SH2 Domains from Suppressor of Cytokine Signaling-3 and Protein Tyrosine Phosphatase SHP-2 Have Similar Binding Specificities[†]

David De Souza,[‡,§] Louis J. Fabri,[‡,||] Andrew Nash,[‡,||] Douglas J. Hilton,[‡,§] Nicos A. Nicola,[‡,§] and Manuel Baca*[,‡,§]

The Walter and Eliza Hall Institute of Medical Research, and The Cooperative Research Centre for Cellular Growth Factors, Parkville, Victoria 3050, Australia, and AMRAD Operations Pty. Ltd., Richmond, Victoria 3121, Australia

Received April 12, 2002

ABSTRACT: Suppressor of cytokine signaling-3 (SOCS-3) and the protein tyrosine phosphatase SHP-2 both regulate signaling by cytokines of the interleukin-6 family, and this is dependent upon recruitment to tyrosine 757 in the shared cytokine receptor subunit gp130. To better explore the overlap in ligand binding specificities exhibited by these two signaling regulators, we have mapped the phosphopeptide binding preferences of the SH2 domains from SOCS-3 and SHP-2. Degenerate phosphopeptide libraries were screened against recombinantly produced SH2 domains to determine the sequences of optimal phosphopeptide ligands. We found that the consensus ligand binding motif for SOCS-3 was pY-(S/A/V/Y/F)-hydrophobic-(V/I/L)-hydrophobic-(H/V/I/Y), while the consensus motif for SHP-2 was pY-(S/T/A/V/I)-X-(V/I/L)-X-(W/F). We validated these data through the design of phosphopeptide ligands based on the consensus motifs and found that these bound to SOCS-3 and SHP-2 with high affinity. Finally, we have compared the affinity of SOCS-3 for binding to phosphopeptides representing putative docking sites in the gp130, leptin and erythropoietin receptors. While SOCS-3 binds with much higher affinity to a gp130 phosphopeptide than to phosphopeptides derived from the other receptors, multiple SOCS-3 binding sites are predicted to exist in the leptin and erythropoietin receptors which may compensate for weaker binding to individual sites.

The duration and intensity of cytokine-induced biological responses are tightly controlled, requiring a delicate balance between positive and negative effectors of signaling. Transient phosphorylation of specific tyrosine residues plays a key role in the transmission and suppression of cytokine signals, in large part by creating docking sites for proteins that contain Src homology 2 (SH2)[1] or phosphotyrosine-binding (PTB) domains. SH2 domains are protein modules of approximately 100 amino acids, and the specific interaction of an SH2-containing protein with its cognate partner is critically dependent on a phosphotyrosine (pY) residue in addition to the flanking amino acid sequence (*1*). While the pY residue interacts with conserved portions of the SH2 domain, residues immediately C-terminal to the pY residue interact with variable regions of the SH2 domain, thus leading to recognition of different phosphorylated proteins by different SH2 domains (*2*−*6*). A detailed understanding of the binding specificity of SH2 domains can therefore be instructive in determining upon which signaling pathways different SH2-containing proteins act.

The suppressor of cytokine signaling-3 (SOCS-3) is an SH2-containing protein that regulates signaling by a variety of cytokines (for a review, see ref *7*). SOCS-3 is a member of the SOCS family of proteins and is characterized by a central SH2 domain required for the recruitment to phosphorylated cytokine receptors, a C-terminal SOCS box motif which is believed to play a role in targeting SOCS-associated proteins to the proteasome and an N-terminal region believed to be important for the inhibition of Janus kinase (JAK) catalytic activity (*8*−*10*). Transcription of the *SOCS-3* gene, like other members of the *SOCS* family, is rapidly induced in response to cytokine, suggesting that this protein acts in a negative feedback loop to regulate cytokine signal transduction (*8*).

Recently, we and others characterized a high-affinity binding site for SOCS-3 on the shared cytokine receptor subunit gp130, the common coreceptor used by cytokines of the interleukin-6 family (*11*, *12*). This site, corresponding to the region of murine gp130 surrounding tyrosine 757 (pY757), had been previously characterized as the docking site for SHP-2, an SH2-containing protein tyrosine phosphatase that is involved in positive and negative regulation of cytokine signaling (*13*−*16*). Given that SOCS-3 and SHP-2 recognize the same binding site on gp130, we had proposed that the SH2 domains from these proteins must

[†] Supported by the National Health and Medical Research Council (Australia), the National Institutes of Health, Bethesda (CA22556), and the Cooperative Research Centres Program of the Australian Government and AMRAD Corporation, Melbourne. M.B. is supported by the Australian Research Council and is an Australian Research Fellow.

* Corresponding author address: The Walter and Eliza Hall Institute of Medical Research, Post Office, Royal Melbourne Hospital, Parkville, Victoria 3050, Australia. Phone: +61 3 9345 2510. Fax: +61 3 9345 2616. E-mail: baca@wehi.edu.au.

[‡] The Walter and Eliza Hall Institute of Medical Research.

[§] The Cooperative Research Centre for Cellular Growth Factors.

[||] AMRAD Operations Pty. Ltd.

[1] Abbreviations: SH2, Src homology 2; pY, phosphotyrosine; SOCS, suppressor of cytokine signaling, SHP-2N, N-terminal SH2 domain of SHP-2; SHP-2C, C-terminal SH2 domain of SHP-2; SHP-2NC, tandem SH2 domain construct of SHP-2; PBS, phosphate-buffered saline; TFA, trifluoroacetic acid; NHS, *N*-hydroxysuccinimidyl; gp130(750−764), gp130-derived phosphopeptide residues 750−764.

10.1021/bi0259507 CCC: $22.00
Published on Web 06/19/2002

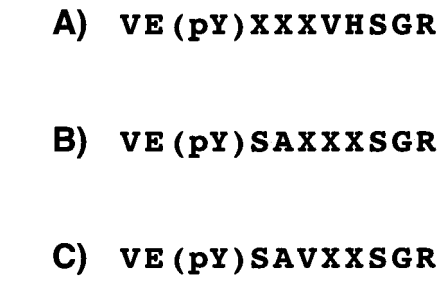

FIGURE 1: Sequences of degenerate phosphopeptide libraries. Each library was based on the sequence surrounding Y757 in gp130 and had two or three randomized positions, denoted by "X", containing an equal mixture of 19/20 natural L-amino acids. Positions pY+1, pY+2, and pY+3 were randomized in library A; positions pY+3, pY+4, and pY+5 in library B; and positions pY+4 and pY+5 in library C. The total degeneracy of libraries A and B was $19^3 = 6859$ peptides and $19^2 = 361$ peptides for library C.

share a similar ligand binding specificity and might therefore bind to common sites on other cytokine receptors. Indeed, the SHP-2 binding sites on the leptin and erythropoietin receptors have subsequently been shown to act as docking sites for SOCS-3 (*17*, *18*). It is therefore important to compare the ligand binding properties of the SH2 domains from SHP-2 and SOCS-3, as well as assess the affinity of SOCS-3 for binding to phosphopeptides derived from the putative docking sites in the gp130, leptin, and erythropoietin receptors.

Here, we report the phosphopeptide binding specificity for the SH2 domains from SHP-2 and from SOCS-3. Degenerate phosphopeptide libraries were screened to reveal preferred ligand sequences, and these data were validated via the design of phosphopeptide ligands which bound to SOCS-3 and SHP-2 with high affinity. We have also investigated the binding of SOCS-3 to phosphopeptides derived from the putative docking sites in the gp130, leptin, and erythropoietin receptors and have used the specificity data to predict additional docking sites that may exist in these receptors.

EXPERIMENTAL PROCEDURES

SH2 Domain Expression. cDNA's encoding the N-terminal (amino acids 1−105), C-terminal (amino acids 109−220), or tandem (amino acids 1−220) SH2 domains from human SHP-2 were subcloned into a pET15b vector (Novagen) and expressed as hexahistidine-tagged proteins (hereafter referred to as SHP-2N, SHP-2C, and SHP-2NC respectively). Constructs encoding either full length murine SOCS-3 or the SH2 domain (amino acids 22−185) were similarly prepared. The recombinant proteins were purified from bacterial cells (lysed in 7 M guanidinium hydrochloride) using immobilized nickel ion affinity chromatography. Purified proteins were refolded by dialysis into phosphate-buffered saline (PBS) containing 0.02% (v/v) Tween-20 and 0.5 mM tris(2-carboxyethyl)-phosphine. The amino acid sequences of the SH2 domains from human and murine SHP-2 are identical, while human and murine SOCS-3 share 96% sequence identity.

Design and Synthesis of Phosphopeptide Libraries. Partially degenerate phosphopeptide libraries were designed to probe the ligand binding specificity of the SH2 domains from SHP-2 and SOCS-3. Three such libraries were synthesized (Figure 1), and each degenerate position within a library contained an equal mixture of all natural L-amino acids with the exception of cysteine. The core sequence of each library was based on the sequence surrounding tyrosine 757 in murine gp130. The libraries were synthesized on a Rink amide methylbenzhydrylamine resin using Fmoc amino acids and *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate activation. All nondegenerate positions were double-coupled, and following amino acid coupling, the resin was treated with acetic anhydride to acetylate any remaining free amine groups. Phosphotyrosine was coupled as the Fmoc-*O*-benzyl-L-phosphotyrosine derivative (Novabiochem). For each degenerate position, the terminal Fmoc group was removed and the resin was dried and split into 19 equal lots. To each resin aliquot an appropriate activated Fmoc amino acid was coupled for 2 h after which time the resin samples were recombined, washed with dimethylformamide, and treated with acetic anhydride. This procedure was repeated for each degenerate position. Cleavage and deprotection of the peptide library was effected by treatment with 95% (v/v) trifluoroacetic acid (TFA) containing 5% (v/v) triisopropylsilane. TFA was removed by evaporation with a stream of nitrogen, the peptide library mixture was precipitated with diethyl ether, solubilized in 50% (v/v) aqueous acetonitrile containing 0.1% (v/v) TFA, filtered, and lyophilized. The sequences of the synthetic libraries were confirmed by N-terminal sequencing, and each degenerate position was found to contain an approximately equal proportion of each of the 19 different amino acids.

Phosphopeptide Library Screening. SHP-2N, SHP-2C, or the SOCS-3 SH2 domain were covalently immobilized onto *N*-hydroxysuccinimidyl (NHS)-activated Sepharose (Amersham Biosciences) in refolding buffer at a density of 3 mg of protein per mL of activated resin. "Control" resin was prepared by derivatizing NHS-activated Sepharose with ethanolamine. Immobilized protein resin (250 μL) was incubated with 4 mg of phosphopeptide library solubilized in 2 mL 50 mM sodium phosphate (pH 7.5) containing 150 mM NaCl, 2 mM DTT, and 0.2% (v/v) Tween-20. The mixture was gently shaken at room temperature for 1 h and then transferred to a disposable 10 mL chromatography column. After the peptide solution was allowed to drain, the resin was washed with 10 × 8 mL PBS containing 0.2% (v/v) Tween-20 and 1 × 8 mL 2 mM sodium phosphate (pH 7.5), using nitrogen pressure to reduce the total washing time to ∼15 min. Bound phosphopeptides were eluted from the resin with 2 × 250 μL of 1% (v/v) aqueous TFA, and the combined eluate was lyophilized. Control experiments were performed in the same way using "control" resin in place of the SH2 protein resin. Dried samples were resuspended in 60 μL of 30% (v/v) aqueous acetonitrile containing 0.1% (v/v) TFA and sequenced on an Applied Biosystems 494 Procise protein sequencer. To correct for the differential recoveries during sequencing, variable abundance of amino acids in the degenerate mixture, and potential differences in nonspecific adsorption to the control resin, the ratio of each amino acid at a given cycle of the phosphopeptide mixture retained on the SH2 resin to the same amino acid in the same cycle of the peptide mixture from the control experiment was calculated (*1*).

Biosensor Analysis. The biotinylated phosphopeptide biotin-STASTVE(pY)STVVHSG [gp130(750−764)] was immobilized onto a streptavidin-coated biosensor chip (SA, Biacore) at a density of ∼200 response units. Binding of recombinant SHP-2 or SOCS-3 constructs was measured by

diluting protein solutions into PBS containing 0.1% (v/v) Tween-20 and injecting these over the chip at a flow rate of 15 μL/min. For kinetic analysis of the binding, protein samples were analyzed immediately after dilution to minimize losses of protein due to nonspecific adsorption effects. After each binding measurement, residual protein was desorbed from the chip surface by injecting 6 M guanidinium hydrochloride (pH 8.0), followed by PBS containing 0.1% (v/v) Tween-20. Binding profiles were analyzed by using BIAEVALUATION software version 3.0 (Biacore). To correct for nonspecific binding effects, the sensorgrams for binding to the noncognate phosphopeptide biotin-FNSKDQM-(pY)SDGNFTD [gp130(674−688)] were subtracted from those for binding to the cognate phosphopeptide. The equilibrium dissociation constant K_D was calculated from a Scatchard-type analysis of the equilibrium response measurements obtained at different concentrations of protein (*19*).

For competition binding assays, serial dilutions of soluble competing peptides were incubated overnight with a fixed subsaturating concentration of full length SOCS-3 (100 nM) or SHP-2NC (50 nM) in PBS containing 0.1% (v/v) Tween-20 and 0.5 mg/mL BSA. These samples were then injected over the gp130(750−764) phosphopeptide chip and the level of bound protein was recorded at a fixed time point within the sensorgram and compared to the corresponding level of bound protein in the absence of any competitor. These fractional binding values (f) were fitted to the equation $f = 1/(1+(c/IC_{50})^m)$, where c = the concentration of soluble peptide ligand, m = the curvature constant, and IC_{50} values correspond to the concentration of ligand required to displace 50% of the bound protein.

RESULTS

Design and Screening of Phosphopeptide Libraries. A series of randomized phosphopeptide libraries were designed and synthesized based on the sequence surrounding the murine gp130 receptor at tyrosine 757 (Figure 1). Each of these libraries contained either two or three randomized amino acid positions within the five residues immediately downstream from the phosphotyrosine. For each randomized site, 19/20 natural amino acids were represented, cysteine being omitted to avoid problems with disulfide formation and sequencing detection. The molecular diversity of these libraries was either 361 or 6859 different peptide sequences depending on whether two or three amino acid positions were randomized.

Our procedure for screening these libraries to identify peptides which bind to recombinantly produced SH2 domains was similar to that used by Songyang et al. (*1*) but with several notable changes. Rather than express GST-SH2 domain fusion proteins for noncovalent immobilization on glutathione resin, we produced hexahistidine-tagged proteins. This produced a much higher yield of purified recombinant protein, particularly of SOCS-3, as the cell lysis and purification steps were performed in the presence of guanidinium hydrochloride which facilitated the solubilization of protein contained within inclusion bodies. The purified and refolded proteins were covalently coupled to NHS-activated Sepharose resin, and the function of the immobilized proteins was confirmed by demonstrating specific capture of a fluorescently labeled gp130(750−764) phosphopeptide (fluorescein-STASTVEpYSTVVHSG, data not shown). Covalent attachment of the SH2 domain onto resin permitted a more stringent washing protocol to be used during the library screening procedure, and bound phosphopeptides could be more aggressively eluted using dilute TFA which was found to be far more efficient than competitive elution by phenyl phosphate (*1*). Typically, screening experiments resulted in the capture of ~2000 pmol of specifically bound phosphopeptides as detected by N-terminal sequencing, while equivalent control experiments (using ethanolamine-derivatized Sepharose) captured less than 50 pmol of nonspecifically bound material.

Analysis of SHP-2 Selected Sequences. Screening of libraries A and B (Figure 1) against the N- and C-terminal SH2 domains of SHP-2 (i.e., SHP-2N, SHP-2C) revealed the preferred ligand sequences for the five residues immediately following phosphotyrosine (i.e., positions pY+1 to pY+5; Figure 2). These data showed that the two SH2 domains from SHP-2 bound to highly related sequences and that the pY+1, pY+3, and pY+5 positions seem to be the most selective determinants for SHP-2 binding. The preferred amino acids were threonine, alanine, and valine at the pY+1 position; valine, isoleucine, and leucine at pY+3; and tryptophan and phenylalanine at pY+5. Some minor differences were detected between the selectivity of each SH2 domain, such as the relative preference for isoleucine versus serine at the pY+1 position. In contrast to the strong selection of specific amino acids in the pY+1, pY+3, and pY+5 positions, the pY+2 and pY+4 positions appeared tolerant of many amino acid types. The specificity profile for the pY+3 position was the same, regardless of which library (i.e., library A or B; Figure 1) was used, indicating that preferences at this position are unlikely to be influenced by the identity of the flanking ligand sequence.

Analysis of SOCS-3 Selected Sequences. Screening of libraries A and C (Figure 1) against the SH2 domain from SOCS-3 was used to determine preferred phosphopeptide ligand binding sequences. (Figure 3). Our initial attempts to simultaneously define the pY+3, pY+4, and pY+5 ligand preferences with library B gave very low recoveries of phosphopeptide from the affinity column, presumably because too few peptides in the mixture had sufficient affinity to remain bound during the washing procedure. For this reason, a simplified library (library C, Figure 1) was generated in which only the pY+4 and pY+5 positions were randomized, and this resulted in much greater capture of peptides from the library mixture. Analysis of the screening data showed a general preference for hydrophobic amino acids at each of positions pY+1 through pY+5, in addition to strong selection of histidine at pY+5. The most selective position appeared to be pY+3, where approximately 65% of the selected sequences had either valine, isoleucine, or leucine. Interestingly, SOCS-3 appears to exhibit defined preferences at each ligand position between pY+1 and pY+5 in contrast to the SH2 domains from SHP-2 which did not show strong selectivity for specific amino acids at the pY+2 and pY+4 positions of the ligand.

Relative Affinities for gp130-Derived Phosphopeptide. On the basis of a comparison of the specificity data, the sequence surrounding the Y757 docking site in murine gp130 (-pYSTV-VH-) seems better optimized for SOCS-3 binding compared to either of the SH2 domains from SHP-2. To test this, we

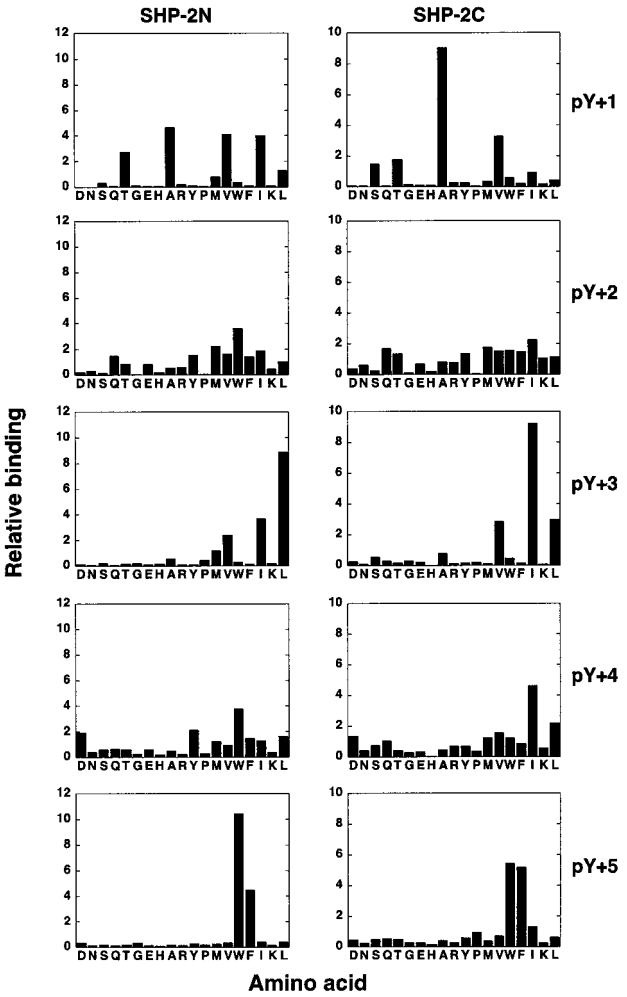

FIGURE 2: Phosphopeptide binding specificities of the SH2 domains from SHP-2. The degenerate phosphopeptide libraries A and B (Figure 1) were incubated with Sepharose resin containing either immobilized SHP-2N or SHP-2C. The column was washed, and bound peptides were eluted with aqueous TFA. The eluted peptide mixture was analyzed by N-terminal sequencing, and the results were compared to those from the eluate of Sepharose resin alone. Data are shown for the relative amino acid composition at each of the randomized positions within the libraries. The data have been normalized so that the sum of all values for a given position equals the number of amino acids (i.e., 19). Data for the pY+3 position is derived from screening library B, although data from library A (not shown) was very similar.

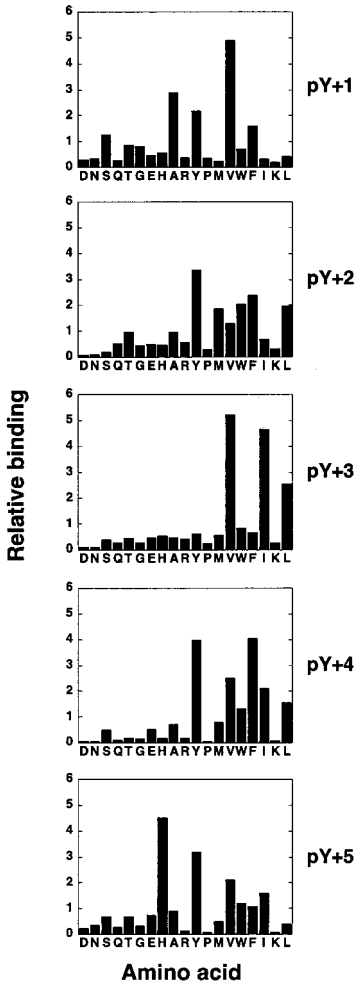

FIGURE 3: Phosphopeptide binding specificity of the SH2 domain from SOCS-3. The degenerate phosphopeptide libraries A and C (Figure 1) were incubated with Sepharose resin containing immobilized SOCS-3 SH2 domain. Experimental and data analysis procedures were as described in the Material and Methods section and Figure 2 legend.

Table 1: Binding Affinity of gp130(750−764) Phosphopeptide to SOCS-3 and SHP-2 Protein Constructs

protein construct	K_D, nM	protein construct	K_D, nM
SOCS-3 full length	42[a]	SHP-2N	1200
SOCS-3(22−185)	140	SHP-2C	550

[a] Taken from Nicholson et al. (*11*).

measured the affinity of each of the SH2 domains for binding to a biotinylated gp130(750−764) phosphopeptide corresponding to the Y757 docking site (Table 1). Previous measurements of the binding affinity of full length SOCS-3 for this phosphopeptide (*11*) showed that the equilibrium dissociation constant (K_D) was 42 nM. We also measured the affinity of the SOCS-3 SH2 domain construct [SOCS-3(22−185)] used in the current screening studies, and a similar K_D value (140 nM) was obtained. The K_D's of the individual SH2 domains from SHP-2 were 1200 nM for SHP-2N and 550 nM for SHP-2C, demonstrating that binding of these domains to the gp130-derived phosphopeptide is weaker than SOCS-3 but similar to each other as expected based on their similar ligand binding preferences.

We also analyzed binding by the SHP-2NC tandem SH2 domain construct to test whether bivalent interaction by the dual domains, as found in wild-type SHP-2, increases the affinity of binding. Consistent with this expectation, the K_D for this interaction as calculated by biosensor analysis was 12 nM, that is, considerably higher affinity binding than by either SH2 domain on its own (Figure 4A). However, when a nonbiotinylated form of the gp130(750−764) phosphopeptide was used as a soluble competitor to displace SHP2-NC from the same immobilized phosphopeptide, the IC_{50} value of this displacement was 3000 nM, that is, 250-fold higher than the K_D. We suspected that the large discrepancy between K_D and IC_{50} values was a consequence of the tetrameric nature of streptavidin allowing two immobilized phosphopeptides to simultaneously engage the tandem SH2 domain construct, thereby resulting in the artificially low K_D. To test this possibility, low surface coverage of the biotinylated gp130(750−764) phosphopeptide was generated by mixing this peptide in a 1:9 (w/w) ratio with the nonbinding gp130(674−688) biotinylated phosphopeptide prior to

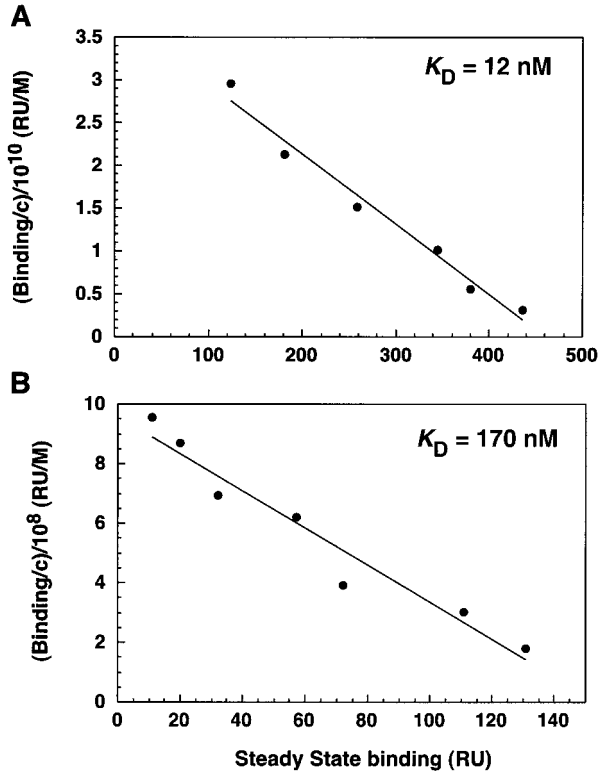

FIGURE 4: Binding of SHP-2NC to gp130(750−764) phosphopeptide under different immobilization conditions. (A) A 2-fold serial dilution series of SHP-2NC (4.20−135 nM) was injected over the streptavidin biosensor chip on which biotinylated gp130(750−764) phosphopeptide was immobilized. A Scatchard-type analysis of this binding data (*19*) was used to calculate the equilibrium dissociation constant (K_D). (B) Biotinylated gp130(750−764) phosphopeptide was mixed 1:9 with the nonbinding gp130(674−688) biotinylated phosphopeptide prior to immobilization on a streptavidin biosensor chip. A 2-fold serial dilution series of SHP-2NC (11.5−736 nM) was injected over the chip, and a Scatchard-type analysis of this binding data was used to calculate the K_D.

immobilization. In this way, capture of the specific phosphopeptide to streptavidin was essentially monomeric and subsequent binding of SHP-2NC would be free of avidity artifacts. Under these conditions, the K_D for binding of SHP-2NC to the gp130(750−764) phosphopeptide was 170 nM (Figure 4B). By contrast, the kinetics of SOCS-3 binding to this modified biosensor chip was essentially the same as observed on the standard gp130(750−764) phosphopeptide chip (data not shown).

Determination of Minimal Binding Epitope. Previous studies had shown that the shortest gp130 phosphopeptide which maintains maximum binding to SOCS-3 was an acetylated eight residue peptide spanning the pY−2 to pY+5 positions (*11*). To determine the same relationship between peptide length and affinity for SHP-2, we measured the affinities of a series of soluble truncated peptide analogues for binding to SHP-2NC in a competitive displacement assay (Table 2). These data showed that the dual SH2 domain-containing SHP-2 construct behaves much like SOCS-3 in that residues beyond the pY−2 or pY+5 position make no contribution to binding but that further truncation results in a reduction in binding affinity. Thus, the SH2 domains from SOCS-3 and SHP-2 have ligand binding sites which accommodate interactions with at least an eight residue ligand. On this basis, we prepared the eight residue phosphopeptide VE(pY)STVVH and measured its binding affinity for SHP-2 and SOCS-3. While the affinity for SHP-2 was the same as for longer peptides, the affinity for SOCS-3 was reduced about 3-fold, presumably because of differences between binding a peptide containing a free versus acetylated pY−2 residue at the amino terminus.

Table 2: Binding of gp130(750−764) Phosphopeptide Analogues to SOCS-3 and SHP-2NC[a]

peptide	SOCS-3 binding IC_{50}, nM	SHP-2NC binding IC_{50}, nM
STASTVE(pY)STVVHSG	110[b]	3000
STASTVE(pY)STVVHS	61[b] (0.55)	2400 (0.80)
STASTVE(pY)STVVH	77[b] (0.70)	2600 (0.87)
STASTVE(pY)STVV	1100[b] (10)	9800 (3.3)
Ac-TVE(pY)STVVHSG	120[b] (1.1)	2500 (0.83)
Ac-VE(pY)STVVHSG	110[b] (1.0)	3900 (1.3)
Ac-E(pY)STVVHSG	970[b] (8.8)	34000 (11.3)
Ac-(pY)STVVHSG	9600[b] (87)	41000 (13.7)
VE(pY)STVVH	340 (3.1)	2500 (0.83)

[a] Ac = acetyl; pY = phosphotyrosine. [b] Taken from Nicholson et al. (*11*). Values in parentheses are the ratio of IC_{50} analogue/IC_{50} gp130(750−764).

Design of High-Affinity Phosphopeptide Ligands. To validate the usefulness of the screening data in predicting the sequences of high-affinity ligands, two phosphopeptides were designed which were predicted to have enhanced affinity for either SOCS-3 or SHP-2 relative to the gp130-derived phosphopeptide. The designed SOCS-3 ligand, VE(pY)VYVIH, contained a total of three amino acid substitutions relative to the gp130-derived peptide and bound with approximately 8-fold higher affinity (Figure 5A). The sequence of the optimized SHP-2 ligand was designed to maximize interaction with both SH2 domains. This peptide, VE(pY)TQLLW, contained five amino acid substitutions relative to the gp130-derived sequence and bound to SHP-2NC with 32-fold higher affinity (Figure 5B). The improved affinities of these peptides for SOCS-3 and SHP-2 demonstrates the predictive usefulness of the screening data in identifying or designing high-affinity ligands for these SH2 domains.

Binding of Leptin and Erythropoietin Receptor Peptides to SOCS-3. SOCS-3 has been reported to bind to specific sites on each of the gp130, leptin, and erythropoietin receptors (*11*, *17*, *18*, *20*), but no comparison has been made of the relative binding affinity of SOCS-3 for each of these sites. To address this, we measured the relative affinity of SOCS-3 for binding to phosphopeptides correponding to each of these sites, that is, the regions surrounding Y757 on the gp130 receptor, Y985 and Y1077 on the leptin receptor, and Y425[2] on the erythropoietin receptor (Table 3). A comparison of the affinities of each peptide showed that SOCS-3 binds to the leptin and erythropoietin receptor-derived phosphopeptides with much lower affinity than to the gp130-derived phosphopeptide.

DISCUSSION

SOCS-3 and SHP-2 play critical roles in regulating cytokine signaling pathways. Deletion of either gene in mice results in embryonic lethality (*21*−*23*), highlighting their essential and nonredundant roles in cell signaling. SOCS-3

[2] Y425 in the erythropoietin receptor is alternatively referred to as Y401 by other authors.

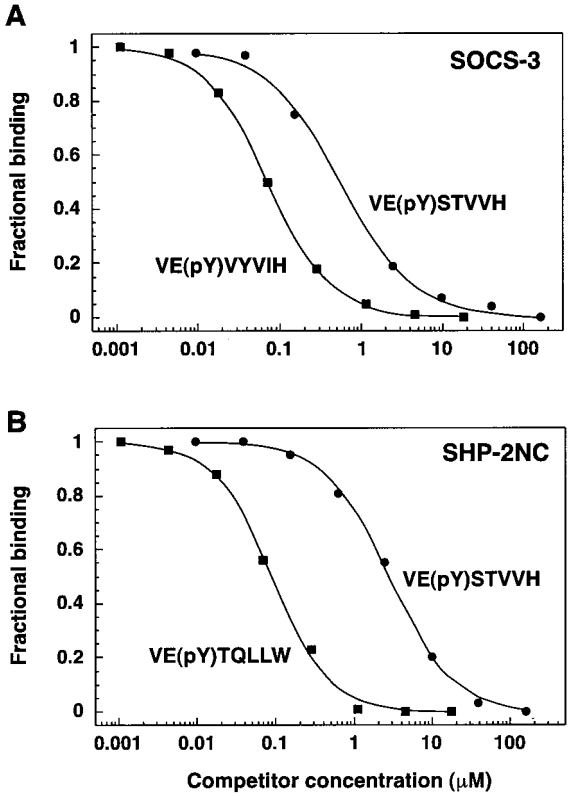

FIGURE 5: High-affinity binding by designed SOCS-3 and SHP-2 ligands. (A) Solution binding of phosphopeptides to SOCS-3. The IC_{50} for inhibition of SOCS-3 binding to immobilized gp130(750−764) phosphopeptide was 71 nM for soluble VE(pY)VYVIH as compared to 550 nM for VE(pY)STVVH. (B) Solution binding of phosphopeptides to SHP-2NC. The IC_{50} for inhibition of SHP-2NC binding to immobilized gp130(750−764) phosphopeptide was 87 nM for soluble VE(pY)TQLLW as compared to 2800 nM for VE-(pY)STVVH.

Table 3: Binding of Murine gp130, Leptin, and Erythropoietin Receptor Peptides to SOCS-3[a]

peptide	receptor source	IC_{50} (μM)
Ac-VE(pY)STVVHSG	gp130, Y757	0.10
Ac-VK(pY)ATLVSND	leptin R, Y985	8.3
Ac-VC(pY)LGVTSVN	leptin R, Y1077	26
Ac-FE(pY)TILDPSS	EPOR, Y425	69

[a] leptin R = leptin receptor; EPOR = erythropoietin receptor.

recognizes the same binding site on gp130 used by SHP-2, and recently, other reports have shown that SOCS-3 also binds to the SHP-2 docking sites on the leptin and erythropoietin receptors (*17*, *18*, *20*). The shared use of these receptor binding sites has raised questions as to the relative roles that SOCS-3 and SHP-2 play in regulating cytokine signaling (*11*), and a better understanding of the ligand binding specificity of each protein may allow for the design of SOCS-3 or SHP-2-specific receptors or peptide antagonists to help address this issue. To determine the phosphopeptide binding specificity of SOCS-3 and SHP-2, we have screened degenerate libraries to define the sequences of preferred peptide ligands. A comparison of the data for each of the two SH2 domains from SHP-2 shows that the ligand binding specificities are highly related, with both domains showing a preference for binding to phosphopeptides containing a pY-(S/T/A/V/I)-X-(V/I/L)-X-(W/F) consensus motif. The similarity in specificity between the N- and C-terminal SH2 domains of SHP-2 is further illustrated by the similar affinity each domain exhibits for the gp130(750−764) phosphopeptide (K_D = 1200 and 550 nM respectively). As anticipated, the binding specificity of SOCS-3 was related to that of SHP-2. Optimal SOCS-3 phosphopeptide ligands were found to have the general consensus sequence pY-(S/A/V/Y/F)-hydrophobic-(V/I/L)-hydrophobic-(H/V/I/Y) which overlaps with the SHP-2 consensus sequence in the pY+1 and pY+3 positions, but differs significantly at the pY+5 position.

SHP-2 is composed of a tandem pair of SH2 domains followed by a C-terminal tyrosine phosphatase domain (*24*). Previous studies (*1*) had characterized the ligand binding specificity of the N-terminal SH2 domain of SHP-2; however, only the pY+1 to pY+3 positions were examined, and the C-terminal SH2 domain was not studied. This earlier work had found isoleucine and valine to be the most favored residues in the pY+1 and pY+3 positions of phosphopeptide ligands, while leucine and proline were also selected at pY+3. Our data is mostly in agreement with this, except we also observed the strong selection of SHP-2N-binding peptides which contained threonine and alanine in the pY+1 position, while proline was not strongly selected at pY+3. These differences may have resulted from the use of peptide libraries containing different sequences in the nonrandomized regions outside pY+1 to pY+3; however, threonine is found in the pY+1 position of several known SHP-2 docking sites (*25*). Importantly, the current data shows that both SH2 domains from SHP-2 recognize very similar sequences, which is perhaps unsurprising given that the C-terminal SH2 domain shares 48% sequence identity with the N-terminal domain. This result implies that both SH2 domains can bind the same monophosphorylated motifs, thereby resulting in an enhanced binding affinity, due to avidity effects, relative to a single SH2 domain interaction. Indeed, the SHP-2NC construct bound with higher affinity to the gp130(750−764) phosphopeptide relative to binding by the individual SH2 domains. Thus, SHP-2 not only shows enhanced binding to biphosphorylated motifs as previously reported (*26*) but also to monophosphorylated motifs.

The most selective position for binding to either of the SH2 domains from SHP-2 was the pY+5 position; ~80% of the phosphopeptides selected by the SHP-2N and 55% selected by the SHP-2C had either tryptophan or phenylalanine in this position. Interestingly, the related amino acid tyrosine was not significantly selected at pY+5. Crystallographic analysis (*27*) has shown that the side chain of the pY+5 amino acid can indeed form important contacts with the N-terminal SH2 domain in the case of a phosphopeptide containing phenylalanine at this position, but not one containing proline. Consistent with this, truncated peptides used in the current study revealed a 4-fold reduction in affinity when the pY+5 residue was removed (Table 2), while earlier work showed that substitution of the pY+5 phenylalanine from an insulin receptor substrate-1 (IRS1)-derived phosphopeptide resulted in a marked decrease in binding affinity (*27*).

The SH2 domains from SOCS-3 and SHP-2 are unlike most other SH2 domains in that they interact with residues N-terminal to phosphotyrosine in addition to the more usual C-terminal determinants. Unfortunately, our attempts to synthesize a phosphopeptide library containing randomized sequence at the pY−1 and pY−2 positions were unsuccessful

Table 4: Amino Acid Sequence Surrounding Cytoplasmic Tyrosine Residues within Murine Erythropoietin Receptor[a]

Tyrosine number	Sequence
Y367	Y L **V** **L** D K
Y425	Y T **I** **L** D P
Y453	Y L **Y** **L** **V** **V**
Y455	Y L **V** **V** S D
Y467	Y **S** S G G S
Y484	Y **S** H P **Y** E
Y488	Y E N S **L** **V**
Y503	Y **V** A C S

[a] Residues which conform to the consensus SOCS-3 binding motif pY-(S/A/V/Y/F)-hydrophobic-(V/I/L)-hydrophobic-(H/V/I/Y) are highlighted in boldface.

due to difficulties in synthesis, thus precluding an investigation of the ligand binding specificity at these positions. However, it is known that the SH2 domains of both SHP-2 and SOCS-3 share a preference for valine at the pY−2 residue, this resulting from substitution of the highly conserved αA2 arginine residue ordinarily found in SH2 domains with glycine (*11*, *28*). The similarity in the ligand binding characteristics of the SH2 domains of SOCS-3 and SHP-2 is therefore likely to extend to the identity of the pY−2 amino acid and thus to the length of the phosphopeptide binding grooves, with both proteins recognizing extended phosphopeptide ligands spanning the pY−2 to pY+5 positions.

The biological expression of SOCS-3 is induced in response to cytokine signaling and it subsequently acts to downregulate that signal (*8*). Recent data has shown that inhibition of interleukin-6, leptin, and erythropoietin signaling by SOCS-3 is dependent upon recruitment to specific phosphorylated receptor docking sites and that these are the same docking sites recognized by SHP-2 (*11*, *12*, *17*, *18*, *20*). To date, the affinity of SOCS-3 for binding to each of these receptor docking sites had not been compared. To assess the relative binding of SOCS-3 to these docking sites, we measured the binding affinity of phosphopeptides corresponding to each of these receptor sites. These data suggested that SOCS-3 would bind to the gp130 receptor with much higher affinity than to the leptin or erythropoietin receptors. However, while earlier studies had shown that SOCS-3 only bound to a single site on the gp130 receptor (*11*), the leptin receptor contains two binding sites at Y985 and Y1077, albeit each of lower affinity than gp130 binding. The presence of two binding sites on leptin receptor may compensate for the low affinity each individual site exhibits for SOCS-3, either because two SOCS-3 molecules can be simultaneously recruited or because the presence of two sites may result in higher affinity binding due to avidity effects. Only one erythropoietin receptor binding site has been defined for SOCS-3 at Y425 (*18*), but binding of SOCS-3 to this phosphopeptide was even weaker than binding to the leptin receptor peptides. However, a comparison of the erythropoietin receptor sequence with the consensus SOCS-3 binding motif (Table 4) suggests that additional SOCS-3 docking sites may exist. The motifs at Y367, Y425, Y453, and Y455 each contain a valine or leucine at the critical pY+3 position and may therefore act as SOCS-3 docking sites. In particular, the sequence surrounding Y453 (pY-LYLVV) appears an especially favorable SOCS-3 binding site given the high similarity between this sequence and the consensus SOCS-3 binding motif. Y453 is known to be phosphorylated during erythropoietin signaling and has been previously characterized as the docking site for SHP-1, a negative regulatory protein with homology to SHP-2 (*29*). Thus, like the leptin receptor, the erythropoietin receptor may have multiple sites which can recruit SOCS-3.

In conclusion, we have shown that the ligand binding specificities of SOCS-3 and SHP-2 are related, thus explaining their shared use of docking sites on different cytokine receptors. These data should prove useful in predicting new docking sites or distinguishing which of these may simultaneously recruit both proteins. In addition, knowledge of phosphopeptide ligand preferences may aid in the design of SOCS-3- or SHP-2-selective cytokine receptors or antagonists as tools for dissecting the relative roles each protein plays in regulating different signaling pathways.

REFERENCES

1. Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., and Lechleider, R. J. (1993) *Cell 72*, 767−778.
2. Waksman, G., Kominos, D., Robertson, S. C., Pant, N., Baltimore, D., Birge, R. B., Cowburn, D., Hanafusa, H., Mayer, B. J., Overduin, M., Resh, M. D., Rios, C. B., Silverman, L., and Kuriyan, J. (1992) *Nature 358*, 646−653.
3. Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D., and Kuriyan, J. (1993) *Cell 72*, 779−790.
4. Eck, M. J., Atwell, S. K., Shoelson, S. E., and Harrison, S. C. (1994) *Nature 368*, 764−769.
5. Eck, M. J., Shoelson, S. E., and Harrison, S. C. (1993) *Nature 362*, 87−91.
6. Pascal, S. M., Singer, A. U., Gish, G., Yamazaki, T., Shoelson, S. E., Pawson, T., Kay, L. E., and Formankay, J. D. (1994) *Cell 77*, 461−472.
7. Krebs, D. L., and Hilton, D. J. (2000) *J. Cell Sci. 113*, 2813−2819.
8. Starr, R., Willson, T. A., Viney, E. M., Murray, L. J., Rayner, J. R., Jenkins, B. J., Gonda, T. J., Alexander, W. S., Metcalf, D., Nicola, N. A., and Hilton, D. J. (1997) *Nature 387*, 917−921.
9. Zhang, J. G., Farley, A., Nicholson, S. E., Willson, T. A., Zugaro, L. M., Simpson, R. J., Moritz, R. L., Cary, D., Richardson, R., Hausmann, G., Kile, B. J., Kent, S. B., Alexander, W. S., Metcalf, D., Hilton, D. J., Nicola, N. A., and Baca, M. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 2071−2076.
10. Yasukawa, H., Misawa, H., Sakamoto, H., Masuhara, M., Sasaki, A., Wakioka, T., Ohtsuka, S., Imaizumi, T., Matsuda, T., Ihle, J. N., and Yoshimura, A. (1999) *EMBO J. 18*, 1309−1320.
11. Nicholson, S. E., De Souza, D., Fabri, L. J., Corbin, J., Willson, T. A., Zhang, J. G., Silva, A., Asimakis, M., Farley, A., Nash, A. D., Metcalf, D., Hilton, D. J., Nicola, N. A., and Baca, M. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 6493−6498.
12. Schmitz, J., Weissenbach, M., Haan, S., Heinrich, P. C., and Schaper, F. (2000) *J. Biol. Chem. 275*, 12848−12856.

13. Stahl, N., Farruggella, T. J., Boulton, T. G., Zhong, Z., Darnell, J. E., Jr., and Yancopoulos, G. D. (1995) *Science 267*, 1349−1353.
14. Symes, A., Stahl, N., Reeves, S. A., Farruggella, T., Servidei, T., Gearan, T., Yancopoulos, G., and Fink, J. S. (1997) *Curr. Biol. 7*, 697−700.
15. Tang, T. L., Freeman, R. M., O'Reilly, A. M., Neel, B. G., and Sokol, S. Y. (1995) *Cell 80*, 473−483.
16. Yamauchi, K., Milarski, K. l., Saltiel, A. R., and Pessin, J. E. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 664−668.
17. Bjorbaek, C., Lavery, H. J., Bates, S. H., Olson, R. K., Davis, S. M., Flier, J. S., and Myers, M. G. (2000) *J. Biol. Chem. 275*, 40649−40657.
18. Sasaki, A., Yasukawa, H., Shouda, T., Kitamura, T., Dikic, I., and Yoshimura, A. (2000) *J. Biol. Chem. 275*, 29338−29347.
19. Ward, L. D., Howlett, G. J., Hammacher, A., Weinstock, J., Yasukawa, K., Simpson, R. J., and Winzor, D. J. (1995) *Biochemistry 34*, 2901−2907.
20. Eyckerman, S., Broekaert, D., Verhee, A., Vandekerckhove, J., and Tavernier, J. (2000) *FEBS Lett. 486*, 33−37.
21. Marine, J. C., McKay, C., Wang, D., Topham, D. J., Parganas, E., Nakajima, H., Pendeville, H., Yasukawa, H., Sasaki, A., Yoshimura, A., and Ihle, J. N. (1999) *Cell 98*, 617−627.
22. Roberts, A. W., Robb, L., Rakar, S., Hartley, L., Cluse, L., Nicola, N. A., Metcalf, D., Hilton, D. J., and Alexander, W. S. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 9324−9329.
23. Saxton, T. M., Henkemeyer, M., Gasca, S., Shen, R., Rossi, D. J., Shalaby, F., Feng, G. S., and Pawson, T. (1997) *EMBO J. 16*, 2352−2364.
24. Freeman, R. M., Plutzky, J., and Neel, B. G. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 11239−11243.
25. Case, R. D., Piccione, E., Wolf, G., Benett, A. M., Lechleider, R. J., Neel, B. G., and Shoelson, S. E. (1994) *J. Biol. Chem. 269*, 10467−10474.
26. Ottinger, E. A., Botfield, M. C., and Shoelson, S. E. (1997) *J. Biol. Chem. 273*, 729−735.
27. Lee, C. H., Kominos, D., Jacques, S., Margolis, B., Schlessinger, J., Shoelson, S. E., and Kuriyan, J. (1994) *Structure 2*, 423−438.
28. Huyer, G., Li, Z. M., Adam, M., Huckle, W. R., and Ramachandran, C. (1995) *Biochemistry 34*, 1040−1049.
29. Klingmuller, U., Lorenz, U., Cantley, L. C., Neel, B. G., and Lodish, H. F. (1995) *Cell 80*, 729−738.

BI0259507