Characterization of SH2−Ligand Interactions via LibraryAffinity Selection with Mass Spectrometric Detection

Michele A. Kelly, Hongbin Liang, Iou-Iou Sytwu, Isodoros Vlattas, Nancy L. Lyons, Benjamin R. Bowen, and Lawrence P. Wennogle*

CIBA Pharmaceuticals Division, 556 Morris Avenue, Summit, New Jersey 07901

Received March 7, 1996; Revised Manuscript Received May 10, 1996[⊗]

ABSTRACT: Synthetic combinatorial libraries have proven to be a valuable source of diverse structures useful for large-scale biochemical screening. Their use has greatly facilitated the study of protein−protein interactions. We have developed a practical technique for screening such libraries by integrating affinity chromatography selection with electrospray ionization mass spectrometric detection, referred to as library affinity selection−mass spectrometry (LAS-MS). The process allows for rapid and efficient screening of solution phase libraries and provides detailed information such as the relative affinities of substrates. The method is generally applicable to include nonpeptide libraries; moreover, electrospray tandem mass spectrometry (ES-MS/MS) yields sequence-specific identification of individual components without the need for chemical tags. This technique is demonstrated using the Src homology 2 (SH2) domain of phosphatidylinositol 3-kinase (PI 3-kinase). The critical importance of methionine in the position +3 (relative to the phosphotyrosine position) is demonstrated in a library built with a phosphotyrosine mimic, (phosphonodifluoromethyl)phenylalanine. The described method has broad applicability to combinatorial library screening.

Combinatorial chemistry is a valuable approach for creating chemical diversity (Felder, 1994; Gallop et al., 1994; Gordon et al., 1994; Houghten, 1993; Houghten et al., 1991; Lam et al., 1991; Lebl et al., 1995) and screening of combinatorial libraries has great potential for rapidly identifying enzyme inhibitors and elucidating protein−protein interactions (Kaur et al., 1995; Songyang et al., 1993; Till et al., 1994). Various techniques for the evaluation of combinatorial libraries have been developed, including the iterative approach (Dooley et al., 1993; Dooley & Houghten, 1993; Furka et al., 1988; Houghten et al., 1991; Pinilla et al., 1993), the spatially addressed approach (Fodor et al., 1991; Gallop et al., 1994), the one-bead, one-peptide method (Chen et al., 1993; Lam et al., 1991; Lebl et al., 1995; Salmon et al., 1993), and the termination synthesis method (Youngquist et al., 1994, 1995). Encoded libraries (Brenner & Lerner, 1992; Kerr et al., 1993; Needels et al., 1993; Ohlmeyer et al., 1993) and biological libraries (Cull et al., 1992; Cwirla et al., 1990; Goodson et al., 1994; Smith, 1985) have further influenced the development of this field. These methods all have certain attributes as well as certain drawbacks. Several of the methods listed above attempt to accommodate nonpeptide libraries in inventive ways; however, many systems deal with resin-bound substrates that may hinder substrate binding. Likewise, tagging of libraries may be burdensome.

SH2 domains are an important protein motif (Moran et al., 1990; Pawson, 1995; Schlessinger, 1994; Stahl et al., 1988), which bind phosphotyrosine-containing sequences and are involved in protein−protein interactions. Of 50 or more SH2 domains identified, short peptides have been shown to suffice for binding, with the amino acids immediately surrounding the phosphotyrosine serving to define specificity for each unique system (Songyang et al., 1993, 1994). In the case of PI 3-Kinase (Herbst et al., 1994; Klippel et al., 1992; Panayotou et al., 1993), the SH2 domain requires a methionine as the third residue on the C-terminal side of the phosphotyrosine (referred to as the +3 or X_3 residue), as originally established using peptides derived from the natural growth factor receptors. Several phosphotyrosine mimics have been identified that are able to maintain high binding affinity to SH2 domains (Burke et al., 1994; Ye et al., 1995). These mimics offer the advantage of a stabilized, phosphatase-resistant substitution and potential cellular permeability.

We have developed a technique called library affinity selection−mass spectrometry (LAS-MS) to exploit the attributes of solution phase libraries, affinity selection, and mass spectrometry to study the Src homology 2 (SH2) domain of phosphatidylinositol 3-kinase (PI 3-kinase). This approach extends an approach taken by Songyang et al. (1993), which utilized affinity selection of a peptide library followed by Edman peptide sequencing. Coupling the affinity selection process with electrospray ionization mass spectrometry (ES-MS) and electrospray tandem mass spectrometry (ES-MS/MS) in the present study led to an effective means to establish the molecular identity of active components and to rank order them. The system avoided the need for library encoding and for covalent immobilization of compounds on beads. The practicality of this method is demonstrated without the need for a sophisticated chromatographic set-up or a dedicated/modified mass spectrometric system. In addition, mass spectrometry is readily applicable to peptidomimics and nonpeptidic small molecule libraries. Recently, mass spectrometry has been used to characterize the integrity of combinatorial libraries (Brummel et al., 1994, 1996; Dunayevskiy et al., 1995; Kaur et al., 1995; Metzger et al., 1994).

* Author to whom correspondence should be addressed.

[⊗] Abstract published in *Advance ACS Abstracts,* August 15, 1996.

S0006-2960(96)00571-5 CCC: $12.00

MATERIALS AND METHODS

Chemicals. Glutathione-agarose was purchased from Pharmacia (Piscataway, NJ). Other chemicals were purchased from Sigma (St. Louis, MO). The standard peptides, (1) GY(Nle)P(Nle)S-amide, (2) G(pY)(Nle)P(Nle)S-amide, (3) G(F_2Pmp)(Nle)P(Nle)S-amide, and (4) GNGDS(pY)-MPMPKS-amide [Nle = norleucine; pY = phosphotyrosine; and F_2Pmp = (phosphonodifluoromethyl) phenylalanine], were synthesized by the FMOC solid phase method. The combinatorial library was purchased from Multiple Peptide Systems (San Diego, CA) and was defined as G(F_2Pmp)X_1-PX_3S-amide, where F_2Pmp is (phosphonodifluoromethyl)-phenylalanine and X = 19 of the 20 standard L-amino acids (minus cysteine).

GST-SH2 Purification. A plasmid directing the expression of a fusion protein including the C-terminal SH2 domain (residues 617−722) of human PI 3-Kinase (Panayotou et al., 1993) was constructed and expressed in an *Escherichia coli* expression system. Liter-scale preparations were grown to an OD of 0.6−0.8 at 30 °C and then induced with 1 mM IPTG until harvest 3 h later. In a typical preparation, 2 g of cell pellet were suspended in 12 mL of 1% Triton in phosphate buffered saline (PBS), sonicated 3 × 20 s and centrifuged at 8000 rpm for 20 min using a SS-34 rotor. The supernatant was mixed with 10 mL of glutathione-agarose to enable the GST-SH2 to bind to the matrix. The mixture was gently shaken at 4 °C for 30 min, loaded onto an empty column, and washed with 200 mL of PBS. GST-SH2 was eluted from the column using 50 mM glutathione in 100 mM NH_4HCO_3, pH 8.6. The fidelity of the SH2 domain was verified by ES-MS analysis (MW_{calc} = 13 126.5 Da, MW_{found} = 13 126.4 Da), protein sequence analysis, and SDS−PAGE after cleavage from the GST fusion partner using factor X and standard techniques.

SH2 Affinity Chromatography. The GST-SH2 fusion protein (12 mg) was coupled to 1 mL of glutathione-agarose. The gel was pre-equilibrated with 50 mM ammonium acetate buffer, pH 7.5 (equilibrating buffer). An SH2 binding assay was performed using a Bolton−Hunter labeled phosphopeptide essentially as described by Piccione et al. (1993). Data analysis was performed to generate IC_{50} values and nonlinear regression curves using PRISM (GRAPHPAD Software Inc., San Diego, CA).

For the batch method of library analysis, the library (2.5 mg) was dissolved in 1 mL of equilibrating buffer and mixed with 1 mL of beads at 4 °C for 1 h. The SH2-GST glutathione-agarose matrix was washed three times with 10 bed volumes of equilibration buffer, and the weakly bound peptides were removed by three successive washes with one bed volume of 1 M ammonium acetate buffer, pH 7.5. More strongly bound library components were eluted successively with 100 mM triethylamine acetate (TA) buffer, pH 4.5; 100 mM TA buffer, pH 3.5; and 0.5% trifluoroacetic acid (TFA). Each step entailed three consecutive washes with two bed volumes of elution buffer, and the eluted peptides were monitored by either ES-MS or ES-MS/MS (see below).

In other applications, a column-based gradient elution was used employing one of three linear gradients over 80 min at a flow rate of 0.5 mL/min. The pH gradient was from pH 7.5 (50 mM ammonium acetate) to 3.5 (100 mM triethylamine acetate), the chaotrope (KSCN) gradient ranged from 0 to 1, M and the displacer (sodium phenyl phosphate) gradient from 0 to 0.25 M. The eluted peptides were monitored by reverse-phase high-performance liquid chromatography (HPLC). A Bio-Rad (Richmond, CA) 250 × 4.6 mm RP 318 reverse-phase HPLC column was utilized with a 30 min linear gradient from 10% B buffer [90% acetonitrile (AcN), 0.09% TFA]/90% A buffer (5% AcN, 0.09% TFA) to 35% B/65% A. HPLC was performed at room temperature at 1 mL/min and was monitored for absorption at 210 nm with a Waters 490 UV detector. The reverse-phase HPLC retention times of the standards 1−4 (see above) were found to be 19.7, 14.5, 18.0, and 10.3 min, respectively.

Mass Spectrometry. Electrospray mass spectrometry (ES-MS) is a soft ionization technique that predominantly produces molecular ions under normal operating conditions. The use of electrospray tandem mass spectrometry (ES-MS/MS) enables collision-induced dissociation of a selected molecular ion, resulting in a fragmentation pattern that contains structural information regarding the selected precursor ion. ES-MS and ES-MS/MS were performed on a PE Sciex API III triple quadrupole mass spectrometer (Concord, ON). A Shimatzu (Kyoto, Japan) LC-10 HPLC was used to deliver a flow rate of 30 μL/min directly to the electrospray interface. Samples were concentrated using a Michrom BioResources (Auburn, CA) small molecule trap placed in-line with the HPLC system and washed with 2:98 A/B (A, AcN; B, 0.1% TFA) to remove interfering salts. Trapping of peptides was maintained using a 2:98 A/B ratio in the LC system, and the elution of peptides was facilitated by increasing the A/B ratio to 80:20 for direct infusion of peptides into the mass spectrometer.

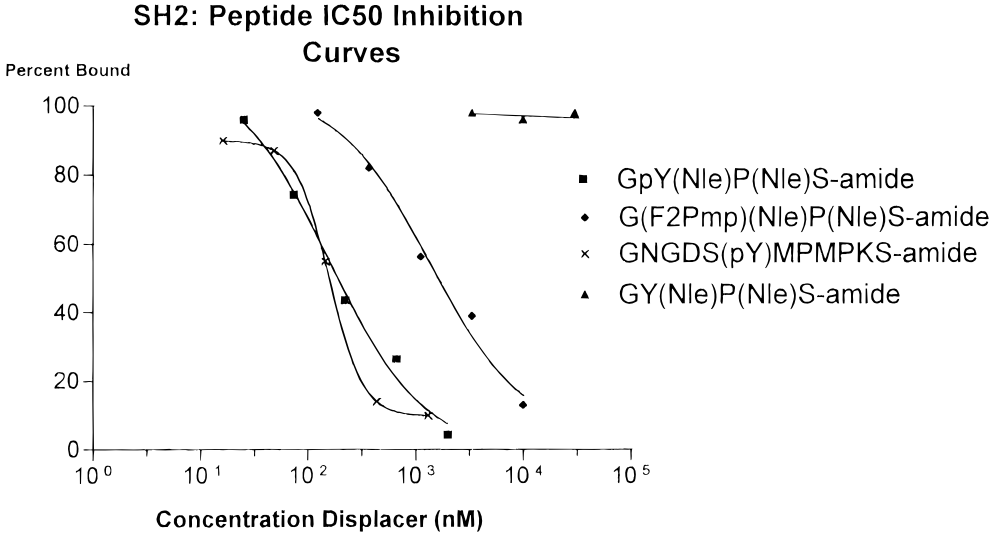
All ES-MS and ES-MS/MS analyses were carried out in the positive ion mode. Typical ES-MS conditions utilized an ionization voltage of 5 kV and orifice voltage of 70 V. Purified air was used as a nebulization gas (40 psi) and ultrapure nitrogen was used as a curtain gas at a flow rate of 1 L/min. ES-MS/MS was carried out using ultrapure argon at 260 × 10^{12} molecules/cm^2 collision gas thickness and a 29 eV collision energy.

The intact peptide library was dissolved in a 2:98 ratio (A/B) and analyzed from a 0.5 mg portion of the original sample. All ES-MS spectra were obtained using multichannel analyzer (MCA) mode and scanning the mass range 630−950 *m/z* in 0.1 *m/z* increments and with a 1 ms dwell time. The calculated amount of each peptide utilized to obtain the ES-MS and ES-MS/MS spectra from the intact library was 450 ng.

ES-MS/MS experiments were carried out using a scan range of 50 − (parent ion + ca. 10) *m/z* units in 0.1 *m/z* increments and a 1 ms dwell time. Limitations to MS/MS on a triple quadrupole instrument are due to practical sequence interpretation limits for compounds ca. >2000 Da and to the instrument's inability to distinguish between amino acids isoleucine and leucine (I/L; MW = 113.1 Da) or lysine and glutamine (K/Q; MW = 128.1 Da). In future work, the amino acid ambiguity can be overcome by eliminating one of the redundant residues from the library or by substituting an isotopic analog for one of the two.

RESULTS

Pilot Experiments. Purified human PI 3-Kinase SH2 domain was produced as a fusion protein with glutathione-S-transferase (GST). The SH2 fusion protein was im-



	IC50 (nM) Ave.	SD
GNGDS(pY)MPMPKS-amide	186	53
GY(Nle)P(Nle)S-amide	> 20000	
GpY(Nle)P(Nle)S-amide	202	24
G(F2Pmp)(Nle)P(Nle)S-amide	1193	263

FIGURE 1: Displacement of the binding of Bolton−Hunter radiolabeled SH2 binding peptide by various displacers. Peptides were tested as described in Materials and Methods for the ability to inhibit binding of a radiolabeled peptide to the C-terminal SH2 domain of PI 3-kinase.

mobilized to glutathione-agarose and used to screen a biased combinatorial library. Pilot experiments with model compounds were performed to test this affinity isolation technique. The four standard compounds, designed to be variations of the PDGFR PI3-kinase-SH2 binding sequence, were as follows: GY(Nle)P(Nle)S-amide; G(pY)(Nle)P(Nle)S-amide; G(F_2Pmp)(Nle)P(Nle)S-amide; and GNGDS(pY)-MPMPKS-amide [Nle = norleucine; pY = phosphotyrosine; and F_2Pmp = (phosphonodifluoromethyl)phenylalanine]. Three gradient elution methods [phenylphosphate displacer, chaotrope agent (KSCN), and pH elution] were evaluated for separating these compounds on the SH2 affinity column. Binding assays of these four standard compounds revealed the IC_{50} values shown in Figure 1.

All three gradient procedures released the standards from the SH2 affinity system and afforded unique separation profiles (Figure 2). The order of elution differed among the various methods, most likely reflecting the differing mechanisms by which the elution agents were disrupting the binding interactions. For example, GNGDS(pY)MPMPKS-amide eluted late in the presence of the displacer gradient, but relatively early when a chaotrope gradient (a technique designed to disrupt hydrophobic interactions) was used. This compound has the highest affinity for the SH2 domain (Figure 1) and is a relatively hydrophilic compound, judging by reverse-phase HPLC (see Materials and Methods), due to its highly charged sequence. Of the three gradient procedures, we focussed on the pH gradient elution for subsequent coupling of this assay with mass spectrometry.

A software program was developed in order to predict the molecular weight distribution of the components of our library. Redundancies of two or more components exist because the same group of building blocks were used for positions X_1 and X_3. For example, each pair of variable residues such as SD and DS will have the same molecular weight. Although ES-MS analysis of the library will not distinguish between these redundancies, ES-MS/MS analysis can, due to its sequencing ability.

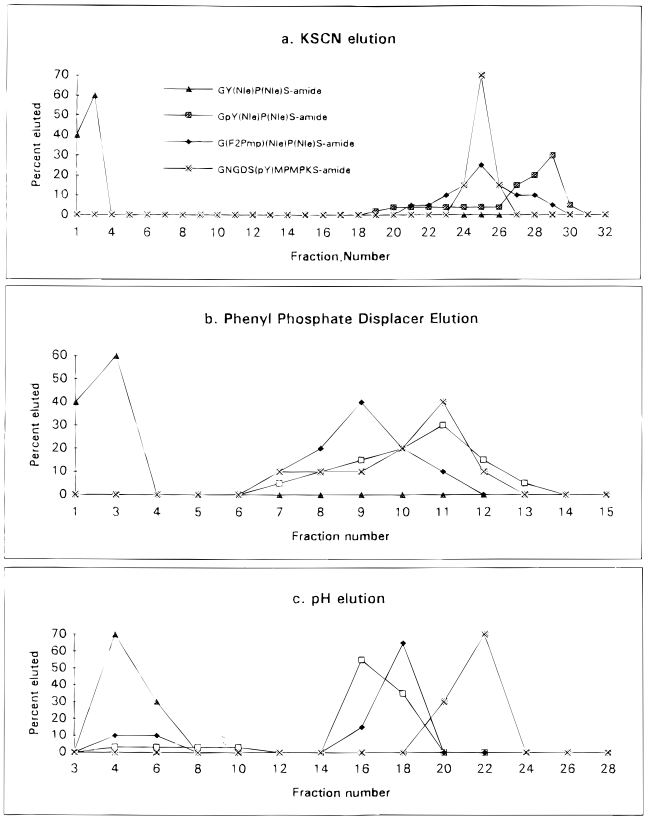


FIGURE 2: Comparison of various elution strategies for the separation of standard peptides on SH2-affinity columns. Standard peptides were run on the SH2-affinity column and eluted by various gradients as described in Materials and Methods. Fractions of the gradient elution were run on reverse-phase HPLC, quantitated using UV absorption at 210 nm, and plotted as the percent of the total peptide eluted. Parallel runs were performed with (a) KSCN, (b) phenyl phosphate, and (c) pH elution.

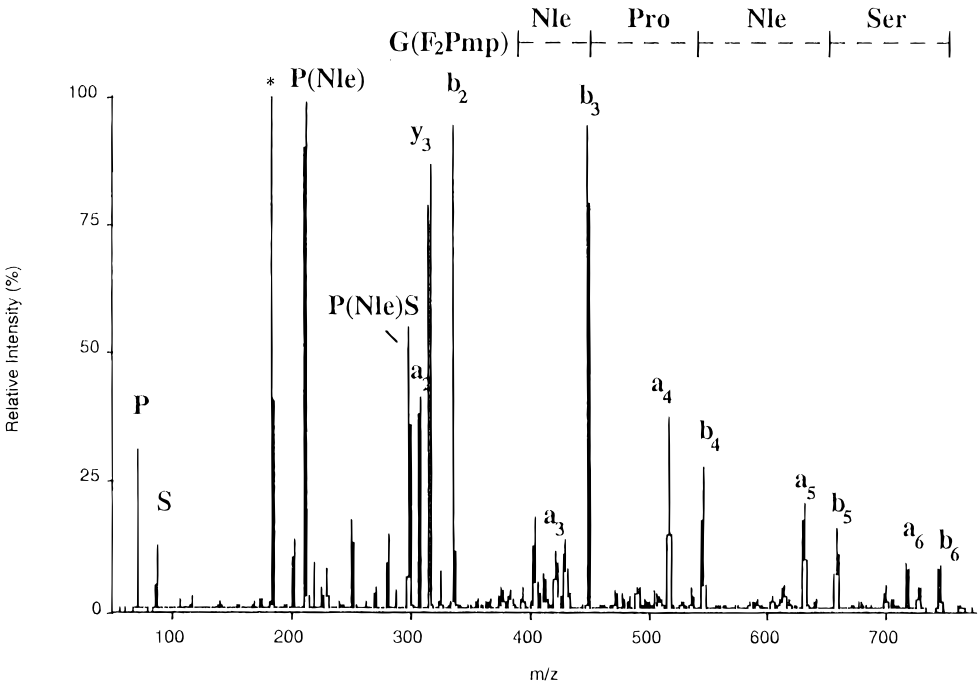


FIGURE 3: ES-MS/MS spectrum of G(F_2Pmp)(Nle)P(Nle)S-amide. The fragmentation pattern assignments are labeled according to the following: single letter designations refer to amino acids and internal fragments forming immonium ions; b-ions refer to ions formed with the charge retained on the N-terminus upon cleavage at the peptide bond; a-ions relate to N-terminal fragment ions with cleavage occurring after the α-carbon (28 Da less than corresponding b-ions); y-ions refer to ions formed with the charge retained on the C-terminus upon cleavage at the peptide bond (complimentary ions to corresponding b-ions). The sequence is designated above the b-ions in the mass spectrum.

The ES-MS system was tested with a standard compound, G(F_2Pmp)(Nle)P(Nle)S-amide, to ensure the effectiveness of the peptide trap for the modified peptides present in the library. This compound was retained by the trap and resulted in the expected molecular ion $\{(M + H)^+\}$ of *m/z* 762.5 (MW_{calc} = 761.4). ES-MS/MS analysis was performed to further explore the capabilities of the system for the present application and to probe for molecular identity and sequence information. The ES-MS/MS spectrum of the ion at *m/z* 762.5 displayed an informative fragmentation pattern (Figure 3). The "b" ions (N-terminal ions formed upon cleavage of the peptide bond) formed a clear pattern starting with the G(F_2Pmp) ion at *m/z* 334.4. From this ion the remaining amino acids could be detected in the correct order: (Nle)P-(Nle)S. Each b-ion had a paired "a" ion at −28 Da (−CO). In addition, proline residues are known to be very labile in MS/MS, where the cleavage takes place at the N-terminal side (so-called "internal ions"). This was the case for the above peptide, where internal fragment ions for P(Nle) and P(Nle)S were detected. These fragmentation patterns were valuable for the characterization of peptides in the library so that the order of the X_1 and X_3 amino acids could be determined via the b-ions $\{X_1$ after G(F_2Pmp)$\}$ and the internal ions $\{X_3$ after P$\}$.

Characterization of a Combinatorial Library by MS. The combinatorial library was defined as: G(F_2Pmp)X_1PX_3S-amide, where X = 19 of the 20 standard L-amino acids (minus cysteine) and, therefore, consisted of 361 (19^2) compounds. F_2Pmp [(phosphonodifluoromethyl)phenylalanine] was utilized as a high affinity phosphotyrosine mimic, as discussed by Burke *et al*. (1994). ES-MS analysis of the intact library revealed the diversity of the 361 components of the library, as shown in Figure 4. The distribution of the peptides properly reflected the expected statistical distribution of molecular weight combinations of amino acids. This mass spectrum also resembled an ES-MS spectrum reported for a 576 component peptoid mixture (Kaur et al., 1995).

An ES-MS/MS spectrum was obtained of the ion at *m/z* 796.6 from the total library mixture. Knowing the four fixed residues in the library, this molecular ion should represent eight distinct peptides containing X_1 and X_3 combinations: EM, PY, IF/LF, ME, YP, FI/FL. Inspection of the MS/MS spectrum (not shown), although complicated by the mixture, indicated the presence of all of the above combinations.

Library Affinity Selection. Affinity selection of the library, followed by ES-MS and ES-MS/MS analysis, was performed using a batch procedure. Five pH elution steps were evaluated: unbound; 1 M salt/pH 7.5 wash; pH 4.5 wash; pH 3.5 and TFA elutions, respectively. Most of the library components were weakly bound and eluted in the of 1 M salt/pH 7.5 and pH 4.5 washes (not shown). The pH 3.5 elution (see below), however, contained a subset of library components, which had stronger affinity for the SH2 protein. The TFA elution (not shown) contained trace amounts of two peptides that were not completely displaced at pH 3.5.

As seen in Figure 5, a small subset of components were detected in the pH 3.5 elution. Virtually all of these compounds could be assigned as methionine-containing peptides. Table 1 lists the molecular ions found over an intensity threshold value of 15 percent in the pH 3.5 elution, as well as the theoretical amino acid combinations associated with each molecular weight. Of four possible unambiguous methionine-containing combinations (TM, NM, HM, and WM), three such ions (TM, NM, and WM) were detected in the pH 3.5 elution. Additionally, four predicted methionine-containing peptides were not detected in the pH 3.5 elution (GM, PM, HM, and RM). Of these, the GM peptide

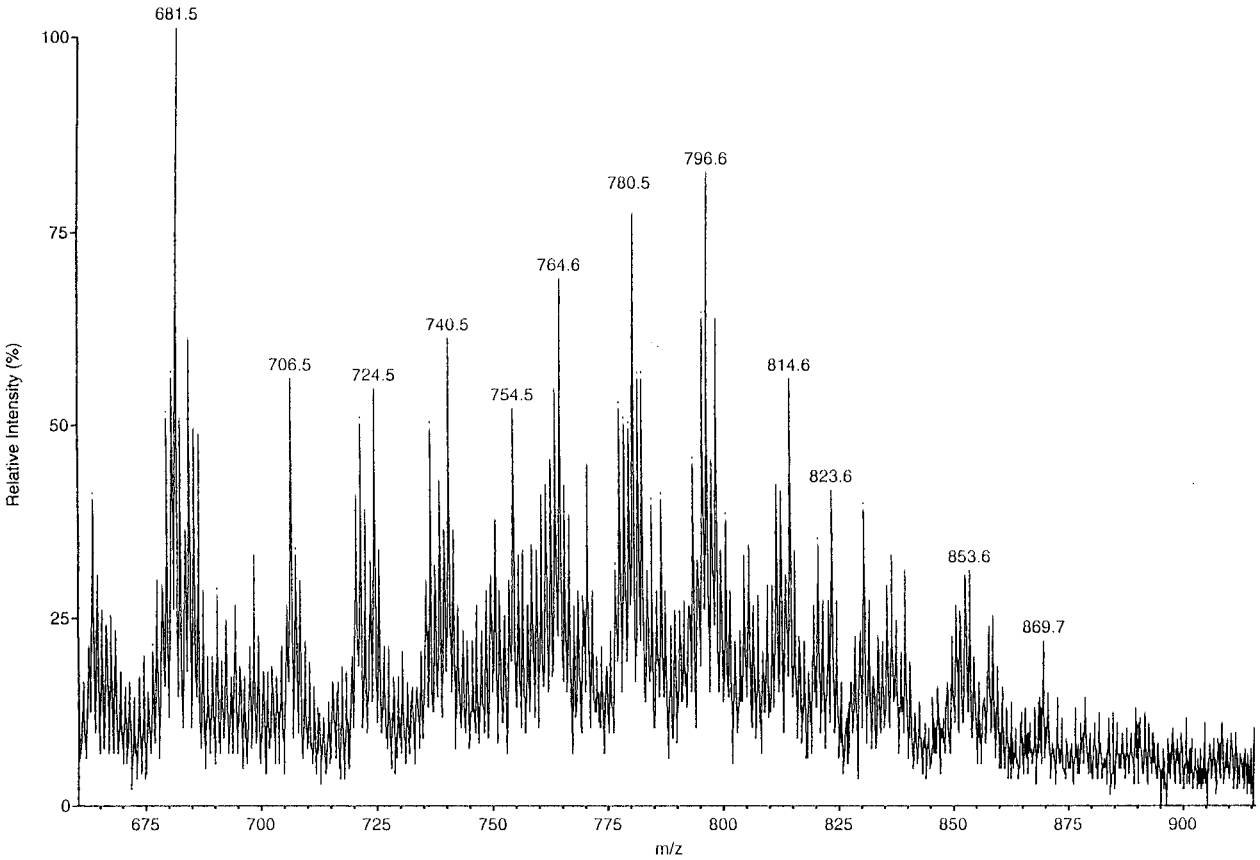


FIGURE 4: ES-MS spectrum of the intact library defined as $G(F_2Pmp)X_1PX_3S$-amide, where F_2Pmp is (phosphonodifluoromethyl)phenylalanine and X = 19 of the 20 standard L-amino acids (minus cysteine). This library contains 361 components, and the mass spectrum was obtained from 450 ng per peptide.

was found in the pH 7.5 and 4.5 elutions, the PM and HM peptides were found in the pH 4.5 elution, and the RM peptide was seen but was below the detection threshold level (15%) in the pH 3.5 elution. Three peptides were detected (*m/z* 735.7, 779.7, and 821.0) that did not match a methionine-containing molecular weight and might be of interest in future work regarding SH2 binding properties. Although the recurrence of methionine in the strongly bound peptides is readily discernible from the data in Table 1, the identity and order of residues in positions X_1 and X_3 must still be probed by ES-MS/MS.

The most abundant peptide in the pH 3.5 elution was *m/z* 798.5, which showed a significant increase from previous elutions and was also found in the TFA elution. This MW matches that for MM, FD, and VY combinations. ES-MS/MS analysis of this ion revealed that only the MM combination was present in the pH 3.5 elution (Figure 6). Given the fact that this ion is also found in the TFA elution, the MM combination can be qualitatively be assigned as a strong binder to SH2.

To probe other redundancies listed in Table 1, ES-MS/MS spectra were obtained for the same ion, *m/z* 780.5, over the pH elutions of 7.5, 4.5, and 3.5. Possible sequences of X_1 and X_3 for this ion are IM/LM, ED, PF, MI/ML, DE, and FP. Changes in the fragmentation patterns can be detected with decreasing pH. Most importantly, the position of the methionine can be seen to favor position X_3 at pH 3.5 for the (I/L)M combination. In addition, the peptide at *m/z* 780.5 was detected in the TFA elution. This implies that the (I/L)M combination and the MM combinations are the strongest binders to SH2, consistent with previously reported data (Songyang et al., 1993). Other ES-MS/MS analyses obtained [e.g., for ions at *m/z* 796.3 and 814.3 (see Table 1)] detected the EM and FM sequences, revealing that methionine predominates at position X_3 in the pH 3.5 elution.

DISCUSSION

We have demonstrated that screening combinatorial libraries in conjunction with ES-MS and ES-MS/MS detection is a feasible approach, yielding a great deal of information concerning peptide disposition and sequence information in a rapid time frame. SH2 domains have been extensively documented by various biophysical approaches, and the forces responsible for high affinity phosphopeptide interaction are well established. The phosphotyrosine group represents the greatest energy of interaction, while residues (chiefly) immediately C-terminal contribute to binding energy and specificity. In the case of PI 3-kinase, the pioneering work by Cantley (Songyang et al., 1993) demonstrated the importance of a hydrophobic residue in the +1 site and a methionine in the +3 site. We confirm this concept using the LAS-MS approach and extend these observations to show (1) that different combinations of such groups have different affinities and can be separated by judicious use of elution conditions and (2) that the same principle applies even with libraries based upon unnatural phosphotyrosine mimics.

Various methods of elution have been exploited for affinity chromatography of proteins and each method has unique advantages and disadvantages. Three possible elution techniques were compared here using model peptides. Each

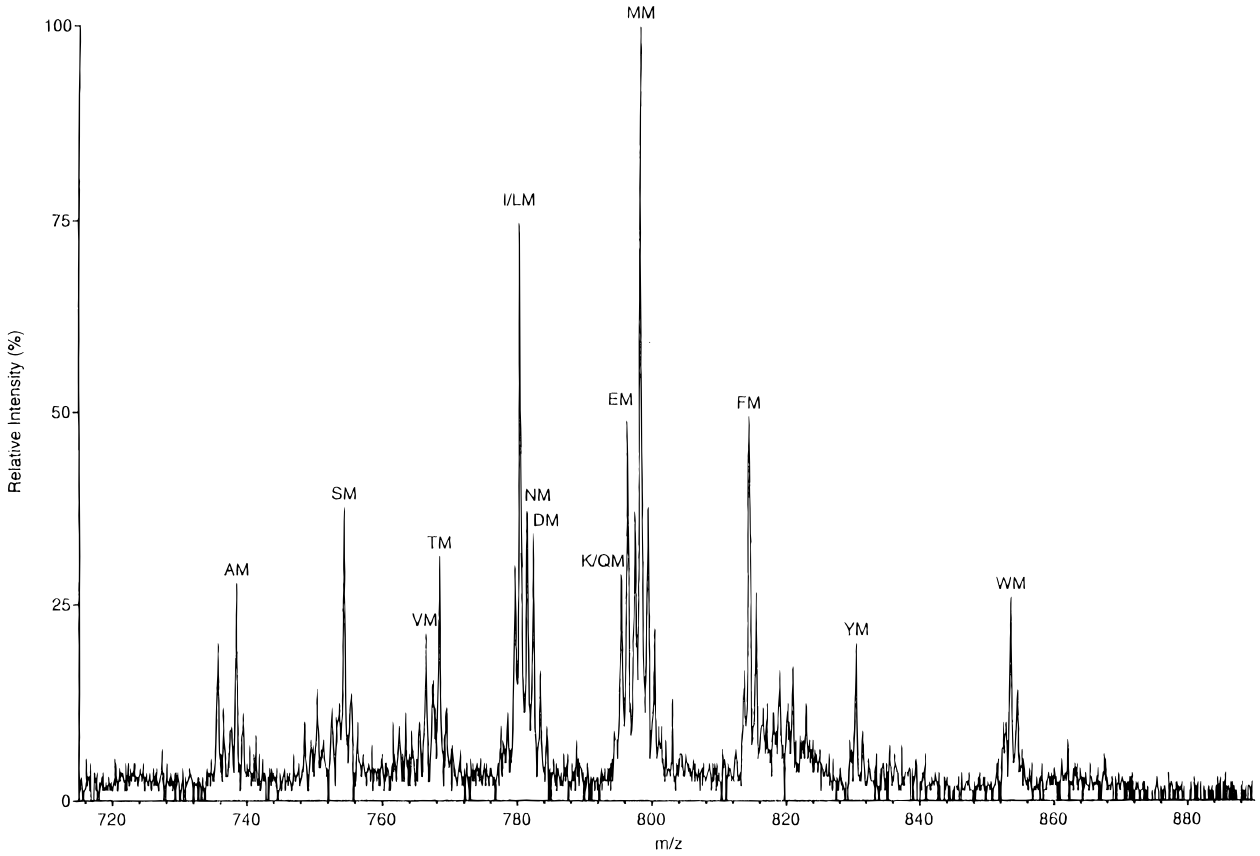


FIGURE 5: ES-MS spectrum of the library affinity selected components in a pH 3.5 affinity elution. Assignments are as designated as follows: sequences inferred for AM, SM, VM, DM, K/QM, and YM; molecular weights validated for TM, NM, and WM; and sequences confirmed for I/LM, EM, MM, and FM (see Table 1).

technique showed an individual character reflecting the physical parameter utilized by each method. Displacer elution using phenylphosphate, as originally used by Cantley (Songyang et al., 1993), is a mild method which released components in a manner reflective of the relative affinities at the phosphate binding pocket. Acid elution was likewise effective, and its elution profile relates to the titration of acidic amino acids in the SH2 domain as well as the titration

Table 1: Library Affinity Selected Compounds Detected by ES-MS (pH 3.5 Elution)

found at pH 3.5[a]					
$(M + H)^+$	intensity (×1000)	Met combinations[b,c]	other combinations[c]	MW validated (ES-MS)[d]	sequence validated (ES-MS/MS)[e]
735.7	340		AK/Q		
738.4	470	AM	TT		
754.5	640	SM	AF		
766.5	360	VM	TE,DD		
767.6	260				
768.5	530	TM		yes	
779.7	510		NE, K/QD, GW,SR		
780.5	1270	I/LM	ED, PF		yes
781.5	630	NM		yes	
782.4	580	DM	VF		
783.5	280				
795.5	490	K/QM			
796.5	830	EM	PY, I/LF		yes
797.6	630				
798.5	1700	MM	FD, VY		yes
799.5	640				
800.4	370				
814.5	840	FM	DY		yes
815.5	450				
821.0	290		ER, VW		
830.5	340	YM	FF		
853.5	440	WM		yes	

[a] Tabulations based upon molecular ions $\{(M + H)^+\}$ and their intensities detected by ES-MS above a threshold of 15% in the pH 3.5 affinity elution (see Figure 5). [b] Asterisked ions assigned as isotope peaks from the corresponding monoisotopic ions. [c] All combinations contain redundancies related to sequence order (i.e., the order of positions X_1 and X_3). [d] MW validated ions derived from pairs with nonredundant molecular weights. [e] Sequence validated ions confirmed using ES-MS/MS to determine amino acid position and order.

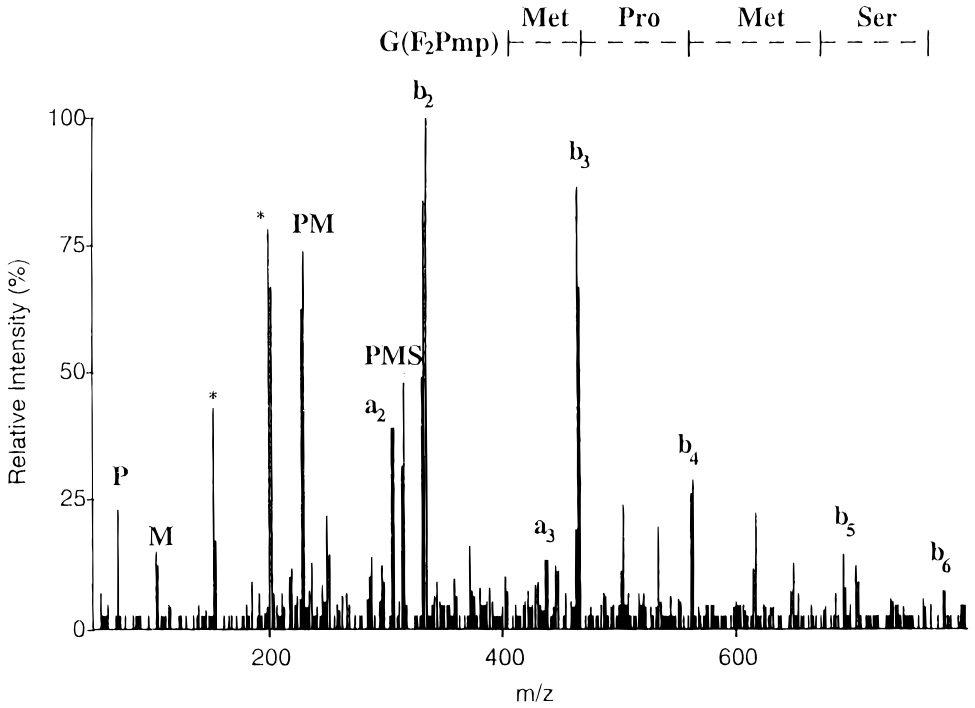


FIGURE 6: ES-MS/MS spectrum of the molecular ion at *m/z* 798.5 from the pH 3.5 elution. Assignments are as designated in Figure 3, and asterisked peaks indicate methionine-specific fragmentation ions. The sequence was determined to be that designated above the b-ions of the mass spectrum.

of phosphonate and phosphate groups. Chaotrope elution is based upon displacement of hydrophobic interactions, a parameter that can vary enormously within a given library. While we have no reason to suggest that one method is preferable to another for the screening of libraries, it is clear that different rank orders of potency are possible and may carry useful information reflective of the active site and the nature of the binding interactions.

Based upon our work with other target systems, it is evident that affinity characteristics and elution techniques will differ when using various biological targets and various libraries. Recent data obtained in our laboratory indicate that the (phosphonodifluoromethyl)phenylalanine library described binds weakly to protein tyrosine phosphatase IB-GST. Under these conditions, most of the library was detected in the 1 M/pH 7.5 elution, and no peptides were retained by the enzyme at low pH. This indicated that F_2-Pmp has weak affinity for tyrosine phosphatase and that another type of elution (e.g., salt gradient) should be used to establish a rank order in this system. As stated for the SH2 affinity system, the acid elution was sufficient to selectively isolate strong binders at pH 3.5 and afforded the minimum cleanup prior to ES-MS analysis.

LAS-MS detection of library components has proven to be a powerful technique for sequence-specific determination of high affinity compounds. ES-MS is a rapid, sensitive technique that has great potential for this type of application due to its molecular specificity, while tandem instruments offer specific structural elucidation. Although there are ambiguities determining certain residues such as I/L and K/Q with a triple quadrupole mass spectrometer, these can be overcome to a degree by eliminating one of the residues from a given library or through the use of isotopic analogues. Other obstacles may present themselves due to the redundant molecular weights that are often present in the structural combinations contained in a peptide library. These too can be overcome by planning a theoretical library that will not contain such redundancies, potentially designed on a subset of an initial, inclusive library screen. In this respect, computer-controlled manipulation of data could prove to be extremely useful in both the planning of libraries and the potential automation of data analysis. Another approach would be to resynthesize individual components of redundant molecular weight for further testing. Furthermore, as combinatorial libraries increasingly move toward the use of unique building blocks (Gordon et al., 1994), these redundancies will become nonissues.

The analysis of intact combinatorial libraries should be a priority before carrying out bioassays, a practical consideration that is often overlooked. ES-MS allows for data to be obtained from which the quality of the library can be assessed. If necessary, a more extensive analysis can be performed on the library using ES-MS/MS. It is not the purpose of this paper to fully detail the completeness of the initial library; however, MS is the primary current method used to judge library fidelity, as has been detailed by recent reports (Dunayevskiy et al., 1995). Additionally, although mass spectral analysis is not strictly quantitative due to the variation in ionization efficiencies in mixtures of components, MS analysis of the initial library and of the affinity-selected fractions (combined with the spreadsheet format of the data readout) can be used to visualize a qualitative representation of fractions from a given library. Furthermore, binding affinities can be determined by resynthesizing and testing the affinity-selected compounds, in a natural progression of the research process.

There is no theoretical upper limit to the size of the library used for the LAS-MS method. The only requirement is that individual active components are present in roughly 10−50 pmol levels for subsequent mass spectral detection and sequencing, the practical limit of the instrumentation. Interestingly, more recent MS methods have demonstrated

sensitivity in the fmol range using nanoelectrospray tandem mass spectrometry (Wilm et al., 1996). The capacity of the SH2 column could potentially limit the library size, particularly if a large percent of the library components had high affinity for the target. In practice, we have not seen such limitations as <10% of sequences in the library were high affinity. It should be noted, however, that the expression system for the recombinant target can easily be scaled up to larger capacity.

The strength of ES-MS/MS detection used in conjunction with library affinity selection lies in its ability to determine the identity of molecular ions of interest. The information obtained directly correlates to individual molecules rather than to a selected mixture. For example, Edman sequencing of affinity-selected peptides is only able to sequence pooled, high affinity components as a mixture. When one amino acid is detected at a given position more frequently than others, it is shown to favor that position. ES-MS/MS can extend relative positional frequency determinations to include the precise pairing and sequence order of amino acids for individual components detected from library affinity selection. Even when molecular weight redundancies exist for paired combinations (for example, see Table 1), ES-MS/MS can determine the amino acid sequence and eliminate the other redundancies concurrently.

Many choices are available for someone interested in applying combinatorial techniques to a novel target, none of which is without its limitations. In the event that little is known about the substrate or binding interaction, it would be most appropriate to consider combinatorial methods that use vast libraries such as the phage library method, a consensus sequence method, or one of the encoded libraries, thus giving the opportunity to screen billions of potential interactions. At the other extreme, once information is known about the type of binding interaction, or once a lead structure is available, the optimization process lends itself nicely to the LAS-MS or the multipin arrays. The choice of methods is further complicated by the influence of the immobilization of library components, the complexities of encoding chemistries, and the availability of expertise and equipment. The potential for artifacts exists, and each method is in the process of being further developed and optimized. As the LAS-MS technique is further utilized and applied to nonpeptide libraries, we anticipate that its potential will be widely demonstrated.

The present work demonstrates how library affinity selection with mass spectrometric detection is possible utilizing the basic resources available in most research facilities, while addressing the many practical considerations that have been identified in the field of combinatorial chemistry. This methodology is ideal for the analysis of soluble, unbound library components, which can be of unlimited design with respect to small-molecule libraries. Significant aspects of the current method include the simplicity of the experimental design using a peptide trap, the ability to analyze the integrity of an intact combinatorial library, the practicality of a well-planned library related to the affinity system, the confirmation that F_2Pmp can replace phosphotyrosine for SH2 binding, the selectivity of the method with the possibility for ranking, and the key advantage of ES-MS/MS for amino acid sequencing, applied to unnatural amino acids.

ACKNOWLEDGMENT

We thank the following for their help with review of the manuscript: E. Felder, H. Rink, and P. Robertson. We also thank E. Gordon and D. Patel for supplying the phosphotyrosine mimic and M. Thoma for computer support.

REFERENCES

Brenner, S., & Lerner, R. A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 5381−5383.
Brummel, C. L., Lee, I. N., Zhou, Y., Benkovic, S. J., & Winograd, N. (1994) *Science 264*, 399−402.
Brummel, C. L., Vickerman, J. C., Carr, S. A., Hemling, M. E., Roberts, G. D., Johnson, W., Weinstock, J., Gaitanopoulos, D., Benkovic, S. J., & Winograd, N. (1996) *Anal. Chem. 68*, 237−242.
Burke, T. R., Jr., Smyth, M. S., Otaka, A., Nomizu, M., Roller, P. P., Wolf, G., Case, R., & Shoelson, S. E. (1994) *Biochemistry 33*, 6490−6494.
Chen, J. K., Lane, W. S., Brauer, A. W., Tanaka, A., & Schreiber, S. L. (1993) *J. Am. Chem. Soc. 115*, 12591−12592.
Cull, M. G., Miller, J. F., & Schatz, P. J. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1865−1869.
Cwirla, S. E., Peters, E. A., Barrett, R. W., & Dower, W. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6378−6382.
Dooley, C. T., & Houghten, R. A. (1993) *Life Sci. 52*, 1509−1517.
Dooley, C. T., Chung, N. N., Schiller, P. W., & Houghten, R. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 10811−10815.
Dunayevskiy, Y., Vouros, P., Carell, T., Wintner, E. A., & Rebek, J. J. (1995) *Anal. Chem. 67*, 2906−2915.
Felder, E. R. (1994) *Chimia 48*, 531−541.
Fodor, S. P., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T., & Solas, D. (1991) *Science 251*, 767−773.
Furka, A., Sebestyen, F., Asgedom, M., & Dibo, G. (1988) in *14th Int. Congr. Biochem.* (*Prague*), 47.
Gallop, M. A., Barrett, R. W., Dower, W. J., Fodor, S. P., & Gordon, E. M. (1994) *J. Med. Chem. 37*, 1233−1251.
Goodson, R. J., Doyle, M. V., Kaufman, S. E., & Rosenberg, S. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 7129−7133.
Gordon, E. M., Barrett, R. W., Dower, W. J., Fodor, S. P., & Gallop, M. A. (1994) *J. Med. Chem. 37*, 1385−1401.
Herbst, J. J., Andrews, G., Contillo, L., Lamphere, L., Gardner, J., Lienhard, G. E., & Gibbs, E. M. (1994) *Biochemistry 33*, 9376−9381.
Houghten, R. A. (1993) *Trends Genet. 9*, 235−239.
Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T., & Cuervo, J. H. (1991) *Nature 354*, 84−86.
Kaur, S., Huebner, V., Tang, D., McGuire, L., Drummond, R., Csetjey, J., Stratton-Thomas, J., Rosenberg, S., Figliozzi, G., Banville, S., Zuckermann, R., & Dollinger, G. (1995) in *43rd ASMS Conference on Mass Spectrometry and Allied Topics*, p 30, Atlanta, GA.
Kerr, J. M., Banville, S. C., & Zuckermann, R. N. (1993) *J. Am. Chem. Soc. 115*, 2529−2531.
Klippel, A., Escobedo, J. A., Fantl, W. J., & Williams, L. T. (1992) *Mol. Cell. Biol. 12*, 1451−1459.
Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M., & Knapp, R. J. (1991) *Nature 354*, 82−84.
Lebl, M., Krchnak, V., Sepetov, N. F., Seligmann, B., Strop, P., Felder, S., & Lam, K. S. (1995) *Biopolymers 37*, 177−198.
Metzger, J. W., Kempter, C., Wiesmuller, K. H., & Jung, G. (1994) *Anal. Biochem. 219*, 261−277.
Moran, M. F., Koch, C. A., Anderson, D., Ellis, C., England, L., Martin, G. S., & Pawson, T. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8622−8626.
Needels, M. C., Jones, D. G., Tate, E. H., Heinkel, G. L., Kochersperger, L. M., Dower, W. J., Barrett, R. W., & Gallop, M. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 10700−10704.
Ohlmeyer, M. H., Swanson, R. N., Dillard, L. W., Reader, J. C., Asouline, G., Kobayashi, R., Wigler, M., & Still, W. C. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 10922−10926.
Panayotou, G., Gish, G., End, P., Truong, O., Gout, I., Dhand, R., Fry, M. J., Hiles, I., Pawson, T., & Waterfield, M. D. (1993) *Mol. Cell. Biol. 13*, 3567−3576.

Pawson, T. (1995) *Nature 373*, 573−580.
Piccione, E., Case, R. D., Domchek, S. M., Hu, P., Chaudhuri, M., Backer, J. M., Schlessinger, J., & Shoelson, S. E. (1993) *Biochemistry 32*, 3197−3202.
Pinilla, C., Appel, J. R., & Houghten, R. A. (1993) *Gene 128*, 71−76.
Salmon, S. E., Lam, K. S., Lebl, M., Kandola, A., Khattri, P. S., Wade, S., Patek, M., Kocis, P., Krchnak, V., Thorpe, D., et al. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 11708−11712.
Schlessinger, J. (1994) *Curr. Opin. Genet. Dev. 4*, 25−30.
Smith, G. P. (1985) *Science 228*, 1315−1317.
Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J., et al. (1993) *Cell 72*, 767−778.
Songyang, Z., Shoelson, S. E., McGlade, J., Olivier, P., Pawson, T., Bustelo, X. R., Barbacid, M., Sabe, H., Hanafusa, H., Yi, T., et al. (1994) *Mol. Cell. Biol. 14*, 2777−2785.
Stahl, M. L., Ferenz, C. R., Kelleher, K. L., Kriz, R. W., & Knopf, J. L. (1988) *Nature 332*, 269−272.
Till, J. H., Annan, R. S., Carr, S. A., & Miller, W. T. (1994) *J. Biol. Chem. 269*, 7423−7428.
Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., & Mann, M. (1996) *Nature 379*, 466−469.
Ye, B., Akamatsu, M., Shoelson, S. E., Wolf, G., Giorgetti-Peraldi, S., Yan, X., Roller, P. P., & Burke, T. R., Jr. (1995) *J. Med. Chem. 38*, 4270−4275.
Youngquist, R. S., Fuentes, G. R., Lacey, M. P., & Keough, T. (1994) *Rapid Commun. Mass Spectrom. 8*, 77−81.
Youngquist, R. S., Fuentes, G. R., Lacey, M. P., & Keough, T. (1995) *J. Am. Chem. Soc. 117*, 3900−3906.

BI960571X