# *Communication*

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 271, No. 4, Issue of January 26, pp. 1813–1816, 1996


# Identification of Residues within the SHC Phosphotyrosine Binding/Phosphotyrosine Interaction Domain Crucial for Phosphopeptide Interaction*

(Received for publication, October 24, 1995, and in revised form, November 30, 1995)

**Vijay Yajnik‡, Pamela Blaikie‡, Peer Bork§, and Ben Margolis¶‖**

*From the ¶Howard Hughes Medical Institute and Department of Internal Medicine and Biological Chemistry, University of Michigan Medical School, Ann Arbor, Michigan 48109-0650, ‡Department of Pharmacology, New York University Medical Center, New York, New York 10016, and §Max Delbrueck Center for Molecular Medicine, Berlin-Buch, Germany and European Molecular Biology Laboratory, Heidelberg, Germany*

**Shc is an Src homology 2 (SH2) domain protein thought to be an important component of the signaling pathway leading from cell surface receptors to Ras. A new phosphotyrosine interaction (PI) domain (also known as the phosphotyrosine binding (PTB) domain) has been described in the amino terminus of Shc. The Shc PI domain binding specificity is dependent on residues lying amino-terminal to the phosphotyrosine rather than carboxyl-terminal as is seen with SH2 domains. We randomly mutagenized the Shc PTB/PI domain in an effort to identify residues in the domain crucial for interaction with phosphotyrosine-containing peptides. We then screened the mutants for binding to the tyrosine-phosphorylated carboxyl-terminal tail of the epidermal growth factor (EGF) receptor. Most striking were mutations that altered a phenylalanine residue in block 4 of the domain severely impairing PI domain function. This phenylalanine residue is conserved in all but one subfamily of PI domains that have been identified to date. Reconstitution of this phenylalanine mutation into full-length Shc created a protein unable to interact with the EGF receptor in living cells.**

Controlled protein interactions play an important role in signal transduction. One example is the interaction of cytoplasmic signaling proteins with tyrosine-phosphorylated growth factor receptors (1–3). One large class of molecules that binds to tyrosine-phosphorylated receptors contains SH2[1] domains. Recently, a second domain has been described that binds to tyrosine-phosphorylated peptides. This domain that we called the PI domain, also known as the phosphotyrosine binding (PTB) domain, was first described in the amino terminus of Shc (4–6). Shc also contains an SH2 domain and appears to be important in Ras activation. Shc becomes tyrosine-phosphorylated after cell activation with a large number of cytokines and growth factors (7–9) creating a high affinity binding site for the SH2 domain of Grb2. Grb2 is in turn complexed to the Son of Sevenless Ras guanine nucleotide exchange factor that converts Ras from its inactive GDP-bound state to its active GTP-bound state (1–3).

The exact role of the SH2 domain *versus* the PI domain in Shc signaling has not been determined. The binding of the PI domain to phosphotyrosine-containing peptides differs from SH2 domain binding in that its specificity is determined by residues lying amino-terminal to the phosphotyrosine (6, 10–15). Due to this difference in specificity it is assumed that the structural basis of interactions between phosphopeptides and the PI domain will differ from those used by SH2 domains. We have identified domains in several different proteins that have primary sequence similarity with the PI domain of Shc (16). In an effort to understand residues in the PI domain important in mediating binding to phosphorylated peptides, we performed random mutagenesis of the Shc PI domain and screened the mutants for binding to tyrosine-phosphorylated EGF receptor. We have identified several different mutations within the domain that impair its binding ability.

## EXPERIMENTAL PROCEDURES

*Random Mutagenesis of Shc Amino Acids 76–209*—Using the oligonucleotides (forward 5′CCG GAA TTC ATG AAC AAG CTG AGT GGA3′ and reverse 5′CCG CTC GAG TCA CAG GTC CTC CTC GCT GAT CAG CTT CTG CTC CTG CAG ATT CCT GAG ATA CTG TTT GAA3′) and plasmid Shc 1–209/pGSTag as a template (4), standard PCR was used to create a Shc 1–209 construct containing an *Eco*RI site on the 5′-end and a *Pst*I site followed by an in-frame Myc epitope tag, stop codon, and an *Xho*I site on the 3′-end (Fig. 1). This fragment was then cloned into the pTOPE vector (Novatope System, Novagen) using the *Eco*RI and *Xho*I restriction sites to create Shc 1–209/Myc pTOPE. The Shc 1–209 insert in pTOPE was then mutagenized by PCR (17) using pTOPE forward and reverse sequencing primers. The product of the PCR reaction (containing the mutagenized PI domain) was digested with *Bst*EII and *Pst*I restriction enzymes and ligated into Shc 1–209/Myc pTOPE (Fig. 1*A*). The ligation was transformed into competent Novablue (DE3, Novagen) *Escherichia coli* strain and selected on carbenicillin/tetracycline agar plates.

*Screening of Mutants*—Clones containing the mutagenized PI domain were plated onto 10-cm plates, and approximately 1500 independent recombinants were screened. Colonies were lifted onto nitrocellulose and lysed *in situ* according to the manufacturer's instruction for the Novatope system. Probing nitrocellulose filters with radiolabeled EGF receptor was performed as described previously (10). After probing with EGF receptor and detection of binding with autoradiography, the filters were then screened with an anti-Myc monoclonal antibody, 9E10 (18), using a biotin-avidin-linked alkaline phosphatase system (Vectastain, Vector Laboratories) as the readout. Clones that did not bind to the EGF receptor yet were positive for Myc antibody binding were selected for sequencing (Sequenase version 2, U. S. Biochemical Corp.). Of the 28 clones sequenced, 20 had single mutations, 6 had two mutations, 1 had three mutations, and 1 had no mutations. The clones with multiple mutations were not analyzed except for one of the clones with two mutations where the mutations were separated by restriction digest and subcloned to yield two clones with single mutations. In total, 22

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ Investigator of the Howard Hughes Medical Institute. To whom correspondence should be addressed: Howard Hughes Medical Inst., University of Michigan Medical School, Rm. 4570, MSRB II, 1150 W. Medical Center Dr., Ann Arbor, MI 48109-0650. Tel.: 313-764-3567; Fax: 313-763-9323; E-mail: bmargoli@uv1.im.med.umich.edu.

[1] The abbreviations used are: SH2, Src homology 2; EGF, epidermal growth factor; GST, glutathione *S*-transferase; PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; BSA, bovine serum albumin; PI, phosphotyrosine interaction; PTB, phosphotyrosine binding.

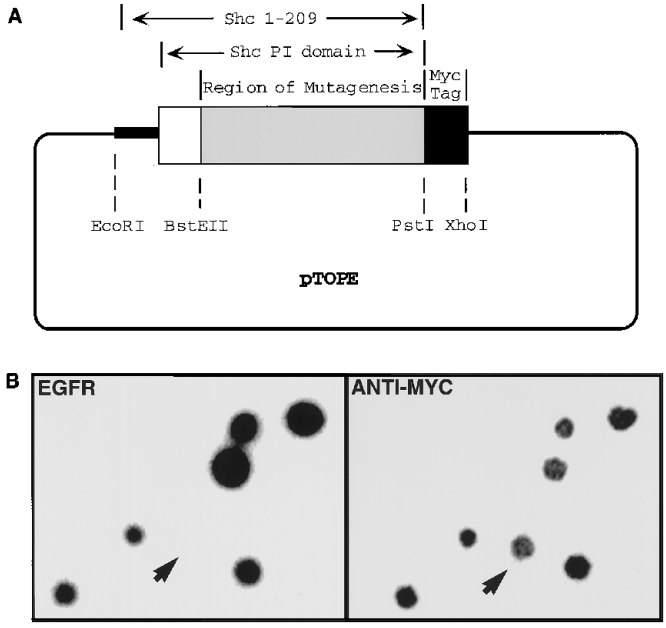


FIG. 1. *A*, Shc 1–209/Myc pTOPE vector. Shc 1–209 was subcloned into the pTOPE vector with a Myc epitope at its carboxyl terminus as described under "Experimental Procedures." The pTOPE vector allows protein expression in bacteria expressing T7 polymerase. cDNA encoding Shc 76–209 that had been subjected to random mutagenesis by PCR was cloned into the *Bst*EII and *Pst*I site. The mutagenesis did not include the Myc tag or pTOPE vector. *B*, screening of bacterial colonies expressing the mutagenized Shc PI domain. The mutagenized Shc 1–209/Myc pTOPE plasmids were used to transform bacteria and plated to yield approximately 150 colonies/100-mm plates. After overnight growth, the colonies were transferred to nitrocellulose and lysed. After blocking the filters, they were probed with the $^{32}$P-labeled carboxyl terminus of the EGF receptor (*EGFR*), washed, and exposed to film. After autoradiography, the filters were reprobed with an anti-Myc antibody using colorimetric detection. The autoradiograms were then aligned with the Myc antibody-stained filters. The *panel* on the *right* shows screening for Myc tag expression using monoclonal antibody 9E10 against the Myc epitope. The *panel* on the *left* shows binding of labeled EGF receptor to the same filter. The *arrow* depicts a colony selected for further analysis with normal binding of the Myc antibody but defective EGF receptor binding.

clones with single mutations were selected for further analysis.

*Dot Blot Analysis of Mutants*—The selected Shc 1–209 mutants were cloned into the GST fusion protein expression vector, pGEX 4T-1, using *Eco*RI and *Xho*I. GST fusion proteins were expressed and purified as described previously (19). Eluted fusion proteins were quantitated using both the Bio-Rad protein detection kit and SDS-PAGE alongside BSA standards. Serial dilutions of 200, 100, and 50 ng of fusion protein in a volume of 2.5 $\mu$l were spotted onto duplicate nitrocellulose filters. One set of filters was hybridized to $^{32}$P-labeled EGF receptor, and the other set was incubated with anti-Myc antibody to ensure an equal protein concentration for each fusion protein. The $^{32}$P-labeled EGF receptor that bound was quantitated by Cerenkov counting. For each mutant, the percent binding was calculated using wild type Shc 1–209 as a control.

*Binding Studies*—HER14 cell lysates were prepared as described previously (4). The lysates were incubated with GST fusion proteins bound to glutathione-agarose beads for 90 min at 4 °C. The beads were then washed three times with radioimmune precipitation buffer (1% sodium deoxycholate, 0.1% SDS, 50 mM Hepes, pH 7.5, 150 mM NaCl, 1.5 mM $MgCl_2$, 1 mM EGTA, 1% Triton X-100, and 10% glycerol), boiled in 1 × sample buffer, and separated by SDS-PAGE. Transfer and immunoblotting were performed as described (20). In other studies, HeLa cells were grown to 50% confluence in 100-mm dishes and transfected with 20 $\mu$g of wild type and mutant Shc DNA constructs using calcium/phosphate precipitation (21). Cells were lysed in 1% Triton X-100 lysis buffer containing protease and phosphatase inhibitors, and lysates were immunoprecipitated for 90 min at 4 °C with the anti-Myc antibody, 9E10 (Oncogene Science), which had previously been bound to protein A-Sepharose beads. The beads were then washed three times with HNTG (20 mM Hepes, pH 7.5, 150 mM NaCl, 10% glycerol, and 0.1% Triton X-100) and boiled in 1 × sample buffer. The bound proteins were separated by SDS-PAGE, transferred to nitrocellulose, and immunoblotted with anti-Myc and anti-phosphotyrosine (19). Blots were visualized using chemiluminescence (Renaissance, DuPont NEN).

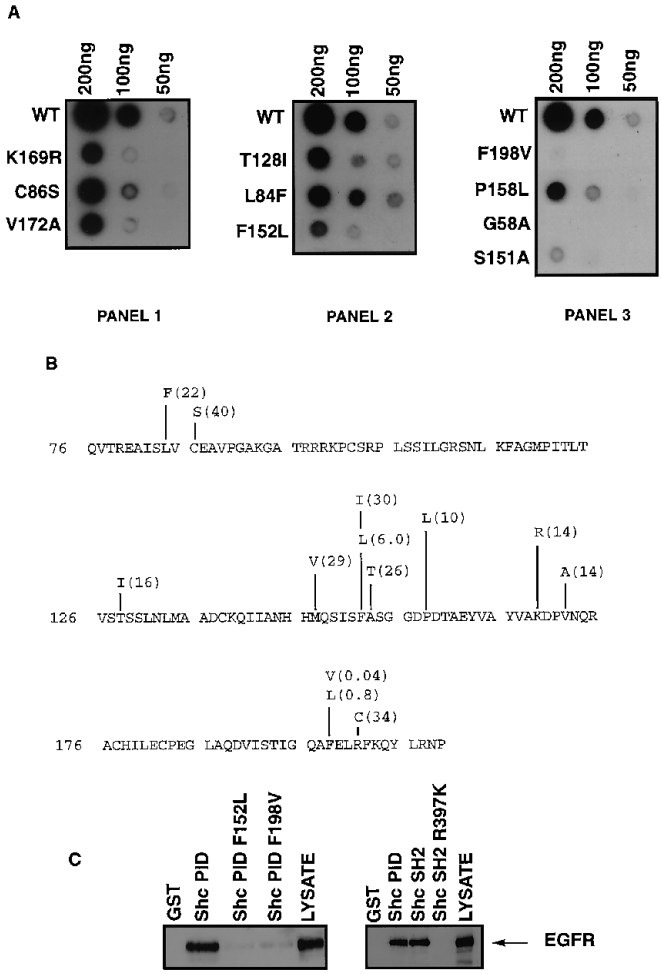


FIG. 2. **Identification and characterization of PI domain mutants.** *A*, binding of $^{32}$P-labeled EGF receptor to PI domain mutants. PI domain mutants identified using the method in Fig. 1 were subcloned into pGEX4T-1 vector. GST fusion proteins were prepared as described under "Experimental Procedures" with 200, 100, and 50 ng spotted on nitrocellulose. The wild type (*WT*) and mutant PI domains were then probed with $^{32}$P-labeled EGF receptor. For quantitation, squares were cut from the filters and Cerenkov counts determined. Duplicate filters spotted with GST fusion protein and probed with anti-Myc were utilized to ensure comparable amounts of protein were spotted for each protein (results not shown). *B*, sites of PI domain mutations. The PI domain point mutations isolated by random mutagenesis are listed above the wild type Shc sequence. The *numbers* in *parentheses* represent the binding of GST fusion proteins containing these mutations to the EGF receptor as a percent of wild type control. The *numbers* represent an average of the binding to 100 and 200 ng of GST fusion protein. *C*, effects of PI and SH2 domain mutagenesis on binding to EGF receptor (*EGFR*). GST fusion proteins (3 $\mu$g of Shc PI domain; 10 $\mu$g of Shc SH2 domain) were immobilized on glutathione beads, and EGF-stimulated HER14 cell lysate was added (1.5 mg of protein). After incubation for 90 min at 4 °C, the beads were washed with radioimmune precipitation buffer and the bound proteins separated by SDS gel. After transferring to nitrocellulose, the bound EGF receptor was detected by immunoblotting with anti-phosphotyrosine antibodies. As a control, lysate (100 $\mu$g) was run directly on the gel (*far right lanes*). It should be noted that although the binding of the wild type SH2 and PI domain (*PID*) appear equivalent, three times as much SH2 domain was used.

## RESULTS AND DISCUSSION

We subcloned the cDNA encoding the amino terminus of murine p52 Shc (Fig. 1*A*) into the pTOPE vector allowing expression in bacteria expressing T7 polymerase. We added a

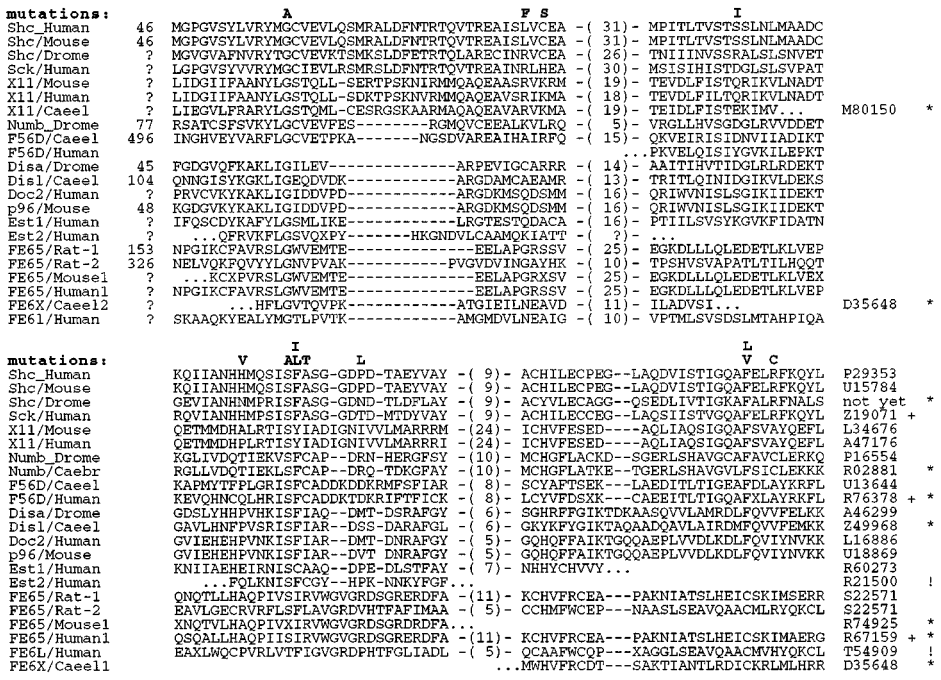


FIG. 3. **Alignment of the conserved blocks of PI domains with mutations generated in this study.** This figure represents an extension of our previously published alignment (16) with data base search and alignments performed as described previously. Above the alignment are the locations of the mutations within the conserved blocks of the PI domain. Except for the mutations of G58A and S151A, the Shc mutations were isolated by screening a library of PI domain mutations (see "Results and Discussion"). !, indicates new identified proteins; *, indicates probable orthologues from other species; +, this entry has been amended with fragments from other expressed sequence tags (*Est*); . . . , only a fragment of this protein is available in the data base.

Myc epitope in-frame at amino acid 209 that contained a unique *Xho*I site at its carboxyl terminus and introduced a *Pst*I site between the Myc epitope and the PI domain. We then created a library of PI domains with mutations running from amino acid 76 to 209 of p52 Shc but excluding the Myc epitope as described under "Experimental Procedures." Bacterial colonies expressing the library of PI domain mutants were lifted onto nitrocellulose and screened for EGF receptor binding and Myc epitope expression. Colonies selected for further study stained positive for Myc (indicating full-length protein expression) but had reduced EGF receptor binding. Such an example is shown in Fig. 1*B*. Plasmids were isolated from these bacterial colonies and subcloned into the pGEX4T-1 vector. This allowed the production of soluble GST fusion proteins that were spotted on nitrocellulose at a known concentration and probed with the radiolabeled EGF receptor. These plasmids were also subjected to DNA sequencing. Several of the mutants that had more than one mutation were not further analyzed. The binding of some of the PI domains with single point mutations and binding less than 50% of control are displayed in Fig. 2*A* while the sites of the mutations are shown in Fig. 2*B*. The binding percentage compared with wild type, displayed in Fig. 2*B*, represents an average of the binding to 100 and 200 ng of GST fusion protein.

Also displayed in Fig. 2*A* is the binding of two mutations that were created based on the previously published PI domain alignment (16). In this alignment two absolutely conserved residues were identified, a glycine at residue 58 and a serine at residue 151 of p52 Shc. The S151A mutation had impaired binding and mapped to a region of Shc where several of the random mutants were located. This mutation resulted in a binding approximately 1.0% of wild type. No binding was detected in the G58A mutation. However, a large fraction of this protein became insoluble when expressed in bacteria, suggesting that this mutation adversely affected the conformation of the protein.

The effects of two of the phenylalanine mutations on PI domain binding are shown using a different assay in Fig. 2*C* (*left panel*). In this assay, the GST fusion proteins were immobilized on glutathione-agarose, and cell lysates containing tyrosine-phosphorylated EGF receptor were added. Both of the phenylalanine mutations, F198V and F152L, markedly reduced the PI domain binding in agreement with the results shown in Fig. 2*A*. Shown for comparison in the *right panel* of Fig. 2*C* is the effect of mutating arginine 397 of the Shc SH2 domain to lysine. This is an arginine within the SH2 domain that is crucial for phosphotyrosine interaction. As expected this mutation completely abolished the ability of the Shc SH2 domain to bind to the EGF receptor.

The alignment presented in Fig. 3 maps the mutations we have generated to the conserved blocks found in PI domains (16). The blocks were confirmed by another round of sequence data base searches in which we were able to retrieve 10 more members of the PI domain family. Among them are the partial sequences of two new proteins, Est2 and FE6L, that have been identified in this work. FE6L (for FE65-like protein) belongs to the FE65 subfamily while Est2 appears to be yet another divergent member of the PI domain family. The other 8 sequences are probably orthologues of existing PI domain containing proteins which, nevertheless, indicate conserved regions in the alignment. The relationship of the PTB domains of IRS-1 and IRS-2 to our PI domain alignment is not clear (6, 22). We have not been able to include them without severely impairing the quality of the PI domain alignment.

Some of the mutations we identified mapped to relatively conserved residues while others did not. The mutations affecting phenylalanine at residue 198 were the most striking because they had the greatest effect on the binding of the Shc PI domain. This phenylalanine is conserved in all PI domains except those of the FE65 subfamily where it is replaced by a cysteine. The fact that mutagenesis of the conserved residues, Phe-198 and Ser-151, severely impairs phosphopeptide binding lends strong support to the validity of PI domain sequence alignment. The mutation of residues analogous to the Phe-198 or Ser-151 in other PI domains could lead to proteins with impaired PI domain function. The finding that one subfamily of PI domains, the FE65-like proteins, has cysteine instead of phenylalanine at the position equivalent to Phe-198 of Shc points to a possible diversification of PI domain binding specificity. Thus it is conceivable that the PI domain represents more of a general protein interaction domain, one subtype of which is involved in phosphotyrosine binding.

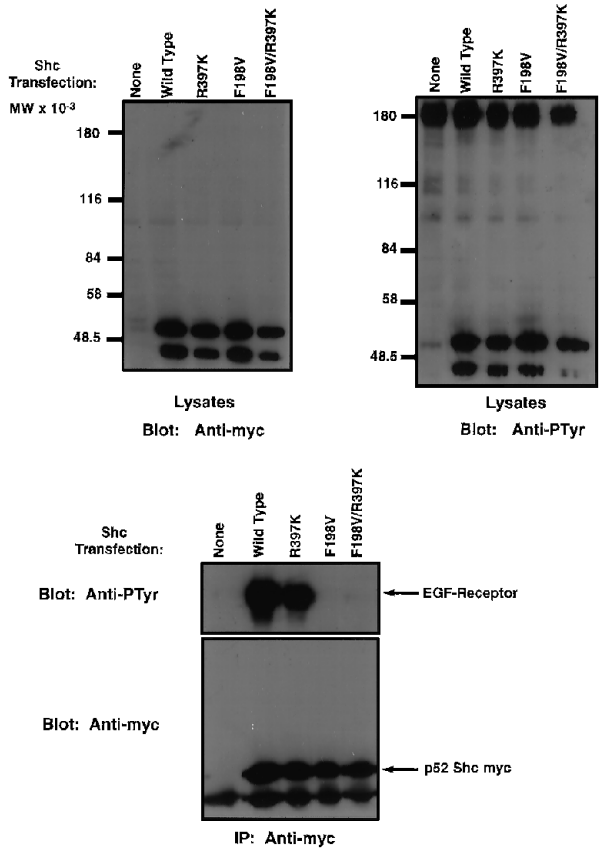


FIG. 4. **Mutagenesis of the PI domain blocks the interaction of Shc with EGF receptor in living cells.** The mutations F198V and R397K either separately or together were reconstituted into full-length murine Shc containing a Myc tag epitope at its carboxyl terminus. Wild type and mutant Shc constructs were then transiently transfected into HeLa cells. Forty-eight hours after transfection the HeLa cells were stimulated with EGF (250 ng/ml) for 5 min. Cells were then lysed in 500 $\mu$l of a 1% Triton-based lysis buffer containing protease and phosphatase inhibitors. The majority of the lysate (400 $\mu$l) was subjected to immunoprecipitation with anti-Myc antibody. After washing with HNTG, the immunoprecipitates (*IP*) were run out on an SDS gel and transferred to nitrocellulose. The high molecular weight portion of the nitrocellulose was then blotted with anti-phosphotyrosine to detect EGF receptor while the lower molecular weights were probed with anti-Myc (*lower panel*). Lysate (20 $\mu$l) was also run directly on an SDS gel, transferred to nitrocellulose, and blotted with anti-Myc or anti-phosphotyrosine (*Anti-PTyr, upper panels*).

The PI domain, not the SH2 domain, is known to be crucial for the interaction of Shc with the nerve growth factor receptor, TrkA (11, 12). However the relative roles of these domains in Shc interactions with EGF receptor are less clear because both the Shc SH2 and PI domains can bind the EGF receptor *in vitro*. We proceeded to reconstitute the PI domain and SH2 domain point mutations we had created into full-length Shc to analyze the role of these domains in the interaction of Shc with tyrosine-phosphorylated EGF receptor in living cells. We performed these studies by transiently transfecting four forms of Myc-tagged Shc into HeLa cells. These forms were wild type Shc, R397K Shc, F198V Shc, and a double mutant, F198V/R397K. The expressions of all these forms of Shc were equivalent as assessed by blotting cell lysates with anti-Myc antibody (Fig. 4, *upper left panel*). By immunoprecipitating with anti-Myc antibody we could analyze the associations of the wild type and mutant forms of Shc with endogenous EGF receptor present in HeLa cells. We found that while mutations in the SH2 domain had only minor effects on binding, mutations in the PI domain severely impaired binding (*lower panel*). Despite the ability of PI domain mutations to interfere with Shc binding, we saw no significant effect of these mutations on EGF-induced Shc tyrosine phosphorylation (*upper right panel*). This is not surprising as there have been several previous studies demonstrating that Shc does not need to bind to EGF receptor to become tyrosine-phosphorylated (23–26). It may be that a low affinity interaction between Shc and the EGF receptor is all that is required for Shc tyrosine phosphorylation or that a separate EGF-responsive kinase can also phosphorylate Shc. Regardless, these results confirmed that the F198V mutation abrogates PI domain function in living cells and points to the important role PI domains play in Shc interactions with growth factor receptors. The isolation of these PI domain mutation will be an important tool to understand the role of this domain in Shc signaling as well as the role of the PI domain in other proteins.

*Note Added in Proof*—While this manuscript was in press, the structure of the PI/PTB domain bound to a phosphopeptide was published (Zhou, M-M., Ravichandran, K. S., Olejniczak, E. T., Petros, A. M., Meadows, R. P., Sattler, M., Harlan, J. E., Wade, W. S., Burakoff, S. J., and Fesik, S. W. (1995) *Nature* **378,** 584–592). This structure indicates that Phe-198 is a crucial residue making direct contact with the phosphopeptide at residues −3 and −5 to the phosphotyrosine.

## REFERENCES


1. Cohen, G. B., Ren, R., and Baltimore, D. (1995) *Cell* **80,** 237–248
2. Pawson, T. (1995) *Nature* **373,** 573–580
3. Schlessinger, J. (1994) *Curr. Opin. Genet. & Dev.* **4,** 25–30
4. Blaikie, P., Immanuel, D., Wu, J., Li, N., Yajnik, V., and Margolis, B. (1994) *J. Biol. Chem.* **269,** 32031–32034
5. Kavanaugh, W. M., and Williams, L. T. (1994) *Science* **266,** 1862–1865
6. Gustafson, T. A., He, W., Craparo, A., Schaub, C. D., and O'Neill, T. J. (1995) *Mol. Cell. Biol.* **15,** 2500–2508
7. Pelicci, G., Lanfrancone, L., Grignani, F., McGlade, J., Cavallo, F., Forni, G., Nicoletti, I., Pawson, T., and Pelicci, P. G. (1992) *Cell* **70,** 93–104
8. Cutler, R. L., Liu, L., Damen, J. E., and Krystal, G. (1993) *J. Biol. Chem.* **268,** 21463–21465
9. Pronk, G. J., McGlade, J., Pelicci, G., Pawson, T., and Bos, J. L. (1993) *J. Biol. Chem.* **268,** 5748–5753
10. Margolis, B., Skolnik, E. Y., and Schlessinger, J. (1995) *Methods Enzymol.* **255,** 360–369
11. Van der Geer, P., Wiley, S., Lai, V. K-M., Olivier, J. P., Gish, G. D., Stephens, R., Kaplan, D., Shoelson, S., and Pawson, T. (1995) *Curr. Biol.* **5,** 404–412
12. Dikic, I., Batzer, A. G., Blaikie, P., Obermeier, A., Ullrich, A., Schlessinger, J., and Margolis, B. (1995) *J. Biol. Chem.* **270,** 15125–15129
13. Kavanaugh, W. M., Turck, C. W., and Williams, L. T. (1995) *Science* **268,** 1177–1179
14. Songyang, Z., Margolis, B., Chaudhuri, M., Shoelson, S. E., and Cantley, L. C. (1995) *J. Biol. Chem.* **270,** 14863–14866
15. Trub, T., Choi, W. E., Wolf, G., Ottinger, E., Chen, Y., Weiss, M., and Shoelson, S. E. (1995) *J. Biol. Chem.* **270,** 18205–18208
16. Bork, P., and Margolis, B. (1995) *Cell* **80,** 693–694
17. Cadwell, R. C., and Joyce, G. F. (1992) *PCR Methods Appl.* **2,** 28–33
18. Evan, G. I., Lewis, G. K., Ramsay, G., and Bishop, J. M. (1985) *Mol. Cell. Biol.* **5,** 3610–3616
19. Batzer, A. G., Blaikie, P., Nelson, K., Schlessinger, J., and Margolis, B. (1995) *Mol. Cell. Biol.* **15,** 4403–4409
20. Margolis, B., Li, N., Koch, A., Mohammadi, M., Hurwitz, D., Zilberstein, A., Ullrich, A., Pawson, T., and Schlessinger, J. (1990) *EMBO J.* **9,** 4375–4380
21. Chen, C., and Okayama, H. (1987) *Mol. Cell. Biol.* **7,** 2745–2752
22. Sun, X. J., Wang, L-M., Zhang, Y., Yenush, L., Myers, M. G., Glasheen, E., Lane, W. S., Pierce, J. H., and White, M. F. (1995) *Nature* **377,** 173–177
23. Li, N., Schlessinger, J., and Margolis, B. (1994) *Oncogene* **9,** 3457–3465
24. Gotoh, N., Tojo, A., Muroya, K., Hashimoto, Y., Hattori, S., Nakamura, S., Takenawa, T., Yazaki, Y., and Shibuya, M. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 167–171
25. Soler, C., Alvarez, C. V., Beguinot, L., and Carpenter, G. (1994) *Oncogene* **9,** 2207–2215
26. Meyer, S., Labudda, K., McGlade, J., and Hayman, M. J. (1994) *Mol. Cell. Biol.* **14,** 3253–3262