Use of random peptide phage-displayed libraries for studying protein phosphorylation and phosphotyrosine-dependent protein-protein interactions

✍ Luciana Dente,[1,2] Gianni Cesareni[1]

[1]Department of Biology "Enrico Calef", University of Rome Tor Vergata, Via della Ricerca Scientifica, 00133 Rome, Italy. [2]Department of Physiology and Biochemistry, University of Pisa, via Carducci, 56010 Pisa, Italy. E-mail: dente@bio.uniroma2.it

ABSTRACT

Protein-phosphorylation is a post-translation modification that alters the binding or the enzymatic properties of the modified proteins. In particular, phosphorylation of tyrosine residues is a crucial event in protein-protein interactions occurring during intracellular signal transduction. Several approaches have been used to study these interactions. Combinatorial libraries have proven to be a valuable tool to characterize ligand-target interactions and to analyze binding specificity. Phage displayed libraries offer the advantage of coupling the capsid exposed ligand to its coding sequence, inserted in the phage genome. Furthermore, peptides displayed on phage capsid can be subjected to enzymatic modification by incubation with appropriate enzymes. Therefore, using a receptor specific for the modified products it is possible to select clones bearing the peptides with the desired properties. Here we describe a method to generate "dedicated" libraries, displaying phosphotyrosine containing peptides, using cytosolic phosphotyrosine kinases (PTKs). These are specialized enzymes, implicated in several events of transduction of external signals into the cells. Random peptide phage libraries can be phosphorylated to different extent to assess the PTK substrate requirements or to analyze phosphotyrosine-dependent protein-protein interactions. This approach is useful for studies aimed at understanding the rules underlying protein interaction mechanisms and to design specific inhibitors that could compete in these interactions.

Keywords: phage display, protein phosphorylation, libraries

Biotecnología Aplicada 2003;20:134-138

RESUMEN

Empleo de bibliotecas de péptidos presentados en la superficie de fagos para el estudio de la fosforilación de proteínas y de la interacción proteína-proteína dependiente de fosfotirosina. La fosforilación de proteínas es una modificación post-traduccional que altera las propiedades enzimáticas o de unión de las proteínas modificadas. En particular, la fosforilación de los residuos de tirosina es un evento importante en las interacciones proteína-proteína que ocurren durante la transducción intracelular de señales. Para el estudio de estas interacciones se han utilizado varias metodologías. Las bibliotecas combinatorias han demostrado ser una herramienta valiosa para caracterizar las interacciones entre un blanco y su ligando, así como analizar la especificidad de la unión. Las bibliotecas presentadas en la superficie de los bacteriófagos presentan la ventaja del acoplamiento del ligando expuesto en la cápsida y su secuencia codificante, insertada en el genoma del fago. Además, los péptidos presentados en la superficie del fago pueden sufrir modificaciones enzimáticas mediante la incubación con las enzimas apropiadas. Por lo tanto, si se usa un receptor específico para los productos modificados es posible seleccionar clones de fago que portan péptidos con las propiedades deseadas. Este trabajo describe un método para generar bibliotecas "especializadas" que presentan péptidos con residuos de fosfotirosina, usando fosfotirosinas-quinasas citosólicas (PTK). Estas son enzimas especializadas, implicadas en varios procesos de transducción de señales externas al interior celular. Las bibliotecas de péptidos aleatorios presentadas en fagos pueden ser fosforiladas en diferentes grados para evaluar los requerimientos de sustrato de las PTK o para analizar las interacciones proteína-proteína dependientes de fosfotirosina. Esta estrategia es muy útil para estudios encaminados al entendimiento de las reglas que fundamentan los mecanismos de interacción de proteínas y para el diseño de inhibidores específicos que sean capaces de competir con estas interacciones.

Palabras Claves: exposición en la capsida de los fagos, fosforilación de proteinas, bibliotecas

*I*ntroduction

Random peptide libraries displayed on the surface of filamentous phage represent a complete repertoire of where to find simple ligands for most ligates. Linear peptides, affinity purified by binding to a given bait immobilized to a solid support, mimic the binding properties of the natural ligand, even if this is a folded protein domain or a non proteinaceous molecule such as carbohydrates.

Thus, the information provided by the selected peptides is relevant to the characterization of the binding mode of the natural ligand. An extension of this approach allows the use of peptide libraries to determine the substrate specificity of peptide-modifying enzymes. The random library is incubated with the enzyme of interest and clones that have been modified are enriched by affinity selection against a specific receptor, which recognizes products of the enzymatic reaction.

This method was first exploited for identifying specific protease substrates by Matthews and Wells [1], who constructed a "dedicated" library, inserting a randomized protease substrate sequence between a bind-

✍ Corresponding author

1. Matthews DJ and Wells JA (1993) Substrate phage: selection of protease substrates by monovalent phage display. Science 260, 1113-1117.

ing domain for an affinity support and the M13 pIII protein sequence. Phages were then fixed to the affinity support and treated with the protease of interest: phages with good protease substrates were released, whereas phages with substrates that resisted proteolysis remained bound. Similar approaches have been extensively used to search for sensitive or resistant substrates for several other enzymes. Here we review the exploitation of these techniques to study the specificity of tyrosine phosphorylation and of protein-protein interactions mediated by phosphotyrosines. Random peptide phage libraries have been used to study the substrate specificity of cytosolic protein tyrosine kinases, PTKs. These enzymes are grouped in two Src-like and non-Src families on the basis of the homology to the viral Src protein, known for its trasforming activity and its association to several malignancies [2]. *In vivo* PTK are tightly regulated: inter and intramolecular interactions modulate the catalytic activity [3]; the choice of the substrate is usually determined by the residues surrounding the tyrosine that has to be phosphorylated [4, 5, 6]. Generally, PTK are very specific in recognizing the substrates in the right context inside the cells, but they become less specific if the reaction conditions are altered. We exploited this property to generate a «completely phosphorylated» library, where all the tyrosine residues are phosphorylated, regardless the surrounding residues [6, 7]. This kind of library can be used to study target recognition in protein-protein interactions, mediated by phosphotyrosine (pY) binding domains. We have demonstrated the power of the technique by analyzing the binding potential of some Src-homology 2 (SH2) and a phosphotyrosine-binding (PTB) domain [6, 7]. Recognition specificity is always modulated by the specific context in which the phosphotyrosine is embedded in the protein target. The analysis of a random repertoire of sequences will enable to understand the rules underlying the recognition mechanisms.

General Considerations

Library source

The construction of random peptide libraries displayed on filamentous phages relies on the genetic fusion of oligonucleotides of a random sequence to the coding sequence of a capsid protein, usually the receptor protein pIII [8] or the major coat protein pVIII [9]. The two fusion methods, given the different amount of the two proteins on each viral capsid (3-5 copies for pIII and about 3000 copies for pVIII) may be alternatively used depending on the goal. The types of libraries used in the studies described here are summarized in Table 1. The presence of a tyrosine (Y) in a fixed position in the peptides («dedicated» library) is not a prerequisite for a successful screening as totally random libraries, displayed either on pIII or pVIII proteins, yielded similar results.

Phosphorylation of peptide phage libraries

We use two different approaches to phosphorylate phage peptide libraries.

1) One approach entails a short incubation of the phages with the PTK: the reaction time and the amount of enzyme has to be carefully titrated to set up the conditions where most phages displaying tyrosine-containing peptides are not labeled at all, while only phages displaying a «good peptide substrate» are phosphorylated. This property can be used to isolate specific substrates for purified TK enzymes (see below).

2) The alternative approach entails incubation of the library for a longer reaction time or at higher enzyme concentration: under these conditions almost all the tyrosine containing peptides are phosphorylated, irrespectively of their amino acid sequence context. The peptide library becomes a «fully modified-library», where virtually every phage displaying a peptide containing a Y is phosphorylated. This kind of library can be exploited to select peptides, which are recognized by phosphotyrosine binding domains.

Table 1. Peptide phage libraries used in Y-phosphorylation experiments. The first two libraries are "dedicated": a tyrosine (Y) is fixed in central position; x indicates each of the 20 possible amino acids. The N-terminus corresponds to the vector sequence preceding the inserted peptide.

Library (vector)	N-terminus	Displayed peptide	Carrier protein	Reference
8mer-Y (pGEM3)	MAELE	xxxYxxxx	(pIII)	[5]
15mer-Y (pC89)	AEGEF	xxxxxxxYxxxxxxx	(pVIII)	[16]
7mer (F1)	AEGEF	xxxxxxx	(pVIII)	[7]
9mer (pC89)	AEGEF	xxxxxxxxx	(pVIII)	[16]

Affinity purification of phosphorylated phages

The two approaches described above require different selection procedures (Figure 1):

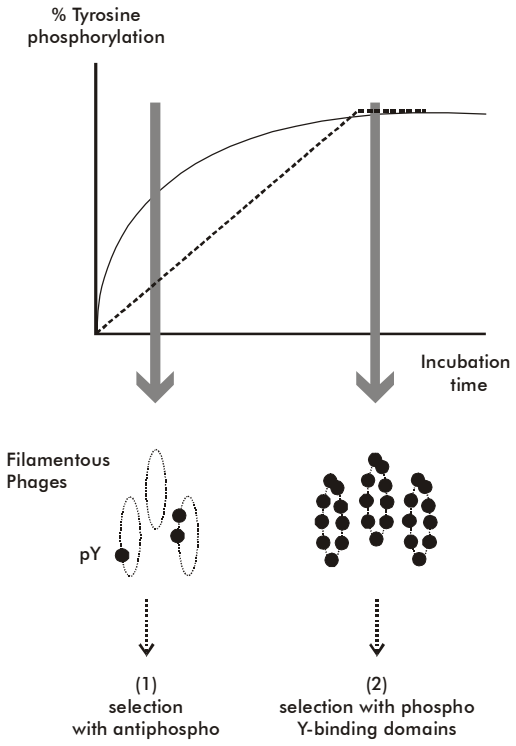

Figure 1. Schematic representation of the phosphorylation kinetic of a phage library: After a short incubation time only few phages displaying the best substrate peptides are phosphorylated (line ——): they can be isolated by selection with anti-pY antibodies (1). At the end of the reaction all the Y-containing peptides will be phosphorylated (line ▪▪▪▪): from this population, distinct phosphoY-binding domains (2) select their specific targets on the bases of the amino acids surrounding the phospho-Y. Filamentous phages are represented as dotted ovals.

2. Cooper JA (1990) The Src family of protein-tyrosine kinases. In Kemp B. (ed.) Peptides and Protein Phosphorylation. CRC Press, Boca Raton, FL, 85-113.

3. Superti-Furga G and Courtneidge S.A. (1995) Structure-function relationships in Src family and related protein tyrosine kinases. Bioessays 17, 321-330.

4. Songyang Z, and Cantley LC, (1995) Recognition specificity in protein tyrosine kinase-mediated signaling. Trends Biochem. Sci. 20, 470-475.

5. 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.

1) Monoclonal antibodies raised against phosphotyrosine can be used for discriminating substrate from non-substrate phages, after mild phosphorylation of the library. However, it is important to take into consideration that some of the available anti-phosphotyrosine antibodies are sensitive to sequence variations that are near the phosphorylated residue or, also, recognize peptides that do not contain phosphotyrosine. For example, we have shown that two anti-pY mAbs, pY-alpha 1 and pY-20 respectively identify the recognition consensus PWXGTT and WLDAR from a non-phosphorylated non-peptide library [6]. These represent a second specificity that prevails over the common pY specificity when the library is poorly phosphorylated. It interferes with the selection procedure when multivalent pVIII libraries are used, while it has not been reported for pIII library selections [5].

We developed a protocol for pVIII library screening, involving sequential panning cycles with the two antibodies to favor the selection of the few phages displaying peptides that are phosphorylated in stringent conditions. In this way the specificity for non phosphorylated peptides does not prevail and the pool of selected phages is typically retained in at least two orders of magnitude better than the background by any of the immobilized anti pY mAbs [6].

2) "Fully phosphorylated" libraries, obtained as described above, can be exploited to select ligands of pY-binding proteins. These proteins usually contain small binding modules, specialized for pY recognition (SH2 and PTB domains) that even when isolated from their original context and expressed in bacteria as recombinant proteins retain their binding potential [10]. The specificity of the recognition of SH2 domains is generally provided by 2-3 residues following the pY [11], while few residues at the amino-terminal side of the phosphotyrosine are determinants for PTB binding [12]. We have shown that screening phosphorylated random peptide libraries uncovers the differences between the recognition specificity of several SH2 and PTB domains [6, 7]. The coding regions of these protein domains were fused to the Glutathione S-transferase (GST) gene in the pGEX2T plasmid (Pharmacia). This vector provides the production of a large amount of recombinant protein that can be easily recovered from bacterial extract by purification on glutathione agarose. Recombinant proteins eluted with glutathione were linked to a solid support, such as polystyrene beads or again glutathione agarose. In case of low protein yield or unstable proteins that are difficult to purify, it is possible to avoid this step and to use GST-fused proteins, still linked to the glutathione-agarose, without eluting them [13]. This approach is less time-consuming, but can increase the background level. Bound phages can be recovered either by drastic pH changes (as in usual phage-display protocols) or by direct transduction in bacteria.

Protocol 1. Determination of TK substrate specificity

Phage phosphorylation

- Phage particles are collected from culture supernatant by a double precipitation with 20% PEG (polyethylene glycol 8000), 2.5 M NaCl and finally dissolved in 1/100 of the initial culture volume (Titre 10^{13} particles/mL).

- Conditions of phosphorylation depend upon the chosen PTK. As an example: phosphorylation reaction using Fyn (TK $p55^{fyn}$; Upstate Biotechnology Inc) was carried out at 30 ºC 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} phages particles (10 µL).

- Incubation was initiated by the inclusion of the enzyme. After 3 minutes the reaction was terminated by incubation at 70 ºC for 5 minutes. The incubation time was empirically determined to favour the phosphorylation only of the "best" substrates.

- Optionally, $^{32P}\gamma$-ATP was included in the phosphorylation reaction as a tracer in order to estimate phage phosphorylation efficiency. In any case, a radioactive band of tyrosine phosphorylated pVIII proteins can be detected by SDS-Tricine PAGE electrophoresis [14] and autoradiography (see fig.1 in [6]) only after the first enrichment cycle. Before amplification it is not possible to detect the few pVIII molecules that are preferentially phosphorylated in a complex library such as the random non-peptide-pVIII library that we have been using [9].

Panning phosphorylated libraries (affinity selection)

- Polystyrene beads or wells of ELISA-treated microtiter plates or any other solid support, such as Sepharose [15] are incubated with 10 µg/mL of purified anti-phosphotyrosine antibody in PBS (phosphate-buffered saline: 37mM NaCl, 3mM KCl, 8mM Na_2HPO_4, 1.5mM KH_2PO_4) at 4 ºC overnight.

- Unbound proteins are removed by several washings with PBS-0.5 % Tween.

- Blocking is performed with PBS-0.5%Tween - 3% BSA (bovine serum albumin) at 4 ºC for 4 hours.

- After removing the blocking solution, phosphorylated phages (about 1 µL: 10^9 transduction units) in PBS-1%/BSA are incubated with the target at 4 ºC (1 hour to overnight).

-Unbound phages are collected for subsequent titration (fraction N).

- Several washings with PBS-0.5%Tween are performed to eliminate non-specific bound phages.

- Bound phages elution (fraction A) is obtained by adding 50 mM glycine-HCl (pH2.2) at 37 ºC for 10 min. The solution is neutralized by adding an equal volume of 0.2 M Tris/HCl pH 9.

-Alternatively, fresh bacteria are added to resin-bound phages for 10 minutes at 37 °C; incubation is prolonged for another 20 minutes after the L- broth inclusion. Finally, transducted bacteria are selected on antibiotic plates.

- Fractions N and A are titrated: the rate between the two values (N/A) is generally about 10^5 after the first panning.

- Amplification of output phages (fraction A) is obtained following standard phage protocols.

Repeated cycles of amplification, phosphorylation and selection with anti-pY antibodies are necessary to select specific phages (increment of the N/A rate allows to monitor the relative enrichment). To overcome

6. Dente L, Vetriani C, Zucconi A, Pelicci G, Lanfrancone L, Pelicci PG, Cesareni G (1997) Modified phage peptide libraries as a tool to study specificity of phosphorylation and recognition of tyrosine containing peptides. J. Mol. Biol. 269, (5), 694-703.

7. Pelicci G, Dente L, De Giuseppe A, Verducci-Galletti B, Giuli S, Mele S, Vetriani C, Giorgio M, Pandolfi PP, Cesareni G, Pelicci PG (1996) A family of Shc related proteins with conserved PTB, CH1 and SH2 regions. Oncogene12, 633-641.

8. Scott JK and Smith GP (1990) Searching for peptide ligands with an epitope library. Science 249, 386-90

9. Felici F, Castagnoli L, Musacchio A, Jappelli R, and 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.

10. Pawson T (1995) Protein modules and signaling networks. Nature 373, 573-580.

11. Songyang Z, et al. (1993) SH2 domains recognize specific phosphopeptide sequences Cell 72, 767-778.

12. Kavanaugh WM, Turek CW, Williams LT. PTB domain binding to signaling proteins through a sequence motif containing phosphatyrosine. Science, 1995;268: 1177-9.

13. Tommasi di Vignano A, Di Zenzo G, Sudol M, Cesareni G, Dente L (2000) Contribution of the different modules in the utrophin carboxy-terminal region to the formation and regulation of the DAP complex. FEBS letters 471, 229-234.

14. Schägger H and 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.

15. Sparks AB, Adey NB, Cwirla S, and Kay BK (1996) Screening phage-displayed random peptide libraries. In Phage Display of Peptides and Proteins (Academic Press, Inc.) 227-253.

selection problems deriving from the antibody-specificity for residues surrounding the phosphorylated tyrosine, we suggest to use at least two different antibodies and to alternate them during panning cycles. To identify $p55^{fyn}$ substrates we used three amplification/selection cycles involving alternate selections with mAb pY20 (Transduction Laboratory, Kentucky), followed by MAb alpha 1 (from ATCC-CRL-1955 hybridoma) and finally again pY20.

Peptide sequence deduction

Selected phages are picked at random, single strand DNA is purified [16] and the relevant region of the phagemid pVIII gene is sequenced by the dideoxy chain termination method. Peptide sequences are aligned and the frequency of conserved amino acids surrounding the pY is determined.

Protocol 2. Determination of pY-Binding Modules specificity

Phage phosphorylation

The phosphorylation protocol is similar to the one described above, except that the reaction has to be prolonged to five hours in order to phosphorylate all the tyrosine containing peptides displayed on phages. The two tyrosines that are present in the pVIII protein are buried in the portion of the protein that is in contact with the DNA and it is not exposed to the solvent

Target purification

We found that GST-fusion proteins, obtained by cloning the coding region of the target of interest in one of the expression vectors of the pGEX family (Pharmacia) work very well in our panning experiments. Other fusion systems as His-tagged or Maltose-binding protein chimeras usually give a higher background.

Following the producer's protocols it is possible to recover about 5 µg of recombinant protein from 1 mL of bacterial culture. After purification on glutathione agarose, a small amount of resin-linked protein, resuspended in loading dye is analyzed by SDS gel electrophoresis to control the yield and the purity of the preparation. Normally the protein is sufficiently purified from bacterial contaminants to be directly used for panning selections without the need to elute it from the resin.

Panning

- Approximately 2-4 µg of bait protein linked to the resin and 10^9 transduction units from the phosphorylated library (1 µL) in 100 µL PBS, 3% bovine serum albumin (BSA) are incubated at 4 ºC (1 hour to overnight).

- Blocking, washing and final elution is carried out as described above for antibody selection.

- The number of panning cycles needed to select specific phage clones is determined by measuring the relative enrichment. The ratio between the titre of the imput phages and the titre of the eluted phages (N/A) is monitored at each cycle.

- When this value increases by at least two orders of magnitude above the background level, independent phage clones are isolated. Binding potential can be eventually confirmed by Phage ELISA [17] and selected phages are finally sequenced to derive the peptide sequences.

Discussion

The dramatic effects caused in cancer by the altered regulation of these post-translation modifications proves the importance of phosphorylation events during different stages of cell proliferation and differentiation (review:[18]).

The identification of the substrates or substrate motifs of protein kinases and of p-Y binding proteins is extremely important to understand cellular functions and to design specific inhibitors to block altered protein interactions. Biochemical studies of the physiological targets of protein kinases together with the resolution of the some tridimensional structures or homology-based molecular models suggest that linear peptides can mimic natural substrates either of catalytic PTK domains or pY-binding modules. The use of the random peptide library approach represents a fast, efficient and specific method to identify substrate motifs [4, 11]. Phage display libraries offer the further advantage that the random sequence peptides are not free in the solution but fused to the phage capsid and associated to their coding sequence inserted in the phage genome [19]. Therefore, after phage selection and amplification, the amino acid sequence of the displayed peptides can be deduced from the corresponding nucleotide sequence.

Peptide sequences are then aligned and grouped into families of similar sequences thus allowing the identification of a consensus. Table 2 summarize the data obtained by the analysis of the substrate specificity of four closely related PTKs, members of the Src kinase family and a more distantly related kinase Syk. The specificity of the selection procedure is highlighted by the fact that distinct substrate features were identified, even though these kinases share a high homology in structure and function [3].

Position+1, with respect to the phosphorylated tyrosine, is preferentially occupied by a G or G/W for Fyn and c-Src, while an acidic residue is preferred by Blk, Lyn and Syk. Position +2 does not seem to affect substrate specificity, while hydrophobic side chains are strongly preferred at position +3 in all the TKs examined. Only Syk does not require specific residues in that position. On the other hand, Syk is the only TK that prefers a negatively charged residue in position -1, where the other TKs require a hydrophobic residue or also a T in the case of Fyn.

Experimentally determined substrate features were confirmed by comparison with the sequences of the physiological substrates of these kinases. For example:

16. Dente L, Cesarini G, Cortese R. pEMBL: a new family of single stranded plasmids. Nucleid Acid Res 1983;72:767-78.

17. Dente L, Cesareni G, Micheli G, Felici F, Folgori A, Luzzago A, Monaci P, Nicosia A, Delmastro P (1994) Monoclonal antibodies that recognize filamentous phage: tools for phage display technology. Gene 148, 7-13.

18. Abram CL, Courtneidge SA. (2000) Src family tyrosine kinases and growth factor signaling. Exp Cell Res. 254(1), 1-13.

19. Smith GP (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 228, 1315.

Table 2. Substrate specificity of PTKs: position 0 indicates the phosphorylated Y; other positions indicate the preferred amino acids surrounding the Y. "h" represents residues with hydrophobic side-chains; x indicates each of the 20 possible amino acids.

TK kinase	-3	-2	-1	0	+1	+2	+3	phage libraries	References
Syk	E/D/x	E/D/x	D	Y	E	x	x	9mer (pIII)	[5]
								Y-15 mer (pIII)	
SRC	E/D	x	L/I	Y	W/G	x	F/W	9mer (pIII)	[5]
Lyn	E/D		L/I	Y	D/E	x	L	9mer (pIII)	[5]
Blk			I	Y	D/E	x	L	9mer (pIII)	[5]
Fyn	E	x	T/hy	Y	G	x	hy	7mer (pVIII)	[16]
								9mer (pVIII)	

the two negatively charged aminoacids that surround the Y preferred by Syk, fit exactly in the phosphorylation site of the anion transporter band 3 (a Syk physiological substrate); the hydrophobic/Y/acidic residues, preferred by Blk and Lyn, are critical features of the ITAM motifs of the b-cell receptors; the T in position -1 at the phosphorylation site of Fyn specific substrates is found also in a p34cDC2 derived peptide, used to discriminate between Src-like PTK activities from other PTKs.

Finally, it has been observed that PTKs preferentially phosphorylate Y residues preceded by acidic residues [4]. Substrate motifs reported in Table 2 confirm this observation, even though the exact position of the negative charges can vary from -1 to -5 with respect to the phosphorylated Y [6]. We found, indeed, that the position of the phosphorylated Y is often near the amino-terminus of the peptide, where a negative charge is provided by the vector sequence preceding the cloning site (Table 1).

While the catalytic specificity of PTK can be evidenced only after a short reaction time, when the library phosphorylation is carried out under saturation conditions, the specificity of the TK enzymes is broaden, and virtually every phage displaying a peptide containing a Y can be phosphorylated, irrespective of the amino acid sequence context. These fully modified libraries have been exploited to select the specific phosphotyrosine peptides that bind SH2 or PTB domains. In Table 3 we have summarized the results obtained by screening phage libraries, phosphorylated with $p55^{fyn}$ PTK, with the SH2 domains of Grb2, of three members of the Shc family (3) and of the PTB domain of Shc-A (2) The binding specificity of Grb2 SH2 was also confirmed using a different phosphorylated library [20].

Table 3. Consensus sequences derived from selected peptides with pY-binding domains [16]. X indicate any amino acid; (E/M*) specificity revealed using a different library [20].

pY- binding Domains	-3	-2	-1	0	+1	+2	+3
GRB2-SH2	x	x	V	pY	(E/M*)Q	N	W/F
SHC-SH2	x	x /N	I/V	pY	E/G	x	L/I/V
SLI-SH2	x	x	L/I/V	pY	E/G	x	W/Y/F
RAI-SH2	x	x	I/V	pY	E/G	x	I/L
SHC-PTB	N	P	T	pY	x	x	W/Y

In conclusion, by using this approach it was possible to uncover the differences in the recognition specificity of the different p-Y domains and to prove that the bias imposed on the method by the kinase substrate specificity, if any, does not influence the consensus sequence of the binding peptide. Although this method does not directly identify substrate proteins, potential substrates can be predicted by searching protein sequence data banks with the consensus sequence. Indeed, the comparison with the sequences of the natural targets (when known) indicates a good agreement among the experimentally determined structural features and the residues surrounding the phosphorylated Y. In some cases we defined an extended consensus that includes more residues than those that are crucial for peptide recognition in natural targets. These residues occupy positions that are not conserved in the natural targets, but that could increase the strength of the binding and could therefore be a useful tool to identify potential inhibitors [21].

To test these hypothesis it is usually necessary to synthesize the corresponding peptides designed on the bases of the substrate consensus and to compare the binding affinity for the protein of interest: high affinity binders may be used to compete physiologically important protein-protein interactions.

20. Gram H, Schmitz R, Zubewr JF and Baumann F, (1997) Identification of phosphopeptide ligands for the Src-homology 2 (SH2) domain of Grb2 by phage display. Eur. J. Biochem. 246, 633-637.

21. Cody WL, Lin Z, Panek RL, Rose DW, Rubin JR (2000) Progress in the development of inhibitors of SH2 domains. Curr Pharm Des. 6(1):59-98.