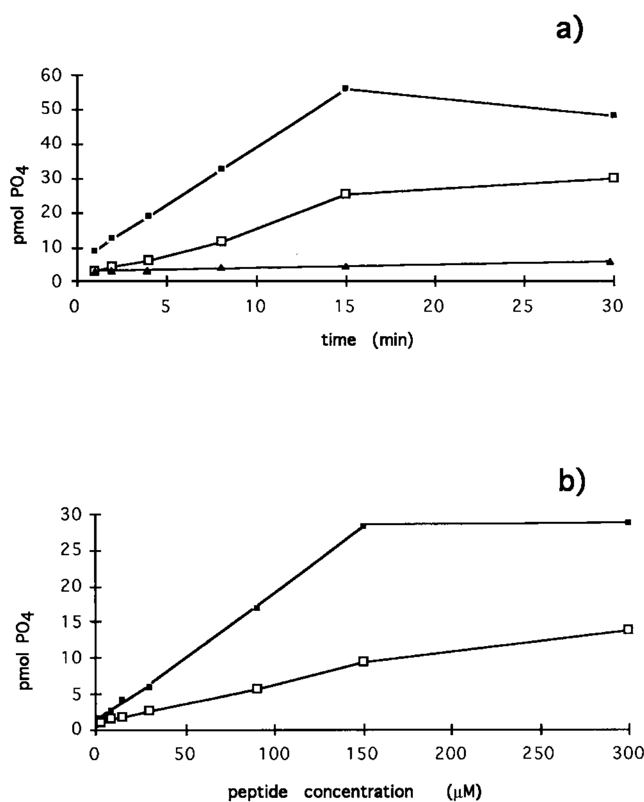

Figure 2. Kinetic analysis of peptide phosphorylation. Comparison of substrate activity of peptides pF4 (EFGTYGTLS, (■)); pS22 (PRDLYEEL, (□)); and pG30 (GEFYQNTS (▲). (a) Synthetic peptides were incubated for different times at a concentration of 300 μM with 3.5 units of Fyn kinase and 0.1 μM ATP. (b) Different concentrations of peptides were incubated for ten minutes in the same conditions as in (a). The amount of incorporated PO_4 was determined by using [γ-^{32}P]ATP as a tracer at a specific activity of 1 μCi/nmol.

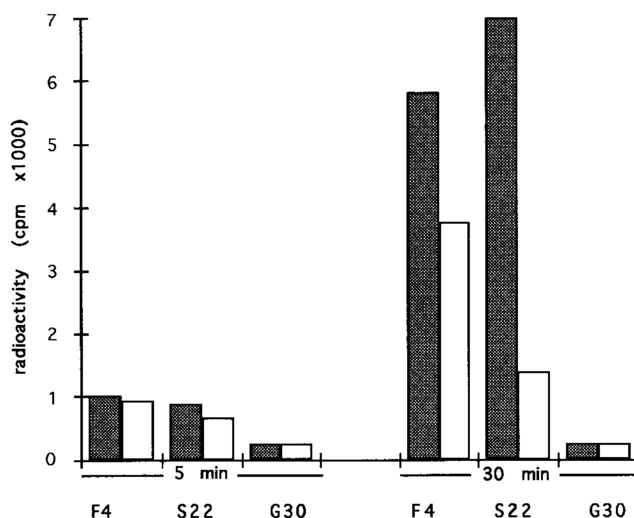

Figure 3. Influence of the Fyn SH2 binding peptide on phage phosphorylation. Purified phage preparations were phosphorylated for five and 30 minutes with (empty bars) and without (shaded bars) pre-incubation with the EPQpYEEIPI peptide. After SDS-PAGE, bands corresponding to pVIII proteins were excised from the gel and the amount of incorporated phosphate determined in a scintillation counter.

display multiple copies of peptides pF4 and pS22, respectively. These multivalent phage structures could mimic a natural substrate containing several potential phosphorylation sites.

The experiment in Figure 3 (shaded bars) indicates that, in contrast with the results obtained with peptides in solution, pS22, displayed by a phage, is not a worse substrate than pF4. This conclusion is more evident at relatively long incubation times. To prove that this effect is mediated by the ability of the kinase SH2 domain to bind phosphotyrosine containing peptides, we have repeated the experiment in the presence of excess of a phosphotyrosine containing peptide (EPQpYEEIPI) that efficiently binds the SH2 of Fyn (Alonso *et al.*, 1995). In these conditions phosphorylation of the phage displaying pS22 is markedly reduced, suggesting that SH2 mediated processivity of the kinase reaction plays a role in determining the phosphorylation level of structures displaying several copies of potential PTK targets.

Selection of phosphopeptides recognised by the pY binding domains of Shc

Some protein-protein interactions that underlie the process of signal transduction are mediated by protein domains that bind phosphotyrosine (SH2, PTB). Recognition specificity is modulated by the specific context in which the phosphotyrosine is embedded in the protein target (Songyang *et al.*, 1993; 1994; 1995a). The availability of a peptide library containing phosphotyrosine, instead of tyrosine, would permit one to exploit phage display technology to study the mechanism at the basis of phosphotyrosine peptide recognition. While studying the substrate specificity of the Fyn kinase, we have observed that, if the phosphorylation reaction time is extended (five hours), the specificity of the enzyme is broadened, and virtually every phage displaying a peptide containing a Y can be phosphorylated, irrespective of the amino acid sequence context.

To demonstrate the versatility of this new type of library, we have panned them over a GST fusion protein containing the PTB amino-terminal domain of the protein Shc (Pelicci *et al.*, 1992). Initially, for this specific purpose, we have utilized a biased library of the type X(7)YX(7) (O. Minenkova unpublished) where X stands for any amino acid. After three selection cycles the amino acid sequences of the peptides were deduced from the DNA sequences of the pVIII genes of the selected phage clones.

Table 4. Peptides displayed by phages that bind the PTB-SHC domain

A.																							
						E	**F**	L	**N**	K	D	**Y**	G	I	**Y**	G	V	H	D	N	A	S	4
						E	**F**	N	**N**	S	F	**Y**	M	S	**Y**	N	P	P	N	A	L	L	3
						E	**F**	N	**N**	C	D	**Y**	G	V	**Y**	N	R	S	H	M	N	L	
						E	**F**	Y	**N**	A	**T**	**Y**	S	A	**Y**	I	A	P	P	Q	S	M	3
						E	**F**	Y	**N**	**P**	**T**	**Y**	A	P	**Y**	D	Y	P	H	T	A	P	
						E	**F**	A	**N**	**P**	**T**	**Y**	M	V	**Y**	S	S	S	A	V	T	K	
						E	**F**	T	**N**	**P**	L	**Y**	M	S	**Y**	T	P	P	Q	P	S	K	
			E	F	R	T	I	L	**N**	H	I	**Y**	D	T	**Y**	N	P	N	Q				
			E	F	H	R	**F**	D	**N**	M	**T**	**Y**	D	T	P	V	R	W	Q				
			E	F	P	N	**Y**	H	**N**	Q	**T**	**Y**	Y	T	T	T	Y	T	P				
	E	F	Y	P	G	Y	**Y**	E	**N**	Y	G	**Y**	L	E	I	P	P						2
			E	F	E	H	**Y**	H	**N**	A	**T**	**Y**	A	S	L	I	P	T	P				
B.																							
							E	F	**N**	**P**	**T**	**Y**	E	S	**W**	A	P						2
							E	F	**N**	G	**T**	**Y**	G	R	**Y**	E	P						2
							E	F	**N**	H	**T**	**Y**	E	L	T	P	R						
					E	F	P	S	**N**	G	L	**Y**	S	Q	**W**								
						E	F	H	**N**	L	Y	**Y**	F	A	**Y**	K							
						E	**F**	F	**N**	**P**	V	**Y**	L	E	**Y**	K							
Consensus:							**F/Y**	×	**N**	**P**	**T**	**Y**	×	×	**Y/W**								

See the legend to Table 3.

As shown in Table 4A, all the PTB binding peptides can be aligned in such a way that, in each peptide sequence, an N precedes a Y in position −3. This confirms previous reports that pointed to the importance of this N for recognition of natural targets of the PTB domain of Shc (Kavanaugh *et al.*, 1995). In a few cases, but more frequently than one would expect on a pure random basis, we observe a P at position −2, and, at higher frequency a T at position −1. Both these residues were found to recur frequently in natural Shc-PTB targets (Kavanaugh & Williams, 1994). Significantly, most of the peptides are aligned in the pVIII protein context in such a way that the F at the amino terminus of pVIII is found five residues before the phosphorylated Y. In the equivalent position of the remaining peptides we either find an aromatic (Y) or an hydrophobic side-chain (I). This is consistent with the observation that most natural Shc-PTB targets contain an hydrophobic residue at that position (Songyang *et al.*, 1995b).

Furthermore, unlike natural targets, that do not show any preference at the carboxy side of the phosphotyrosine, our selected peptides exibit a marked preference for Y at position +3. In principle, this could simply be the consequence of using a biased library of the type X(7)YX(7) and of the tendency of selecting a second Y, three residues amino-terminal to the fixed Y, to ''utilize'' the hydrophobic character of the F in the amino-terminal sequence of pVIII.

However, when we submitted to an identical selection procedure a library of nonapeptides (Felici *et al.*, 1991) of completely random sequence, we could confirm a preference for an aromatic residue, W or Y, at position +3 (Table 4B).

We have also used this modified phage library to determine the preferred sequence context of phosphotyrosine recognized by several SH2 groups (Pelicci *et al.*, 1996). The results summarized in Table 5, by substantially confirming previous findings obtained with independent approaches (Songyang *et al.*, 1993, 1994), support the robustness and the reliability of the technique. Furthermore these results permit us to establish that the SH2 specificity of two Shc related proteins, Sli and Rai, share the same recognition specificity of the SH2 of Shc. In all cases however we define an extended consensus that includes more residues than those recognised to influence SH2 peptide recognition in natural targets.

Table 5. Consensus sequences of peptides that are either efficiently phosphorylated by Fyn or bind several phosphotyrosine binding domains

Fyn PTK	E	x − x	T/φ	**Y**	G	x	I/V/L	
	−3	−2	−1		+1	+2	+3	+4
GRB/SH2			V	**pY**	Q	N	W/F	
SHC/SH2		N	I/V	**pY**	E/G	T	I/V/L	W/F
SLI/SH2			L/I/V	**pY**	E/G	T	W/Y/F	
RAI/SH2			V/I	**pY**	E/G	T/Y	I/L	
SHC/PTB	N	p	T	**pY**	x	x	Y/W	

Discussion

Target choice, in protein tyrosine kinases, is determined by a fine balance between the substrate specificity of its enzymatic domain and the substrate availability. The latter is usually influenced by enzyme-substrate co-localization mediated by protein recognition modules like the SH2 and SH3 domains.

The cytoplasmic PTKs of the Src family contain, at the amino terminus, an SH2 and an SH3 domain and their activity and target-specificity is influenced by inter- and intramolecular interactions mediated by these domains (Mayer & Baltimore, 1994; Mayer *et al*., 1995; Superti-Furga, 1995). The fine interplay between these two, in principle distinct, determinants (enzymatic and binding specificity) makes prediction of natural protein kinase substrate a difficult task. A feasible and productive approach involves studying the two factors separately. This has been pioneered by Songyang and collaborators (Songyang *et al*., 1993; 1994; 1995a) who developed an approach that makes use of synthetic peptide repertoires containing either tyrosine or phosphotyrosine at specific positions in an otherwise random peptide sequence.

Here we have shown how peptide libraries displayed by filamentous phage can be utilised for a similar scope. The main advantage of the phage approach is that this method does not produce a consensus sequence of a large pool of peptides that share a common property, but rather yields a list of peptide sequences each of which is endowed with the same property. One or more consensus sequences can then be determined by aligning and comparing the amino acid sequences of the selected clones.

A second consequence of these considerations is that, in the search of peptides that bind with high affinity a given target, the synthetic peptide approach is restricted by the need of limiting the selection stringency in order to recover sufficient material for amino acid sequence determination by Edman degradation. By contrast, this limitation is not encountered in the phage display approach since, in principle, even one out of several million clones can be isolated, by repeated enrichment cycles, and subsequently propagated for characterization.

Because of technical reasons synthetic peptide repertoires do not contain C and W and the specific library that was utilized to determine tyrosine kinase specificity did not contain Y, S and T (Songyang *et al*., 1995a). On the contrary amino acid composition of phage peptide libraries is only biased by the degeneracy of the genetic code. This can be minimized by utilization of random codons of the type NNK where K is either G or T. It is worth mentioning, however, that, in our libraries, we found a strong bias against display of peptides containing a single C (Iannolo *et al*., 1995).

Finally the phage display approach requires simple technology that is available practically in any laboratory, even in those that might not have access to equipment for peptide synthesis and sequencing.

Recently, while this manuscript was in preparation, Gram and collaborators (Schmitz *et al*., 1996) have independently proposed a similar approach to study substrate specificity in protein tyrosine kinases. Aside from technical details (see Material and Methods) the main differences with our approach reside in the utilization of pVIII instead of pIII as the carrier capsid protein and in the alternate selection with two different anti-phosphotyrosine antibodies. We found the latter to be necessary to avoid selection of peptides that do not contain tyrosine, whenever the kinase reaction is carried out only for a short time to favour phosphorylation of substrates with a low K_m. In fact we found that, in conditions of poor library phosphorylation, both antibodies select peptides that define two consensus amino acid sequences that do not contain Y (WLDAR and PWXGTT). This finding might have some relevance in the interpretation of peptide phosphorylation patterns obtained by blotting protein extracts with different anti-pY antibodies. In any case this stratagem offers the further advantage of neutralysing the impact of a second possible artifact of the selection procedure. In fact, in principle, utilizing a single anti-pY mAb could result in a consensus sequence that is the weighted sum of the kinase and the antibody sequence context preferences.

In addition we have shown that, by extending the kinase reaction time, the sequence context specificity is weakened and practically any tyrosine containing peptide is phosphorylated. Thus it is possible to assemble a modified peptide library where all the tyrosine residues are changed *in vitro* into phosphotyrosine. This offers a powerful tool to study protein-protein interaction mediated by phosphotyrosine containing peptides. In principle, the concept of modified peptide libraries is not limited to phosphorylation of tyrosine; serine/threonine phosphorylation, glycosylation, ubiquitination or any other modification that can be carried out *in vitro* and that is suspected to modulate protein-protein interaction may be used to build different types of modified libraries.

We have tested the feasibility of this approach by selecting peptides that are efficiently phosphorylated by the kinase domain of the Fyn PTK. Furthermore, as a test for the phosphorylated library we have selected phosphotyrosine containing peptides that bind the PTB domain of Shc.

E– φTYGXφ is the consensus sequence that defines peptides that are efficiently phosphorylated by the Fyn PTK. Cheng *et al*. (1991) reported that a synthetic peptide derived from $p34^{cdc2}$, EKIEGT-pYGVVYK, is a specific and efficient substrate of Src-related Fyn kinase. The sequence of this peptide is in good agreement with the consensus sequence that we have determined. Evidence that the peptide sequence, EKIEGTpYGVVYK, contains the structural features important for substrate activity,

were derived from studies on peptide analogues with single substitutions. By this approach it was shown that T in position −1 is an important activity determinant (Cheng *et al*., 1991).

The synthetic peptide approach (Songyang *et al*., 1995a) allowed definition of the consensus substrate amino acid sequences that, for instance, discriminate between receptor PTKs and cytoplasmic PTKs. It was not possible, however, to reveal subtle differences among members of the same group. For instance, by this approach, the suggested consensus peptide of the enzyme Fes, that is not a Src-like kinase, is similar to that of c-Src and v-Src. This is not in agreement with the finding that these enzymes have very different activity when incubated with the $p34^{cdc2}$ substrate peptide (UBI technical report).

In a recent report, phage display technology helped in defining specific consensi for the Blk, Lyn, c-Src and Syc PTKs (Schmitz *et al*., 1996).

Our results on Fyn confirm a preference for a negative charge, amino-terminal to pY and for a hydrophobic residue at position +3. By contrast the preference for T at position −1 is not found in the preferred substrates of the other PTKs and identifies Fyn substrates. Furthermore the peptide cdc2 is a very poor substrate for c-Src, Blk and Lyn, while our experiments confirm that this peptide contains the structural determinants of Fyn specificity.

Searching the protein data libraries with a pattern, defined by the consensus Fyn substrate that we have determined, could help in identifying proteins that are candidate functional substrates. It must be stressed, however, that kinase substrate choice can be influenced by co-localization mediated by SH3 and SH2 domains. We have mimicked this mechanism, *in vitro*, by showing that phosphorylation of peptides, that are also targets for the SH2 domain, is stimulated by multiple display on the phage surface, probably because the SH2 domain promotes processive phosphorylation. It is possible that processivity plays a role *in vivo* in the case of phosphorylation by Fyn of the three ARAMs (Antigen Recognition Activation Motives) following activation of T cell receptor (Weiss, 1993).

By carrying out phosphorylation of phage peptide libraries in saturation conditions, we have assembled a new type of modified library that can be exploited to approach protein-protein interaction mediated by phosphotyrosine containing peptides. We have already reported that this approach can be utilized to uncover subtle differences between the recognition specificity of the SH2 domains of two proteins related to Shc (Pelicci *et al*., 1996). The different phosphotyrosine sequence contexts that were selected for instance with the SH2 domains of Shc and Grb confirms that the bias imposed on the method by the kinase substrate specificity, if any, is not such to influence the consensus sequence of the binding peptide. Nevertheless, to further reduce any bias, the phage library could be phosphorylated with a cocktail of two or more PTKs having different substrate specificity.

Here we have described the selection of phosphorylated peptides that bind the PTB domain of Shc. The consensus peptide that was deduced F/YxNpTYxxY/W confirms that the N at position −3 is the major recognition determinant, as concluded from the comparison of sequence context in natural substrates (reviewed by van der Geer & Pawson, 1995) or by selection of a degenerate phosphotyrosine containing peptide library (Songyang *et al*., 1995b). The P at −2, although found in all known natural substrates, is not essential for PTB binding. In contrast with the results obtained with synthetic peptides (Songyang *et al*., 1995b) our results point to the importance of T at position −1. Threonine at that position is found in the middle T antigen peptide 250-258, which is the natural peptide with the highest known affinity for the Shc PTB (Wolf *et al*., 1995). Furthermore substituting T with an A at −1 results in a dramatic decrease in affinity (Trüb *et al*., 1995).

Our method selects peptides that bind with high affinity when they are presented linear and unconstrained. Natural PTB binding peptides, presumably, have been subjected to different selective pressures depending on their function and structural context. The net result is not necessarily the highest affinity binding peptide. These considerations are likely to explain why our consensus peptides define an extended consensus with preferred amino acids at positions that are found to be neutral in natural targets. See for instance the preference for N and hydrophobic residues on the amino side of some Shc SH2-binding peptides or for aromatic side-chains at position +3 in the PTB consensus (Table 5). Furthermore, unlike its natural counterpart, a linear peptide without any structural constraints imposed by the host protein scaffold, can freely adapt along its entire length over the domain binding surface.

Although this put some restraint on the utilization of the consensus, obtained by this method, as a tool to predict natural targets for phosphotyrosine binding domains, at the same time it stresses the potential of this approach to isolate high affinity binders that may be used *in vitro* and/or *in vivo* to compete physiologically important protein-protein interactions.

Materials and Methods

Enzymes, antibodies and peptides

$p55^{fyn}$, purified from bovine spleen, was purchased from Upstate Biotechnology Inc; pY mAb "alpha1" (P-Tyr-1) was purified on MabTrap (Parmacia, LKB, Upsala, Sweden) from ATCC-CRL-1955 hybridoma (a gift from C.T. Baldari); Mab PY20 was purchased from Transduction Laboratory (Kentucky). Phosphorylated synthetic peptide EPQpYEEIPI, corresponding to Src SH2 high affinity peptide, was kindly provided by G. Superti-Furga. The remaining peptides described in Results were syn-

thesized and HPLC purified at 80% purity grade, by Genosys Biotecnologies, Cambridge.

Phage libraries and GST fusions

For phosphorylation and panning experiments, we have utilized a pVIII phagemid library that displays nonapeptide of random sequence after the AEGEF amino-terminal peptide of the major capsid protein (Felici *et al.*, 1991).

The expression plasmid directing the synthesis of Shc PTB-GST fusion protein (amino acids 1 to 232 of Shc protein) was obtained by PCR amplification of the corresponding cDNA fragment and subcloning into the pGEX471 vector.

Phage phosphorylation

Phage particles were collected from culture supernatant by a double precipitation with 20% polyethylene glycol 8000, 2.5 M NaCl and finally dissolved in 1/100 of the initial culture volume (titre $\approx 10^{13}$ particles/ml).

A standard Fyn PTK reaction was carried out for five hours at 30°C in 10 μl volume of assay buffer (50 mM Tris-HCl (pH 7), 50 mM $MgCl_2$, 50 mM Na_3VO_4, 500 μM ATP) containing 3.5 units of Fyn and about 10^{10} phage particles. In the panning experiments, that aimed at determining the sequence of the Fyn consensus substrate, the incubation time was reduced to three minutes to favour the phosphorylation of the best substrates.

[γ-^{32}P]ATP was included in the phosphorylation reaction as a tracer in order to estimate phage phosphorylation efficiency. Protein bands corresponding to pVIII were separated from other phage capsid proteins by SDS-Tricine PAGE (Schägger & von Jagow, 1987) excised and counted.

Kinetic analysis on synthetic peptides

Fyn activity was monitored at different peptide concentrations: 3 μM, 9 μM, 15 μM, 30 μM, 90 μM, 150 μM, 300 μM. The assay was carried out in the same buffer conditions used for phage phosphorylation, final concentration of [γ-^{32}P]ATP was 100 μM with a specific activity of 1 μCi/nmol. The reaction was terminated after ten minutes by TCA precipitation and an aliquot was spotted onto PVDF membrane (BIORAD); after washing in 0.75% phosphoric acid, the amount of incorporated radioactivity was measured.

Panning

Phosphorylated peptide libraries were exposed to selection cycles by standard panning methods as previously described (Cesareni *et al.*, 1995). Briefly, polystyrene beads or wells of polystyrene microtitre dishes were incubated with 10 mg/ml of purified anti-pY mAbs or protein domains fused to GST (Pelicci *et al.*, 1996). The number of panning cycles necessary to select specific phage clones was determined by monitoring the ratio between the titre of the phage eluted from the beads at pH 2 and the titre of the imput phages

Selected phages were picked at random, single strand DNA purified and the relevant region of the phagemid pVIII gene sequenced by the dideoxy chain termination method (Sanger *et al.*, 1977).

Micropanning control experiments were performed by binding to anti-pY antibodies 10^8 phage particles of each selected clone, after or before phosphorylation.

Acknowledgements

This work was supported by a grant from ''Associazione Italiana Ricerca sul Cancro''to G.C. and P.G.P. We thank C. Baldari for the hybridoma cell line producing the anti-pY mAb alpha 1, S. Vendetti for help in the production of hybridoma supernatants, O. Minenkova for the 15 aa peptide library, G. Superti-Furga for the EPQ-pYEEIPI peptide.

References

Alonso, G., Koegl, M., Mazurenko, N. & Courtnedge, S. A. (1995). Sequence requirements for binding of Src family tyrosine kinases to activated growth factor receptors. *J. Biol. Chem.* **270**, 9840–9848.

Casnellie, J. E., Harrison, M. L., Pike, L. J., Hellstrom, K. E. & Krebs, E. G. (1982). Phosphorylation of synthetic peptides by a tyrosine protein kinase from the particulate fraction of a limphoma cell line. *Proc. Natl Acad. Sci. USA,* **79**, 282–286.

Cesareni, G., Castagnoli, L., Dente, L., Iannolo, G., Vetriani, C., Felici, F., Luzzago, A., Monaci, P., Nicosia, A. & Cortese, R. (1995). Construction and utilization of peptide libraries displayed by filamentous bacteriophage. In *Immunological Recognition of Peptides in Medicine and Biology,* pp. 43–59, CRC Press.

Cheng, H., Litwin, C. M. E., Hwang, D. M. & Wang, J. H. (1991). Structural basis of specific and efficient phosphorylation of peptides derived from p34cdc2 by a pp60src-related protein tyrosine kinase. *J. Biol. Chem.* **266**, 17919–17925.

Cheng, H. C., Nishio, H., Hatase, O., Ralph, S. & Wang, J. H. (1992). A synthetic peptide derived from p34 cdc2 is a specific and efficient substrate of src-family tyrosine kinases. *J. Biol. Chem.* **267, n:13**, 9248–9256.

Felici, F., Castagnoli, L., Musacchio, A., Jappelli, R. & Cesareni, G. (1991). Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector. *J. Mol. Biol.* **222**, 301–310.

House, C., Baldwin, G. S. & Kemp, B. A. (1984). Synthetic peptide substrates for the membrane tyrosine protein kinase stimulated by epidermal growth factor. *Eur. J. Biochem.* **140**, 363–367.

Hunter, T. (1982). Synthetic peptide substrates for a tyrosine protein kinase. *J. Biol. Chem.* **257**, 4843–4848.

Iannolo, G., Minenkova, O., Petruzzelli, R. & Cesareni, G. (1995). Modifying filamentous phage capsid: Limits in the size of the major capsid protein. *J. Mol. Biol.* **248**, 835–844.

Kavanaugh, W. M. & Williams, L. T. (1994). An alternative to SH2 domains for binding tyrosine-phosphorylated proteins. *Science,* **266**, 1862–1865.

Kavanaugh, W. M., Turck, C. W. & Williams, L. T. (1995). PTB domainbinding to signaling proteins through a sequence motif containing psphotyrosine. *Science,* **268**, 1177–1179.

Kohl, N. E., Mosser, S. D., De Solms, J., Giuliani, E. A., Pompliano, D. L. & Gibbs, J. B. (1993). Selective inhibition of ras-dependent transformation by a farnesyltransferase inhibitor. *Science*, **260**, 1934–1937.

Mayer, B. J. & Baltimore, D. (1994). Mutagenic analysis of the roles of SH2 and SH3 domains in regulation of the abl tyrosine kinase. *Mol. Cell. Biol.* **14**, 2883–2894.

Mayer, B. J., Hirai, H. & Sakai, R. (1995). Evidence that SH2 domains promote processive phosphorylation by protein-tyrosine kinase. *Curr. Biol.* **5**, 296–305.

Parmley, S. F. & Smith, G. P. (1988). Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene*, **73**, 305–318.

Pawson, T. (1995). Protein modules and signalling networks. *Nature*, **373**, 573–580.

Pelicci, G., Lanfrancone, L., Grignani, F., McGlade, J., Cavallo, F., Forni, G., Nicoletti, I., Grignani, F., Pawson, T. & Pelicci, P. G. (1992). A novel transforming protein (SHC) with an SH2 domain is implicated in mitogenic signal transduction. *Cell*, **70**, 93–104.

Pelicci, G., Dente, L., De Giuseppe, A., Verducci-Galletti, B., Giuli, S., Mele, S., Vetriani, C., Giorgio, M., Pandolfi, P. P., Cesareni, G. & Pelicci, P. G. (1996). A family of Shc related proteins with conserved PTB, CH1 and SH2 regions. *Oncogene*, **13**, 633–641.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.

Schägger, H. & von Jagow, G. (1987). Tricine-sodium dodecyl sulphate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368–379.

Schmitz, R., Baumann, G. & Gram, H. (1996). Catalytic specificity of phosphotyrosine kinases Blk, Lyn, c-Src and Syk as assessed by phage display. *J. Mol. Biol.* **260**, 664–677.

Scott, J. K. & Smith, G. P. (1990). Searching for peptide ligands with an epitope library. *Science*, **249**, 386–390.

Songyang, Z., Shoelson, S. E., Chadhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J., Neel, B. G., Birge, R. B., Fajardom, J. E., Chou, M. M., Hanafusa, H., Schaffhausen, B. & Cantley, L. C. (1993). SH2 domains recognize specific phosphopeptide sequences. *Cell*, **72**, 767–778.

Songyang, Z., Shoelson, S. E., McGlade, J., Olivier, P., Pawson, T., Rustello, X. R., Barbacid, M., Sabe, H., Hanafusa, H., Yi, T., Ren, R., Baltimore, D., Ratnofsky, S., Feldman, R. A. & Cantley, L. C. (1994). Specific motifs recognised by the SH2 domains of Csk, 3BP2, fps/fes, Grb-2, HCP, SHC, Syk and Vav. *Mol. Cell. Biol.* **14**, 2777–2785.

Songyang, Z., Carraway III, K. L., Eck, M. J., Harrison, S. C., Feldman, R. A., Mohammadi, M., Schlessinger, J., Hubbard, S. R., Smith, D. P., Eng, C., Lorenzo, M. J., Ponder, B. A. J., Mayer, B. J. & Cantley, L. C. (1995a). Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature*, **373**, 536–539.

Songyang, Z., Margolis, B., Chaudhuri, M., Shoelson, S. E. & Cantley, L. C. (1995b). The phosphotyrosine interaction domain of SHC recognizes tyrosine-phosphorylated NPXY motif. *J. Biol. Chem.* **270**, 14863–41866.

Superti-Furga, G. (1995). Regulation of the Src protein tyrosine kinase. *FEBS Letters*, **369**, 62–66.

Trüb, T., Choi, W. E., Wolf, G., Ottinger, E., Chen, Y. J., Weiss, M. & Shoelson, S. E. (1995). Specifity of the PTB domain of SHC for ß turn-forming pentapeptide motifs amino terminal to phosphotyrosine. *J. Biol. Chem.* **270**, 18205–18208.

van der Geer, P. & Pawson, T. (1995). The PTB domain: a new protein module implicated in signal transduction. *Trends Biochem. Sci.* **20**, 277–280.

Weiss, A. (1993). T cell antigen receptor signal transduction: a tale of tails and cytoplasmic protein-tyrosine kinases. *Cell*, **73**, 209–212.

Wolf, G., Trub, T., Ottinger, E., Groninga, L., Lynch, A., White, M. F., Miyazaki, M., Jongsoon, L. & Shoelson, S. E. (1995). PTB domains of IRS1 and Shc have distinct but overlapping binding specificities. *J. Biol. Chem.* **270**, 27407–27410.

Edited by J. Karn

(Received 4 February 1997; received in revised form 21 March 1997; accepted 4 April 1997)