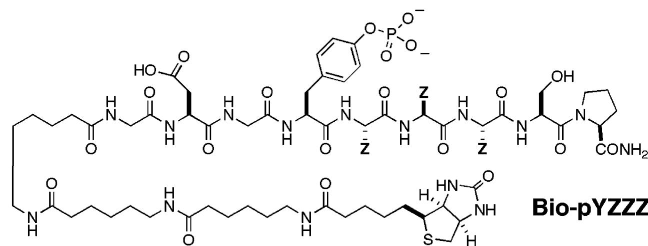
J|A|C|S
COMMUNICATIONS

Published on Web 03/09/2004

Two-Dimensional Diversity: Screening Human cDNA Phage Display Libraries with a Random Diversity Probe for the Display Cloning of Phosphotyrosine Binding Domains

Elizabeth J. Videlock, Victor K. Chung, Michael A. Mohan, Timothy M. Strok, and David J. Austin*
Department of Chemistry, Yale University, New Haven, Connecticut 06520
Received October 12, 2003; E-mail: david.austin@yale.edu

Interaction domains govern many of the protein/protein interactions that are crucial for normal cellular function.[1] Src homology 2 (SH2) domains represent a major class of interaction domains that are responsible for the recognition of phosphotyrosine (pTyr)-containing proteins, an interaction found almost exclusively in metazoan cellular signal transduction.[2] Complete knowledge of all possible interactions between SH2 domains and potential pTyr ligands is fundamentally important for the understanding of signal transduction itself, since any single molecular interaction can alter cellular processes. Toward this end, both synthetic random peptides[3] and phage-displayed cDNA libraries[4] have been used to identify ligands for known SH2 domains. cDNA libraries have also been used to screen for SH2 domains with known phosphotyrosine peptides.[5] Both of these approaches involve the screening of a library of molecules with a single known entity. We endeavored to broaden these efforts by utilizing a library versus library approach, which eliminates the need for a known entity, thereby expanding the realm of isolatable domains. In this communication, we report the use of a randomized pTyr peptide probe, in combination with human cDNA phage display libraries, for both the isolation and binding evaluation of multiple SH2 domains. In addition, we show that these phage-displayed domains are functional, exhibiting both phospho- and dose-dependent binding characteristics.



Design of the bio-pYZZZ diversity peptide was guided by a previous pTyr random peptide library[3b] and is based on the observation that domain recognition is C-terminal to the pTyr.[3a,6] The biotinylated nonapeptide contains three aminohexanoic acid spacer units, a Gly-Asp-Gly sequence preceding the pTyr, and three randomized residues following it. Both tryptophan and cysteine were eliminated from the synthesis to improve overall coupling efficiency and reduce difficulties with oxidation.[3b] A Ser-Pro sequence was added after the randomized region to aid in solubility, and the C-terminal end of the peptide was amidated to reduce overall charge. Amino acid analysis showed good representation ($\leq\pm1.2$ mol % difference between expected and observed values), suggesting degeneracy in the randomized region and providing an average molecular weight.[7] This probe is used as a single molecule in this study, yet it contains $18 \times 18 \times 18 = 5832$ different sequence combinations.

The affinity selection was performed with 11 different T7 cDNA phage display libraries: seven normal and four carcinoma tissue sources.[7] For each library, a single well of an avidin-coated 96-well plate was treated with bio-pYZZZ, followed by blocking with 1 mM biotin. An aliquot of each T7 cDNA phage library was precleared by incubation with an untreated avidin plate before being transferred to the bio-pYZZZ treated plate. After incubation overnight at 4 °C and removal of unbound phage, the remaining binders were amplified by re-infection with log phase *Escherichia coli*. The resulting lysate was used directly in the next round, and this cycle was repeated for a total of three selection rounds. Phage from the third round were eluted using 1% SDS (in TBS buffer), diluted with 2xYT and plated on LB top agarose plates to produce individual phage plaques. A total of 464 phage clones, approximately 36 from each library, were randomly chosen and individually screened for bio-pYZZZ probe-binding affinity on seven avidin-coated 96-well plates. The phages were ranked on the basis of relative titer, and the top 86 clones were subjected to DNA sequence analysis.[7]

Table 1. SH2 Domains Identified in the bio-pYZZZ Selection and Summary of Phage Clone Data

SH2 domain[a]	highest rank[b]	no. of clones (unique)[c]	clone a.a. range[d]	SH2 a.a. range[e]	pYZZZ EC_{50} (nM)	max titer[f] ($\times10^6$)
P85α1	1	30 (6)	300−430	331−414	263	3.7
P55γ1	2	5 (5)	51−204	64−145	453	2.9
P55γ2	3	1 (1)	348−461	356−438	893	1.1
FYN	9	3 (2)	121−269	146−236	466	4.4
VAV	10	2 (1)	634−780	671−752	488	3.4
SCK	33	2 (2)	297−425	328−407	724	2.8
P85α2	40	3 (1)	614−724	622−704	557	0.31
P85β	41	2 (1)	300−427	329−410	455	3.6
SH2β	57	1 (1)	510−671	525−610	778	0.97

[a] Domain names are based on the parent protein, and a numerical suffix is used if the protein contains more than one SH2 domain. [b] Domains are ranked in order of their relative titer from the initial validation screen.[7] [c] Total number of occurrences for each domain and number of unique clones. [d] Amino acid ranges for the isolated clones. [e] Accepted amino acid ranges for each domain, based on values reported in the SMART database.[8] [f] Maximum rescue titer, taken from the average maximum value observed during the concentration-dependent binding study, before normalization.

Of the top 50 ranking phages, 40 encoded SH2 domains. In total, 49 of the 86 phage clones contained SH2 domains, 2 clones expressed non-SH2 protein fragments, and 33 expressed short "nonsense" peptides, arising from gene sequences that are out-of-frame with the cp10 phage coat protein. The DNA sequence for two of the clones could not be determined. The predominance of SH2 domains in this selection is significant since cDNA libraries have the potential to contain *any* protein in the human proteome. All 49 of the SH2 domain-containing genes were in-frame with the cp10 coat protein, as expected for a properly expressed fusion protein, and contained the entire domain sequence (Table 1).[7] In summary, 9 distinct SH2 domains in 20 different clones were isolated from 7 different proteins: the adaptor proteins SCK[9] and SH2β,[10] the phosphatidyl inositol kinase regulatory subunits P85α,[11]

10.1021/ja039006p CCC: $27.50 © 2004 American Chemical Society

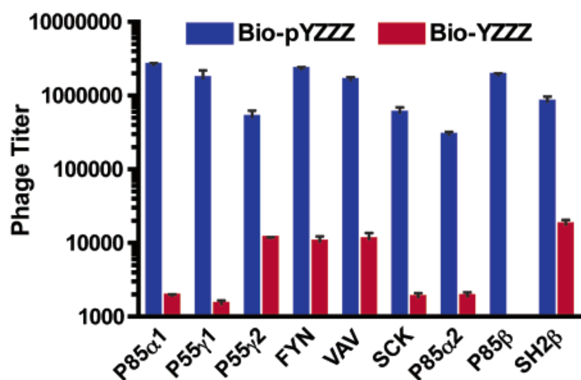


Figure 1. Specificity analysis of phage-displayed SH2 domains with phosphorylated bio-pYZZZ and unphosphorylated bio-YZZZ, plotted on a log scale.

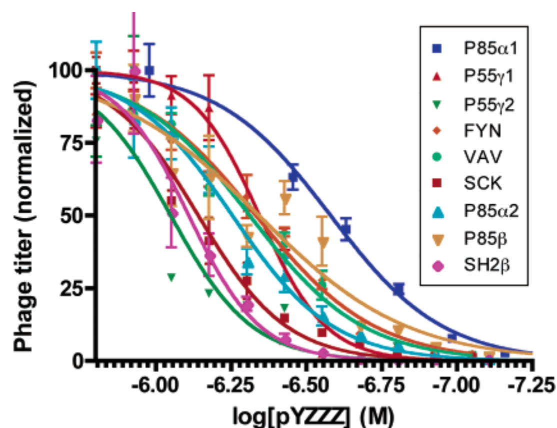


Figure 2. Probe-dependent binding analysis of the isolated SH2 domains with the phosphorylated bio-pYZZZ diversity probe.

P55γ,[12] and P85β,[13] a guanyl-nucleotide exchange factor VAV,[14] as well as two isoforms of the FYN tyrosine kinase.[15] One representative clone for each SH2 domain was selected for further binding analysis.

We investigated the phosphate dependence of phage binding to establish that the phage-displayed SH2 domains exhibit the same phospho-dependent binding characteristics as in their native state. For each SH2 domain, two wells of an avidin-coated 96-well plate were treated with probe: one with 10 μg/mL bio-pYZZZ and the other with 10 μg/mL bio-YZZZ, an unphosphorylated version of the random probe. Each set of wells was incubated overnight at 4 °C with lysate from one of the representative SH2 phage. After being washed, bound phage were eluted with 1% SDS, and the rescue titer was measured. In each case, the unphosphorylated probe bound at or below background levels (10^3−10^4 phage) and at least two log units below that observed for the phosphorylated probe (Figure 1).[7]

An on-phage binding analysis was performed to establish the relative probe affinity for the isolated SH2 domains. Consecutive wells of an avidin-coated 96-well plate were incubated with decreasing concentrations of bio-pYZZZ peptide. After affinity selection, the rescued phage were titered (Table 1), the normalized titers were plotted as a function of incubated probe concentration (Figure 2),[16] and the data were fitted by nonlinear regression analysis.[17] The calculated EC_{50} values are listed in Table 1 and ranged from 263 to 893 nM,[7] well within the expected range for pTyr/SH2 domain interactions.[3]

Since the affinity selection was performed using an unbiased natural protein source, our results further confirm the binding propensity of pTyr-containing peptides for SH2 domains, even in the absence of specific sequence information. It should be noted that since the bio-pYZZZ probe is a collection of randomized sequences, the observed affinity is a result of the binding affinities of *many* different peptides. Despite the degenerate nature of the probe, however, it displays good specificity for the SH2 domain interaction, as evidenced by the phosphate-dependent binding shown in Figure 1. The probe also appears relatively nonselective, binding to SH2 domains known to have distinct ligand selectivity,[3] albeit with different affinities (Figure 2).

In conclusion, we have shown that a library of random phosphotyrosine peptides can be used as a *single* diversity probe, in combination with cDNA phage display affinity chromatography for the isolation of SH2 domains. We have also developed a rapid on-phage binding assay that is compatible with any biotinylated probe. In this manner, once phage-displayed proteins are isolated, they can be directly assayed in a 96-well plate format with any number of probes. We are currently expanding this effort to construct an entire human SH2 domain array on phage to facilitate the proteome-wide analysis of phosphotyrosine recognition and to determine ligand selectivity for each of the isolated domains.

Acknowledgment. We thank Ruben M. Savizky and Michael B. Lazarus (Yale University), Barbara Morris (EMD Biosciences, Inc.), and Tamara L. Smith (The Rothberg Institute) for helpful discussion, EMD Biosciences, Inc. for donation of cDNA libraries, Janet Crawford and Myron Crawford (W. M. Keck Center) for peptide synthesis and amino acid analysis, and the National Institutes of Health (R01 GM59673) for financial support. D.J.A. is the recipient of an Alfred P. Sloan Foundation Fellowship.

Supporting Information Available: Experimental details (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Pawson, T.; Scott, J. D. *Science* **1997**, *278*, 2075−2080.
(2) Yaffe, M. B. *Nat. Rev. Mol. Cell Biol.* **2002**, *3,* 177−186.
(3) (a) Songyang, Z.; Shoelson, S. E.; Chaudhuri, M.; Gish, G.; Pawson, T.; Haser, W. G.; King, F.; Roberts, T.; Ratnofsky, S.; Lechleider, R. J.; Neel, B. G.; Birge, R. B.; Fajardo, J. E.; Chou, M. M.; Hanafusa, H.; Schaffhausen, B.; Cantley, L. C. *Cell* **1993**, *72*, 767−778. (b) Songyang, Z.; Shoelson, S. E.; McGlade, J.; Oliver, P.; Pawson, T.; Bustelo, X. R.; Barbacid, M.; Sabe, H.; Hanafusa, H.; Yi, T.; Ren, R.; Baltimore, D.; Ratnofsky, S.; Feldman, R. A.; Cantley, L. C. *Mol. Cell Biol.* **1994**, *14*, 2777−2785.
(4) Cochrane, D.; Webster, C.; Masih, G.; McCafferty, J. *J. Mol. Biol.* **2000**, *297*, 89−97.
(5) Zozulya, S.; Lioubin, M.; Hill, R. J.; Abram, C.; Gishizky, M. L. *Nat. Biotechnol.* **1999**, *17*, 1193−1198.
(6) Waksman, G.; Shoelson, S. E.; Pant, N.; Cowburn, D.; Kuriyan, J. *Cell* **1993**, *72*, 779−790.
(7) See Supporting Information for detailed experimental protocols, as well as DNA sequence information, amino acid analysis, and binding data.
(8) (a) Schultz, J.; Milpetz, F.; Bork, P.; Ponting, C. P. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5857−5864. (b) SMART Home Page. http://smart.embl-heidelberg.de/.
(9) SCK, Genbank RefSeq XM_042009; *Homo sapiens* neuronal Shc adaptor homolog.
(10) SH2β, Genbank RefSeq NM_015503; *Homo sapiens* SH2-B homolog.
(11) PI3K p85α subunit, Genbank RefSeq XM_043865; *Homo sapiens* phosphoinositide-3-kinase, regulatory subunit, polypeptide 1.
(12) PI3K p55γ subunit, Genbank RefSeq NM_003629; *Homo sapiens* phosphoinositide-3-kinase, regulatory subunit, polypeptide 3.
(13) PI3K p85β subunit, Genbank RefSeq NM_005027; *Homo sapiens* phosphoinositide-3-kinase, regulatory subunit, polypeptide 2.
(14) VAV, Genbank RefSeq NM_003371; *Homo sapiens* VAV 2 oncogene.
(15) FYN, Genbank RefSeq NM_002037; *Homo sapiens* FYN oncogene, transcript variant 1 and NM_153047, transcript variant 2.
(16) The concentration of the random bio-pYZZZ probe reported in the binding study was determined using the molecular weight calculated from an amino acid analysis (see Supporting Information).
(17) Nonlinear regression analysis, using a four parameter logistic equation, was performed using GraphPad Prism version 4.0a for MacOS X, GraphPad Software: San Diego, CA.

JA039006P