Systematic Mapping of Potential Binding Sites for Shc and Grb2
SH2 Domains on Insulin Receptor Substrate-1 and the Receptors
for Insulin, Epidermal Growth Factor, Platelet-derived Growth
Factor, and Fibroblast Growth Factor*

(Received for publication, September 27, 1995, and in revised form, November 27, 1995)

Colin W. Ward‡§, Keith H Gough‡, Melisa Rashke‡, Soo San Wan‡, Gordon Tribbick†, and J Ian-xin Wang‡
From ‡CSIRO, Division of Biomolecular Engineering, Parkville, Victoria 3052, Australia and §Chiron Mimotopes Pty. Ltd., Clayton, Victoria 3168, Australia

Multipin peptide synthesis has been employed to produce biotinylated 11-mer phosphopeptides that account for every tyrosine residue in insulin receptor substrate-1 (IRS-1) and the cytoplasmic domains of the insulin, epidermal growth factor, platelet-derived growth factor- and basic fibroblast growth factor receptors. These phosphopeptides have been screened for their capacity to bind to the SH2 domains of Shc and Grb in a solution phase enzyme-linked immunosorbent assay. The data revealed new potential Grb2 binding sites at Tyr-1114 (epidermal growth factor receptor (EGFR) C-tail); Tyr-743 (platelet-derived growth factor receptor (PDGFR) insert region), Tyr-1110 from the E-helix of the catalytic domain of insulin receptor (IR), and Tyr-47, Tyr-939, and Tyr-727 in IRS-1. None of the phosphopeptides from the juxtamembrane or C-tail regions of IR bound Grb2 significantly, and only one phosphopeptide from the basic fibroblast growth factor receptor (Tyr-556) bound Grb2 but with medium strength. Tyr-1068 and -1086 from the C-tail of EGFR, Tyr-684 from the kinase insert region of PDGFR, and Tyr-895 from IRS-1 were confirmed as major binding sites for the Grb2 SH2 domain. With regard to Shc binding, the data revealed new potential binding sites at Tyr-703 and Tyr-789 from the catalytic domain of EGFR and at Tyr-557 in the juxtamembrane region of PDGFR. It also identified new potential Shc binding sites at Tyr-764, in the C-tail of basic fibroblast growth factor receptor, and Tyr-960, in the juxtamembrane of IR, a residue previously known to be required for Shc phosphorylation in response to insulin. The study confirmed the previous identification of Tyr-992 and Tyr-1173 in the C-tail of EGFR and several phosphopeptides from the PDGFR as medium strength binding sites for the SH2 domain of Shc. None of the 34 phosphopeptides from IRS-1 bound Shc strongly, although Tyr-690 showed medium strength binding. The specificity characteristics of the SH2 domains of Grb2 and Shc are described. This systematic peptide mapping strategy provides a way of rapidly scanning candidate proteins for potential SH2 binding sites as a first step to establishing their involvement in kinase-mediated signaling pathways.

SH2 domains are one of the conserved protein modules that regulate signal transduction pathways involving phospholipid metabolism, protein phosphorylation and dephosphorylation, activation of Ras-like GTPases, gene expression, protein trafficking, and cytoskeletal architecture (1). SH2 domains promote protein-protein interactions by binding short, specific, phosphotyrosine peptide sequences on activated receptors and cytoplasmic proteins (1, 2). The three-dimensional structure of several SH2 domains has now been determined (3–7), and the nature of the phosphopeptide binding site has been described.

Individual SH2 domain binding sites have been identified and their specificity requirements determined by (i) isolation of specific phosphopeptides after receptor activation, (ii) site-specific mutagenesis, (iii) synthetic peptide binding studies (see Refs. 2 and 8), or (iv) analysis of degenerate phosphopeptide libraries (9, 10). It is generally accepted that the SH2 binding sites of receptor tyrosine kinases are located outside the kinase domain proper in the juxtamembrane region, the kinase insert region, or the C-terminal tail (1). However, there are some data (11–13) that suggest that Tyr(P) residues within the catalytic domain proper can act as binding sites for SH2 domain or other signaling proteins. The phosphatidylinositol 3’-kinase binding site in the tyrosine kinase receptor Trk-A is Tyr-751 (11). This Tyr residue is located in the H-helix (see Ref. 14), 24 residues upstream of the C-terminal residue of the catalytic domain as defined by the alignments of Hanks (15). Similarly, the highly conserved activation loop residue Tyr-809 of the CSF-1 receptor (see Ref. 14) has been shown to be required for induction and sustained transcription of c-myc and the control of ligand-dependent cell growth without significantly affecting receptor tyrosine kinase activity or immediate early gene expression (12, 13).

In this paper we describe a systematic approach to identify potential binding sites for SH2 domains on cytoplasmic proteins or transmembrane receptors that includes an analysis of every tyrosine-containing peptide. This approach is rapid and sensitive and complements the studies on direct identification of binding sites by mutagenesis and competition binding or specificity determination with degenerate peptide libraries (9, 10). It confirms many of the sites established by alternative methods and reveals new potential binding sites for Grb2 at Tyr-1110 in IR, Tyr-1114 in EGFR, Tyr-743 in PDGFR, and Tyr-47, Tyr-727, and Tyr-939 in IRS-1 and for Shc at Tyr-960 in IR, Tyr-703 and Tyr-789 in EGFR, and Tyr-557 in PDGFR.

* 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.

§ To whom correspondence should be addressed: CSIRO, Division of Biomolecular Engineering, 343 Royal Parade, Parkville, Victoria, 3052, Australia. Tel.: 0011-61-3-342-2400; Fax: 0011-61-3-347-5481; E-mail: Colin.Ward@mel.dbe.csiro.au.

1 The abbreviations used are: IR, insulin receptor; IRS, insulin receptor substrate; EGFR, epidermal growth factor receptor; PDGFR, platelet-derived growth factor receptor; GST, glutathione S-transferase; FGFR, fibroblast growth factor receptor; bFGFR, basic FGFR; PID, phosphotyrosine interaction domain.
MATERIALS AND METHODS

GST-SH2 Fusion Proteins—cDNAs encoding the whole of Grb2 (residues 1–215) or the SH2 domain of Shc (residues 366–473) cloned into pGEX vectors (16, 17) were kind gifts from Dr. David Bowtell (Howard Florey Institute of Experimental Physiology and Medicine, Parkville, Victoria, Australia) and Dr. Tony Pawson (Samuel Lunenfeld Research Institute, Mount Sinai Hospital and University of Toronto, Ontario, Canada), respectively. The GST fusion proteins were produced in Escherichia coli and purified with glutathione-agarose beads (18). Fusion protein concentrations were estimated by Coomassie dye binding (19) using the kit obtained from Bio-Rad.

Synthesis of Peptides—The 256 peptides synthesized corresponded to the sequences centered around all of the tyrosine residues in rat IRS-1 (20) and the cytoplasmic domains of the human insulin receptor (21), human EGFR (22), mouse PDGFR (23), and chicken basic fibroblast growth factor receptors (24). The peptides were synthesized as 11-mers joined to a biotin cap by the tetrapeptide linker SGSG in the format biotin-SGSGXXXXXYXXXXX where Y is either Tyr or Tyr(P). The biotin-SGSG- cap enabled attachment to streptavidin-coated plates in the enzyme-linked immunosorbent assay. Peptides containing Cys residues were also synthesized, with Ala replacing the Cys residue to check for nonspecific binding due to the free thiol side chain. In no case did this substitution have a significant effect on binding. Throughout the text the peptides are referred to by the protein source and Tyr residue number; e.g., EGFR-1068. Peptides were synthesized using the multipin peptide synthesis approach on polyethylenesupports derivitized with an acid-labile handle (25). The synthesis was carried out on detachable crown-shaped pins headed with hydroxyethylmethacrylate and functionalized with the trifluoroacetic acid-labile 4-hydroxymethylphenoxyacetic acid handle, which yields peptide carboxylates on acid cleavage (25).

Peptide Binding Assay—The GST-SH2 fusion proteins for Shc and Grb2 (10 μg/ml final concentration) were incubated in solution phase with 0 μM, 0.2 μM, 1 μM, 5 μM, and 25 μM (final concentrations) of each Tyr and Tyr(P) peptide. Aliquots (125 μl) of each fusion protein (20 μg/ml) were mixed with 125 μl of biotinylated peptide (0 μM, 0.4 μM, 2 μM, 10 μM, and 50 μM) in individual wells in Titertek™ plates pre-blocked with 1% bovine serum albumin, 10 mM phosphate-buffered saline. The buffer used for the dilution of SH2 fusion proteins and biotinylated peptides was 10 mM phosphate-buffered saline, 0.1% (v/v) Tween 20. The mixtures were incubated for 1 h at room temperature (22°C) with shaking. After 1 h, duplicate 100-μl aliquots were transferred to Streptavidin-coated NUNC Maxisorp™ plates. The NUNC plates were incubated for 1 h at room temperature with shaking and washed with distilled water. The amount of GST fusion protein bound to the trapped biotinylated peptides was determined by enzyme-linked immunosorbent assay. Aliquots (100 μl) of rabbit anti-glutathione S-transferase antibody (AMRAD Corp. Ltd.) at 1:5000 dilution in 1% bovine serum albumin, 10 mM phosphate-buffered saline were added to each well, and the plates were incubated for 1 h at room temperature with shaking. The plates were washed with distilled water, and 100 μl of goat anti-rabbit IgG (H + L) horseradish peroxidase conjugate (Kirkegaard & Perry Laboratories, Inc.) at 1:2000 dilution in 0.1% FIG. 1. Potential Binding Sites for Shc and Grb2 SH2 Domains

Fig. 1. Aligned sequences of human IR, human EGFR, mouse PDGFR, and chicken bFGFR cytoplasmic domains showing the relative location of the Tyr residues and their positions in the protein structure (see Ref. 14). The sequence of human TrkA is also shown to indicate the relative position of the phosphatidylinositol 3'-kinase binding site at Tyr-751, which is located in the catalytic domain proper (11).
bovine serum albumin, 10 mM phosphate-buffered saline was added to each well. The plates were incubated for 1 h at room temperature with shaking. After a final wash with distilled water, 100 μl of substrate, 2,2'-azino-di-[3-ethyl-benzthiazolinesulfonate] was added to each well. The optical density of the plates was read after 45 min at 405/492 nm with a Titertek Multiscan™ plate reader. The values are expressed as net optical densities and represent the difference between the value obtained for the phosphopeptide minus the value obtained for the corresponding nonphosphorylated control peptide. Net optical densities between 0.2 and 0.6 will be referred to as weak binding, those between 0.6 and 1.0 as medium binding, and those between 1.0 and 1.4 as strong binding.

RESULTS

Phosphopeptide Design and Binding Assay—Phosphopeptides accounting for all the Tyr-containing sequences in IR, EGFR, PDGFR, and bFGFR (Fig. 1) and IRS-1 were synthesized as 11-mers with the Tyr or Tyr(P) residue in the sixth position. Preliminary studies using a solid phase, direct binding protocol on peptides that remained attached to the acrylic acid-coated polyethylene rods after synthesis (26) gave strong binding for Tyr(P) peptides compared with the nonphosphorylated or Phe-substituted controls. However, the solid phase assay did not give the required discrimination between specific Tyr(P) peptides, presumably because of the high concentration of peptide on the pin surface. Consequently, the final system chosen was a solution phase, direct binding assay with biotinylated peptides at concentrations of 0, 0.2, 1, 5, and 25 μM. Complex formation between GST-SH2 fusion proteins and biotinylated peptides was measured by enzyme-linked immunosorbent assay using anti-GST antibodies after transfer to streptavidin-coated plates. This assay was quick, sensitive, and discriminating when used with a set of test peptides including known positive controls (data not shown).

Saturation binding curves for the Grb2-fusion protein with 100 μM of phosphopeptide EGFR-1068 (27) and the Shc-SH2-fusion protein with 100 μM PDGFR-708 (28) were carried out and showed linear binding within the range 1–20 μg/ml (data not shown). A concentration of 10 μg/ml of fusion protein was chosen for all screening assays. Ca2+ ions (2 mM), which have been shown to stabilize the conformation of the p85N-SH2 domain and enhance its binding to activated α-PDGFR (29), were found to be inhibitory to Grb2-SH2 binding to EGFR-1068 and Shc-SH2 binding to PDGFR-708 and were not included in the binding screens with Grb2- and Shc-SH2-fusion proteins.

Grb2-SH2 Binding Sites—The results obtained from the phosphopeptide screens with the Grb2-fusion protein are shown in Figs. 2–6. The major sites for Grb2 binding in human EGFR (at 1068, 1086, and 1114) were all located in the C-tail region of the receptor (Fig. 2), while the two major sites of Grb2 binding in the mouse PDGFR (at 684 and 746) were both in the kinase insert region. Chicken bFGFR contained only one potential binding site for Grb2 at Tyr-556, and this showed only weak interaction (Fig. 4). The dominant site in human IR was peptide IR-1110 from the kinase domain (Fig. 5). Phosphopeptides corresponding to sites in the juxtamembrane or C-tail regions were either weak (IR-953, IR-972) or negligible (IR-960, IR-1316, IR-1322; Fig. 5). The strong binding sites in rat IRS-1 were peptides IRS1-47, 895, 893, and to a lesser extent -727 (Fig. 6).

Shc-SH2 Binding Sites—The results obtained from the phosphopeptide screens with the Shc-SH2-fusion protein are shown in Figs. 2–6. The major sites for Shc binding in human EGFR and PDGFR were at Tyr-891 and Tyr-977, respectively. The dominant site in human IR was peptide IR-1110 from the kinase domain (Fig. 5). Phosphopeptides corresponding to sites in the juxtamembrane or C-tail regions were either weak (IR-953, IR-972) or negligible (IR-960, IR-1316, IR-1322; Fig. 5). The strong binding sites in rat IRS-1 were peptides IRS1-47, -895, -893, and to a lesser extent -727 (Fig. 6).
phopeptide screens with the Shc-SH2 fusion protein are also shown in Figs. 2–6. One peptide, EGFR-703 at the start of the kinase domain, bound Shc-SH2 strongly, while peptide EGFR-789 from the kinase domain showed medium binding (Fig. 2). Of the PDGFR peptides, PDGFR-557 bound strongly, while PDGFR-654, -708, and -848 exhibited medium binding to the Shc-SH2 domain (Fig. 3). The Shc-SH2 domain showed medium binding to the chicken bFGFR C-tail phosphopeptide bFGFR-764 and weak binding to bFGFR-556 and -728 in the kinase domain (Fig. 4). Only two phosphopeptides from human IR bound the Shc-SH2-fusion protein. The best binder (medium strength) was IR-960 from the juxtamembrane region, while IR-1198 from the kinase domain showed weak binding (Fig. 5). As shown in Fig. 6, none of the 34 phosphopeptides from IRS-1 bound the Shc-SH2-fusion protein strongly, and only one, IRS1–690, bound with medium strength.

**DISCUSSION**

Songyang and co-workers (9, 10) have used degenerate phosphopeptide or peptide libraries to systematically search for optimal sequences that serve as binding sites for SH2 domains or substrates for protein-tyrosine kinases (30). In this paper we describe a complementary approach in which all tyrosine-containing sequences in signaling proteins are synthesized and screened for their capacity to be bound by specific SH2 domains. It is obvious that the method could be extended to include a systematic analysis of the potential for these sites to be phosphorylated by any specific tyrosine kinase.

To date most known tyrosine autophosphorylation sites are reported to be located within noncatalytic regions of receptors (2). However, the demonstration that some phosphorylation/dephosphorylation sites occur within the catalytic domain of tyrosine kinase proteins and that at least one SH2-containing protein has been shown to bind to a Tyr(P) sequence within the catalytic domain of TrkA (11) highlights the need to examine all Tyr-containing sequences for their potential to bind SH2 (1, 2) or PID domains (31, 32) if phosphorylated. Such studies do not establish that these sites are used in vivo but merely indicate their potential to be phosphorylated and participate in SH2-mediated signaling if they were exposed and accessible. Additional data are required to establish their in vivo significance and must take account of the fact that the stoichiometry of phosphorylation at specific sites at any given time in living cells is not known (33). In addition the net signaling outcomes will be influenced by direct competition between different SH2- or PID-containing proteins for the same or closely overlapping phosphorylation sites and by the effects of secondary phosphorylation of Ser or Thr residues on these phosphotyrosine-mediated interactions.

In this paper we have synthesized and screened all the tyrosine-containing sequences in IRS-1 and the cytoplasmic domains of IR, EGFR, PDGFR, and bFGFR for their capacity to bind to the SH2 domains of the signaling proteins Grb2 and Shc. EGFR—Two of the major sites identified for Grb2 binding on EGFR (Tyr-1068, Tyr-1086) have been reported (27, 33), but the additional major potential binding site at Tyr-1114 in the C-tail has not been described previously. Tyr-1114 has not been identified as a major phosphorylation site in EGFR by biochemical analysis (34). This situation is similar to that found for

![Grb2 and Shc-SH2 domain binding to phosphopeptides from chicken basic FGFR](image1)

**Fig. 4.** Grb2- and Shc-SH2 domain binding to phosphopeptides from chicken basic FGFR. Binding was carried out as described under "Materials and Methods." The net optical density values obtained at 1.0 μM are shown for each Tyr residue in chicken bFGFR. Peptides PDGFR-461, -561, -570, -581, -583, -603, -611, -651, -652, -675, -699, -728, -764, and -774 showed negligible binding by the SH2 domain of Grb2. Peptides bFGFR-461, -561, -570, -581, -583, -603, -611, -651, -652, -675, -699, and -774 showed negligible binding to the Shc-SH2 domain. A schematic representation of the bFGFR is shown to indicate the approximate location of each of the positively binding peptides in the sequence.

![Grb2 and Shc-SH2 domain binding to phosphopeptides from human IR](image2)

**Fig. 5.** Grb2- and Shc-SH2 domain binding to phosphopeptides from human IR. Binding was carried out as described under "Materials and Methods." The net optical density values obtained at 1.0 μM are shown for each Tyr residue in IR. Peptides IR-953, -972, -999, -1075, -1150, -1151, 1198, -1215, -1316, and -1322 showed negligible binding by the SH2 domain of Grb2. Peptides IR-953, -972, -999, -1075, -1110, -1146, -1150, -1151, -1215, -1316, and -1322 showed negligible binding to the Shc-SH2 domain. A schematic representation of the IR is shown to indicate the approximate location of each of the positively binding peptides in the sequence.
Potential Binding Sites for Shc and Grb2 SH2 Domains

Tyr-954 of EGFR, which was identified as the probable binding site for SH-PTP2, although its phosphorylation status has not been established (34). Although EGFR can bind both Shc and Grb2, the predominant linkage observed in rat fibroblasts was between EGFR and Shc (35). At first this seems surprising given the higher binding values obtained with Grb2 at its major sites Tyr-1068, -1086, and -1114 compared with Shc at Tyr-1173 (see Fig. 2). It may reflect the effects of competition in vivo between Grb2 and other SH2-containing proteins such as phospholipase C\textsubscript{y}, which also binds the Tyr-1068 site (36), or the contribution of other binding sites to Shc activation.

The major potential binding site for Shc on EGFR (Tyr-703) and the two medium binding peptides (Tyr-740, Tyr-789) are located in the catalytic domain, while the weak to medium binding peptides (Tyr-876, -954, -974, and -1173) are all in the C-tail region (Figs. 1 and 2). The physiological significance of these multiple Shc binding sites within the catalytic domain and the C-tail is uncertain. Competition binding and dephosphorylation protection experiments showed that Tyr-1173 was a major and Tyr-992 a minor binding site for the Shc2 domain of Shc on EGFR (33). However, other sites have been shown to be able to compensate for the loss of these C-tail sites for Shc, and stable association of Shc with EGFR may not be necessary for in vivo function (37). Phosphorylation of Shc, complex formation of Shc-Grb2 and activation of mitogen-activated protein kinase were found to be normal in an EGFR truncated at residue 1011 and mutated to Phe at Tyr-992, even though the association between the receptor and Shc was hardly detectable (37). Our findings suggest that any one of six additional potential binding sites could be responsible for Shc binding and activation in such an EGFR construct (Fig. 2).

PDGFR—In this report we find two major sites and four intermediate sites for Grb2 binding in mouse PDGFR (Fig. 3). The two major sites, Tyr-684 and Tyr-743, are both in the kinase insert domain (Figs. 1 and 3). The 684 site (Tyr-716 in human PDGFR) has been identified previously as the only major binding site for Grb2 from a study of Grb2 binding to a panel of 10 mutant human PDGFRs and nine synthetic phosphopeptides (38). In addition, mutation of Tyr-699 (the residue homologous to Tyr-684 in mouse PDGFR) removed the capacity of the CSF-1 receptor to bind Grb2 and compromised its mitogenic response to ligand (39). These data suggest that Tyr-743 is either not phosphorylated or not available for Grb2 binding in vivo. Neither the CSF-1 receptor nor c-Kit contain a Tyr residue equivalent to Tyr-743 in mouse PDGFR (40). The autophosphorylation sites identified so far in the PDGFR (41) are located in the juxtamembrane region (Tyr-547 and Tyr-549), the kinase insert domain (Tyr-684, -708, -719, and -739), the catalytic domain (Tyr-825), and the C-tail region (Tyr-977 and -989).

With Shc binding, we show that Tyr-557 at the end of the juxtamembrane region binds the Shc-SH2 domain strongly, while 12 other sites are medium to weak binders (Fig. 3). Yokote et al. (28), using synthetic peptide analyses, showed that multiple autophosphorylation sites (equivalent to 547, 708, 719, and 739) were able to mediate binding to the Shc-SH2 domain but suggested that additional unidentified autophosphorylation sites in PDGFR may also bind Shc since mutant receptors, in which each of these Shc binding site tyrosines were changed to phenylalanine, showed normal (Tyr-708, -719, -739 mutants) or partially (40%) reduced (Tyr-547 mutant) Shc binding and were still capable of phosphorylating Shc at close to normal levels. Our data suggest Tyr-557 could be the unidentified mediator of this Shc binding. It is interesting to note that Tyr-571 in the CSF-1 receptor (homologous to Tyr-557 in mouse PDGFR) occurs in a sequence where the amino acid residues at the +1, +2, and +3 positions are different from those found in PDGFR (40). This may explain the absence of demonstrable Shc binding to the CSF-1 receptor in Rat2 cells stimulated with CSF-1, even though Shc was phosphorylated and shown to be associated with Grb2 (39).

Basic Fibroblast Growth Factor Receptor—In this report on chicken bFGFR, we find that Tyr-556 at the start of the B5 strand in the catalytic domain is the strongest binding peptide for Grb2 (Fig. 4). Its binding is weak to medium (–42%) of that seen with the best binders from EGFR, PDGFR, and IRS-1. The other 14 bFGFR phosphopeptides tested, including Tyr-461, -653, and -766, the known autophosphorylation sites (8, 42, 43), did not bind Grb2 significantly. Tyr-653 is the major autophosphorylation site and is in the activation loop of the catalytic domain (Fig. 1). Tyr-461 in the juxtamembrane region has been implicated as the binding site for Grb2 in cells overexpressing FGR-1, although Grb2 association with FGR-1 was not observed in cells expressing physiological levels of FGR-1 (see Ref. 42).

Fibroblast growth factor-induced mitogenesis is mediated by activation of the mitogen-activated protein kinase cascade (44), presumably through the activation of Shc (see Ref. 42), although the site of Shc binding has not been reported. In this report the C-tail Tyr-764, (equivalent to Tyr-766, the phospholipase C\textsubscript{y} site in human FGR-1), shows medium binding to the SH2 domain of Shc. Mutation of Tyr-766 in FGR-1 had no effect on mitogenesis but inhibited phosphatidylinositol hydrolysis, Ca\textsuperscript{2+} influx (45, 46), and internalization (47). The lack of effect on mitogenesis suggests that Shc activation with the FGR-1 may be mediated by SH2 binding to the minor sites.
In contrast to the activated EGFR, activated IR is thought not to bind Grb2 directly (50, 51) but rather to induce the association of Grb2 with phosphorylated IRS-1 with Grb2 has been reported (52, 53). One of these sites (Tyr-939) is known to be phosphorylated in vivo (31, 32). No sites exist in chick bFGFR (24) that match the specificity requirements of the Shc PID (1, 48, 49).

In contrast to IRS-1—In contrast to the activated EGFR, activated IR is thought not to bind Grb2 directly (50, 51) but rather to induce the association of Grb2 with phosphorylated IRS-1 and Shc (52). As shown in Fig. 5, none of the 34-phosphotyrosine peptides from the juxtamembrane or C-tail regions of IR showed strong binding to Grb2, the only major potential binding site being Tyr-1110 near the end of the E-helix in the kinase domain. The phosphorylation status of Tyr-1110 and its accessibility to SH2-containing proteins in vivo is not known (20).

In contrast to IR, IRS-1 showed strong potential to bind Grb2 at several sites: Tyr-47, Tyr-895, Tyr-939, and to a lesser extent Tyr-727 (Fig. 6). The physical association of phosphorylated IRS-1 with Grb2 has been reported (52, 53). One of these sites (Tyr-895) is known to be phosphorylated in vivo and to bind to a fusion protein containing the SH2 domain of Grb2 (20). Two other phosphopeptide fractions were shown to be bound by the Grb2-SH2 fusion protein but were not identified and did not include the Tyr-939 phosphopeptide (20). The reason for this discrepancy between their findings with the Tyr-939 phosphopeptide and ours is not known. The additional Grb2 binding sites may come from the region 530–768 in IRS-1, since a GST fusion protein containing these IRS-1 residues, but not amino acids 3–67, bound Grb2 when phosphorylated by the IR (53). The region of IRS-1 from 530 to 768 contains 13 tyrosines including the strong binder Tyr-727 (Fig. 6).

Shc can be phosphorylated by purified IR and IGF-1R in vitro, suggesting it is a direct substrate (54). In addition, in 32-D myeloid progenitor cells (which lack IRS-1 and have low levels of IR), expression of IR alone was sufficient for Shc phosphorylation and mitogen-activated protein kinase activation (55). In this report we have identified a potential Shc-SH2 binding site at Tyr-960 in the juxtamembrane region of the IR. In addition, immunoprecipitation and Western blotting has shown that the Shc-SH2 domain fusion protein can bind to autophosphorylated IR in vivo.2 The direct binding of Shc-SH2 domain to IR has not been reported previously (52, 56). Tyr-960 is known to be an autophosphorylation site in IR, although the detection of this phosphorylation was difficult (57, 58). It has also been shown recently to be the site at which the PID of Shc interacts (59). Removal of Tyr-960 by mutagenesis severely impaired insulin-dependent phosphorylation of Shc (60). However, it also impaired the phosphorylation of IRS-1 (57, 61), consistent with the fact that both Shc and IRS-1 compete for this residue in their association with the phosphorylated IR (59). Residues in the C-tail of the IR are also required for Shc phosphorylation (60), but neither of the C-tail phosphopeptides exhibited significant binding to the Shc-SH2 domain (Fig. 6).

As shown in Fig. 6, none of the 34-phosphotyrosine peptides from IRS-1 bound Shc strongly, although Tyr-690 showed moderate binding and Tyr-107 weak to moderate binding. The significance of this potential binding is not known. Neither of these sites has been positively identified as a phosphorylation target for the IR kinase, although additional phosphorylation sites in IRS-1 do exist (20). While some groups have failed to detect Shc binding to IRS-1 (52, 56), Shc has been detected in IRS-1 immunoprecipitates by others (54).

Specificity Requirements—The top 20 phosphopeptides that bound to Grb2 are listed in Table I. The dominant specificity determinant is Asn at the +2 position, as reported previously (10). It is present at this position in the 10 top binders, and there are no other phosphopeptides with Asn at the +2 position in the set of 267 tested. However, as shown in Table I, other

| Peptide | OD | Peptide | OD |
|---------|----|---------|----|
| EGFR-1068 | 1480 | PDGFR-557 | 1208 |
| EGFR-1086 | 1474 | EGFR-703 | 1204 |
| EGFR-1114 | 1463 | EGFR-789 | 917 |
| IRS1-47 | 1459 | IRS-960 | 907 |
| IR-1110 | 1350 | PDGFR-708 | 767 |
| IRS1-895 | 1347 | PDGFR-848 | 710 |
| IRS1-939 | 1297 | PDGFR-654 | 643 |
| PR-684 | 1131 | IRS-107 | 614 |
| PR-743 | 1126 | EGFR-867 | 574 |
| IRS1-727 | 1014 | PDGFR-832 | 573 |
| IRS1-1148 | 958 | IRS-960 | 573 |
| EGFR-1101 | 860 | EGFR-974 | 568 |
| IRS1-46 | 709 | IRS1-1173 | 548 |
| PR-646 | 648 | IRS1–690 | 536 |
| PR-739 | 663 | PDGFR-646 | 533 |
| PR-549 | 593 | IRS-727 | 536 |
| FR-556 | 578 | PDGFR-739 | 525 |
| PR-547 | 511 | IRS-960 | 525 |
| IR-953 | 501 | IRS-1222 | 542 |
| IRS1-760 | 501 | IRS1–479 | 542 |
| IRS1–47 | 598 | IRS1–502 | 542 |
| IRS1–114 | 501 | IRS1–536 | 542 |
| IRS1–46 | 501 | IRS1–536 | 542 |
| IRS1–895 | 501 | IRS1–536 | 542 |
| IRS1–939 | 501 | IRS1–536 | 542 |
| IRS1–114 | 501 | IRS1–536 | 542 |

2 S. L. Macaulay, personal communication.
residues can occur at this position and still show significant binding. The other dominant feature seen in the set of peptides studied is the high frequency Pro at the -2 position, particularly in combination with one of the small amino acids Gly, Asn, or Ser at the -3 position.

The top 29 phosphopeptides bound by the Shc SH2 domain are also shown in Table I. The results reveal a strong tendency for hydrophobic residues, particularly Leu and Met, at the +3 position and for Trp and Ser at the +4 position. There is also a high frequency of small hydrophilic residues at the -3 and -2 positions. Almost any residue can occur at the -1 and +2 positions. While there is a higher occurrence of hydrophobic residues at the +1 position in the peptides examined, almost any residue can be accommodated here, and the two strongest binding peptides have Asp or Lys (Table I).

These results are in good agreement with the findings obtained with degenerate peptide libraries (9, 10). The method complements their approach and allows all tyrosine-containing sequences from any tyrosine kinase or tyrosine kinase substrate to be rapidly screened for their potential to bind specific SH2 domains. Replacement net analyses using the multipin synthesis methodology employed here (26) could also be used to map the effects, on SH2 domain specificity, of amino acid substitutions at each position. This information would allow a more comprehensive search of protein data bases for potential SH2 domain binding sites. Recent studies with phospholipase C have revealed that SH2 domain binding specificity is more constrained at first thought and that residues at the +4 and +5 positions can have an inhibitory effect on the binding of some SH2 domains (62).

Acknowledgments—We thank Drs. Leah Cosgrove, M. J. Frenkel, and S. L. Macaulay for helpful comments and review of the manuscript. Dr. David Bowtell for the cDNA for GST-Grb2, and Dr. Tony Pawson for the cDNA of the SH2 domain of Shc as a GST fusion.

REFERENCES
1. Pawson, T. (1995) Nature 373, 573–580
2. Pawson, T., and Schlessinger, J. (1993) Curr. Biol. 3, 434–442
3. Waksman, G., Koninos, D., Robertson, S. C., Pant, N., Baltimore, D., Birge, R. B., Cowburn, D., Hanafusa, H., Mayer, B. J., Overduin, M., Reish, M. D., Schlessinger, J., and Kuriyan, J. (1992) Nature 358, 645–653
4. Edd, M. J., Shoelso, S. E., and Harrison, S. C. (1993) Nature 367, 89–91
5. Pascal, S. M., Singer, A. U., Gish, G., Yamazaki, T., Shoelso, S. E., Pawson, T., Kay, L. E., and Forman-Kay, J. D. (1994) Cell 77, 461–472
6. Lee, C. H., Koninos, D., Siebes, S., Margolis, B., Schlessinger, J., Shoelso, S. E., and Kuriyan, J. (1994) Structure 2, 423–438
7. Xu, R. X., Word, J. M., Davis, D. G., Rink, M. J., Willard, D. H., and Gampe, S. E. (1994) J. Biol. Chem. 269, 2574–2583
8. Schlessinger, J., and Ullrich, A. (1992) Cell 71, 373–383
9. Songyang, Z., Carraway, K. L., Eck, M. J., Harrison, S. C., Feldman, R. A., Kavanaugh, W. M., and Heidaran, M. A. (1994) Biochemistry 33, 746–754
10. Brown, K. J., and Ullrich, A. (1994) Cell 76, 587–599
11. Obermeier, A., Lammers, R., Wiesmuller, K. H., Jung, G., Schlessinger, J., and Vennström, S. (1992) EMBO J. 11, 3589–3597
12. Yarden, Y., Kuang, W. J., Yang-Feng, T., Couchins, L., and Ullrich, A. (1987) EMBO J. 6, 3341–3351
13. Yamada, Y., Tobe, K., Matuoka, K., Tamemoto, H., Ueki, K., Kaburagi, Y., Asai, S., Noguchi, T., Takahashi, Y., Miyazaki, M., and Ando, M. E. (1994) J. Biol. Chem. 269, 373, 536–539
14. Hubbard, S. R., Wei, L., Elis, L., and Hendrickson, W. A. (1994) J. Mol. Biol. 236, 1533–1534
15. Matsuda, M., Tobe, K., and Kaburagi, Y. (1994) J. Biol. Chem. 369, 2661–2672