Cell, Vol. 85, 695–705, May 31, 1996, Copyright ©1996 by Cell Press

Structure of the IRS-1 PTB Domain Bound to the Juxtamembrane Region of the Insulin Receptor

Michael J. Eck,*‡ Sirano Dhe-Paganon,†‡ Thomas Trüb,†‡ Robert T. Nolte,*‡ and Steven E. Shoelson†‡
*Laboratory of Molecular Medicine
Children's Hospital
†Research Division
Joslin Diabetes Center
‡Department of Medicine
Harvard Medical School
Boston, Massachusetts 02115

Summary

Crystal structures of the insulin receptor substrate-1 (IRS-1) phosphotyrosine-binding (PTB) domain, alone and complexed with the juxtamembrane region of the insulin receptor, show how this domain recognizes phosphorylated "NPXY" sequence motifs. The domain is a 7-stranded β sandwich capped by a C-terminal helix. The insulin receptor phosphopeptide fills an L-shaped cleft on the domain. The N-terminal residues of the bound peptide form an additional strand in the β sandwich, stabilized by contacts with the C-terminal helix. These interactions explain why IRS-1 binds to the insulin receptor but not to NPXpY motifs in growth factor receptors. The PTB domains of IRS-1 and Shc share a common fold with pleckstrin homology domains. Overall, ligand binding by IRS-1 and Shc PTB domains is similar, but residues critical for phosphotyrosine recognition are not conserved.

Introduction

Activated insulin receptor initiates signal transduction by phosphorylating its own intracellular β subunits as well as additional cellular proteins. Insulin receptor substrate-1 (IRS-1) is the best studied of these additional substrates. It serves as a link between the activated receptor and numerous effector proteins that bear Src homology-2 (SH2) domains (Sun et al., 1991, 1993). The first steps in signaling involve autoactivation of the receptor by phosphorylation of tyrosines 1146, 1150, and 1151 in its catalytic cleft (White et al., 1988b; Hubbard et al., 1994), and phosphorylation of tyrosine 960 in its juxtamembrane region. The latter phosphorylation leads to recruitment of IRS-1 (White et al., 1988a; O'Neill et al., 1994). At least two regions of the IRS-1 protein are necessary for efficient interaction with the activated receptor: a pleckstrin homology (PH) domain at the extreme N-terminus (Myers et al., 1995; Voliovitch et al., 1995) and a domain just C-terminal to it (Wolf et al., 1995). We have termed this second region the IRS-1 phosphotyrosine-binding (PTB) domain. The receptor sequence N-terminal to tyrosine 960 is also important for the interaction. In vitro binding experiments show that the IRS-1 PTB domain recognizes phosphotyrosine in the context of an NPXpY sequence motif with a hydrophobic residue at pY−8 (He et al., 1995; Wolf et al., 1995). A second substrate of the insulin receptor, IRS-2, has a similar domain structure and functions analogously in signaling (Sun et al., 1995).

A PTB or PI (phosphotyrosine interaction) domain has been identified independently in Shc, a substrate of numerous receptor-linked tyrosine kinases, including the insulin receptor (Blaikie et al., 1994; Kavanaugh and Williams, 1994; van der Geer et al., 1995). The Shc PTB domain also binds an "NPXpY" or related motif, and it requires a hydrophobic residue at pY−5 (Dikic et al., 1995; Trüb et al., 1995). Several proteins have homology with the Shc PTB domain (Bork and Margolis, 1995), including at least one reported to bind protein (Fiore et al., 1995). Nonetheless, the Shc PTB domain shares no recognizable sequence similarity either with the IRS domains or with SH2 domains.

We have determined X-ray crystal structures of the IRS-1 PTB domain, alone and in complex with a phosphopeptide sequence corresponding to its site of interaction with the insulin receptor. The domain is a 7-stranded β sandwich capped by a long α helix. It closely resembles a PH domain. All nine residues of the Ac-LYASSNPApY-NH_2 peptide contact the IRS-1 PTB domain. The phosphotyrosine interacts extensively with two arginines. The NPXpY motif of the peptide forms a type I β turn, which lies on a platform created by a 3_{10} turn between strands β4 and β5 of the domain. The asparagine at pY−3 appears to be essential for binding, as its side chain forms both intrapeptide and intermolecular hydrogen bonds. Peptide residues pY−3 to pY−8 lie in a groove parallel to the helix and to strand β5, and create an additional strand on the domain β sandwich.

A recent multidimensional nuclear magnetic resonance study has revealed a related structure for the Shc PTB domain (Zhou et al., 1995), and the general orientations for peptide binding to the two PTB domains are similar. However, the residues that contact the phosphotyrosine are different in the two domains. One critical arginine in the Shc domain lies in a long insertion between strands β1 and β2 that is not present in IRS-1. The domain surface used by the IRS-1 PTB domain in phosphopeptide binding is also different from the one used by PH domains to bind phosphoinositides (Ferguson et al., 1995; Hyvönen et al., 1995). Therefore, the IRS-1 and Shc PTB domains and PH domains share similar folds, but they have evolved distinct mechanisms for phospholigand recognition.

Results and Discussion

Delineating and Crystallizing the PTB Domain

Sequence alignments between IRS-1 and IRS-2 reveal two regions of extended homology, termed IH-1 and IH-2 (IRS-1 homology 1 and 2) (Sun et al., 1995). The first is predicted to be a PH domain (Musacchio et al., 1993). The second does not align readily with known protein sequences in data bases. In a previous study, the entire 173-residue IH-2 domain of human IRS-1 was

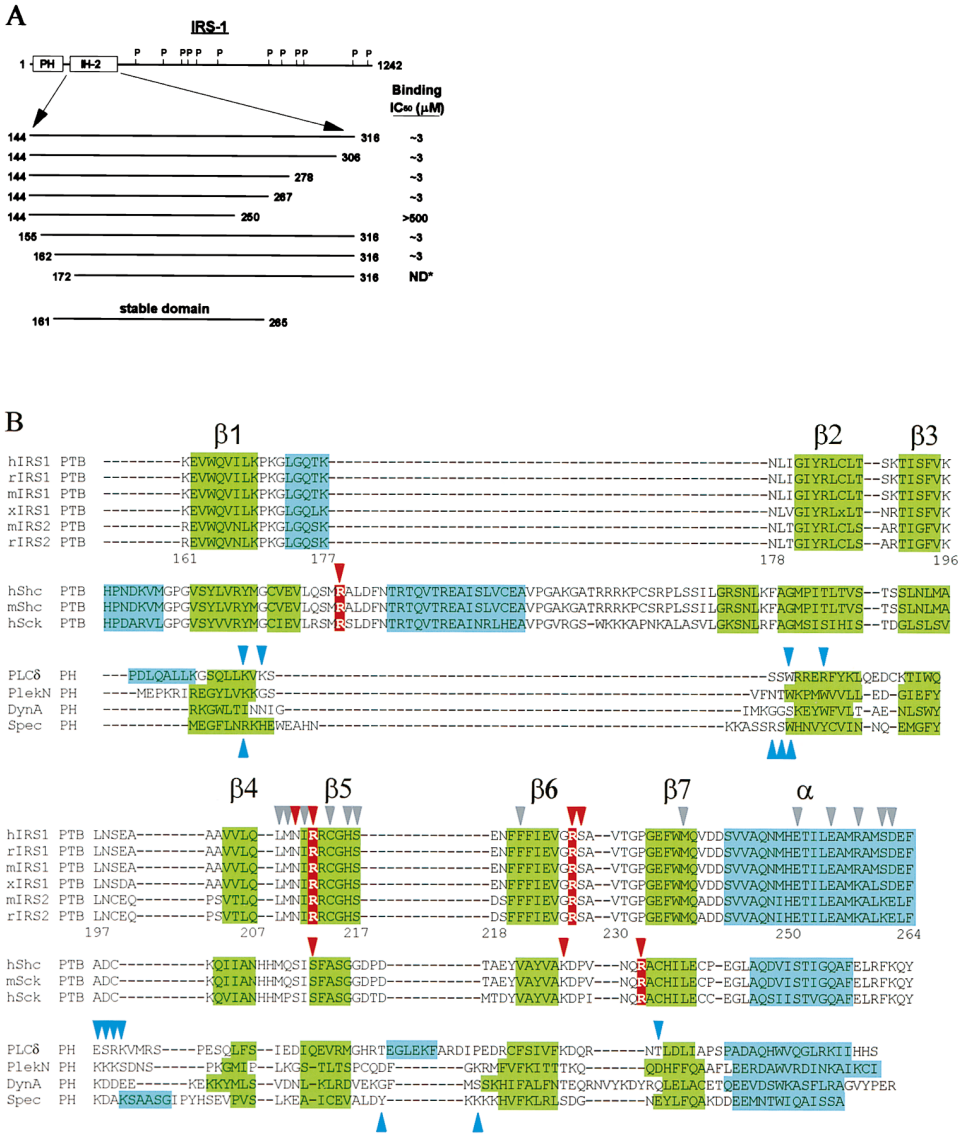

Figure 1. Functional Mapping of the IRS-1 PTB Domain and Structure-Based Sequence Alignments of PTB and PH Domains

(A) Domain structure of IRS-1 and phosphopeptide binding activity of expressed proteins. IRS-1 contains a PH domain and a second homology domain (termed IH-2) having PTB domain activity. Serial truncations at both ends of the IH-2 domain show that PTB activity is preserved with significantly shorter proteins and suggest that residues 162–267 encompass a fully active domain, consistent with the stably folded domain (161–265) observed in our crystal structure. (asterisk) ND, not determined; rapid proteolysis of 172–316 in E. coli suggests that it is not stably folded.

(B) Structure-based sequence alignments between IRS and Shc PTB domains and several PH domains. Sequences of human IRS-1 (161–265) and Shc (40–210) PTB domains, and rat phospholipase Cδ (11–140), human N-terminal pleckstrin (1–103), human dynamin (506–629), and mouse brain spectrin (2199–2304) PH domains, were aligned by superimposing the structures. Numbers refer to human IRS-1. α helices are boxed in turquoise, β-sheet strands in green, and critical arginines for phosphotyrosine recognition are boxed in red. Arrowheads denote residues involved in ligand binding: red for pY binding, gray for residues of IRS-1 interacting with insulin receptor sites other than pY, and blue for residues of PLCδ and spectrin PH domains that bind phosphoinositide polyphosphate.

expressed in Escherichia coli and found to possess PTB activity (Wolf et al., 1995). For crystallographic studies, additional truncated versions were expressed to delimit boundaries of the PTB domain (Figure 1A). Phosphoprotein binding was retained following removal of residues from both amino and carboxyl ends of the homology region. These truncation data suggest that residues 162–267 constitute a stably folded and functional PTB domain, a segment considerably shorter than the 173 residues identified by sequence alignment, but remarkably consistent with the ordered region observed in our crystal structure (161–265). The protein used for crystallization included residues 144–267 of IRS-1, and two additional residues (Gly–Ser) at the amino terminus. This fragment was generated by thrombin proteolysis of a glutathione S-transferase (GST) IRS-1 144–316 fusion

protein. The cleaved protein retained full PTB function and crystallized in the presence and absence of its natural ligand, the juxtamembrane region of the insulin receptor.

The 9-residue peptide ligand in our crystals has the sequence Leu-Tyr-Ala-Ser-Ser-Asn-Pro-Ala-pTyr, which corresponds to receptor residues 952–960, with alanine rather than glutamic acid at position pY−1. The alanine substitution increases binding affinity 30-fold (Wolf et al., 1995).

Structure of the Domain

The IRS-1 PTB domain is a compact 7-stranded β sandwich, capped on one end by a long α helix. This fold is also found in PH domains, a protein module found in many signal transduction proteins. To facilitate comparisons with this structurally related domain, the same nomenclature used for elements of secondary structure in PH domains has been applied to the PTB domain. A ribbon diagram showing the arrangement of secondary structural elements and the position of the bound phosphopeptide is shown in Figure 2A. One face of the β sandwich is formed by β1-β4, arranged in an antiparallel β sheet. Strands β1 and β2 are long and curved, wrapping around to contribute to the β5-β7 face. The polypeptide segment connecting strands β1 and β2 forms 1 1/2 turns of an α helix, which partially "caps" the bottom of the β sandwich. The connections between the other β strands and between β7 and the C-terminal helix are rather short, giving the domain its compact appearance. A 3_{10} helical turn connects strands β4 and β5, a feature of the PTB domain that may be important for recognition of phosphorylated NPXpY peptides (see below).

The free and complexed domains are superimposed in Figure 2B. The structures deviate most in the loop connecting strands β3 and β4, a region that is poorly ordered in the apodomain. The carboxy-terminal helix is a full turn longer in the liganded structure, likely due to stabilization of the helix by interactions with the bound phosphopeptide. The penultimate turn of the helix (residues 257–261), which interacts directly with the Asn-Pro-Ala turn in the complexed peptide, is poorly ordered in the unliganded structure.

Peptide Recognition

The insulin receptor phosphopeptide fills an L-shaped cleft on the surface of the PTB domain (Figures 2A and 3A). The recognition site is formed by the β5 strand, the C-terminal helix, and the 3_{10} turn connecting strands β4 and β5. The entire 9-residue peptide is in contact with the domain; the interaction buries 779 Å^2 of surface area. The amino-terminal residues of the peptide (pY−8 to pY−3) form a β strand that hydrogen bonds to strand β5 in an antiparallel orientation (Figure 4A). The C-terminal residues of the peptide NPApY form a type I β turn, which positions the phosphotyrosine residue so that it extends at a 90° angle from the amino-terminal peptide backbone (Figures 4B). The phosphotyrosine is coordinated primarily by two arginine residues, Arg-212 and Arg-227, which extend from the β5 and β6 strands, respectively (Figures 1B and 5B).

In addition to β sheet (backbone) interactions, there are numerous contacts between the domain and peptide side chains that contribute specificity to the interaction. The leucine at pY−8 is important for high affinity binding to the insulin receptor (Gustafson et al., 1995; Wolf et al., 1995). It is bound in a shallow hydrophobic pocket between Phe-222 and Trp-237 (Figures 3A and 4B). Tyrosine at pY−7 is largely solvent exposed, but it is in van der Waals contact with His-216. The side chain of the pY−6 residue, an alanine in the IR peptide, extends toward Phe-222 and Arg-212 on the domain, and leucine

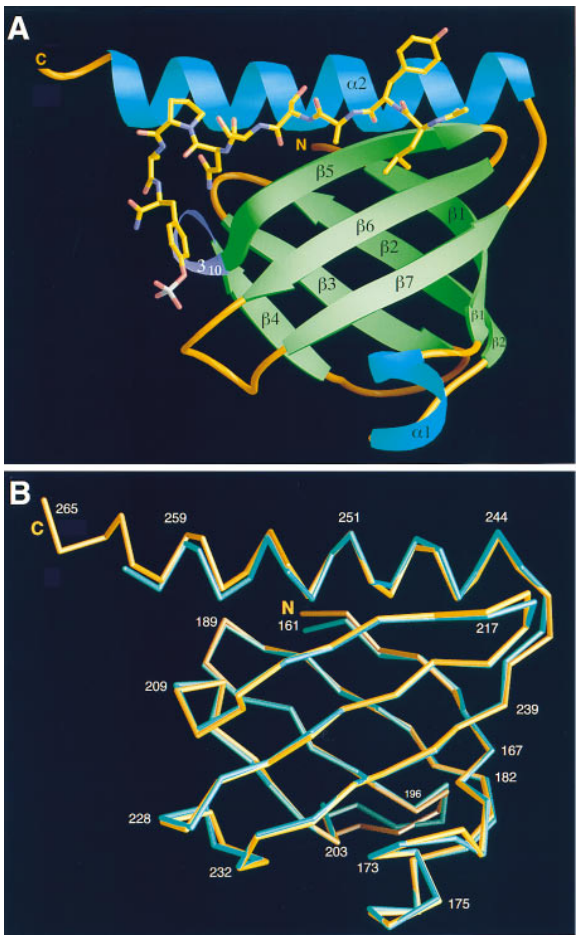

Figure 2. Structure of the Free and Complexed IRS-1 PTB Domains

(A) RIBBONS (Carson, 1991) diagram showing the fold of the IRS-1 PTB domain and the orientation of the bound insulin receptor peptide. The IRS-1 domain is a compact 7- stranded β sandwich (green) closed on top by helix α2 (turquoise). A short helix (α1) and its connecting loops partially cap the bottom of the sandwich. A 3_{10} helical turn (deep blue) connects strands β4 and β5 and forms part of the binding site for residues 952–960 of the insulin receptor peptide (shown in stick representation).

(B) Superposition of free and complexed IRS-1 PTB domains. Cα tracings of the bound (orange) and free (green) PTB domains show minimal differences other than an extension of the helix in the presence of bound peptide. Cα atoms of core residues of the structures superimpose with an RMS deviation of 0.29 Å.

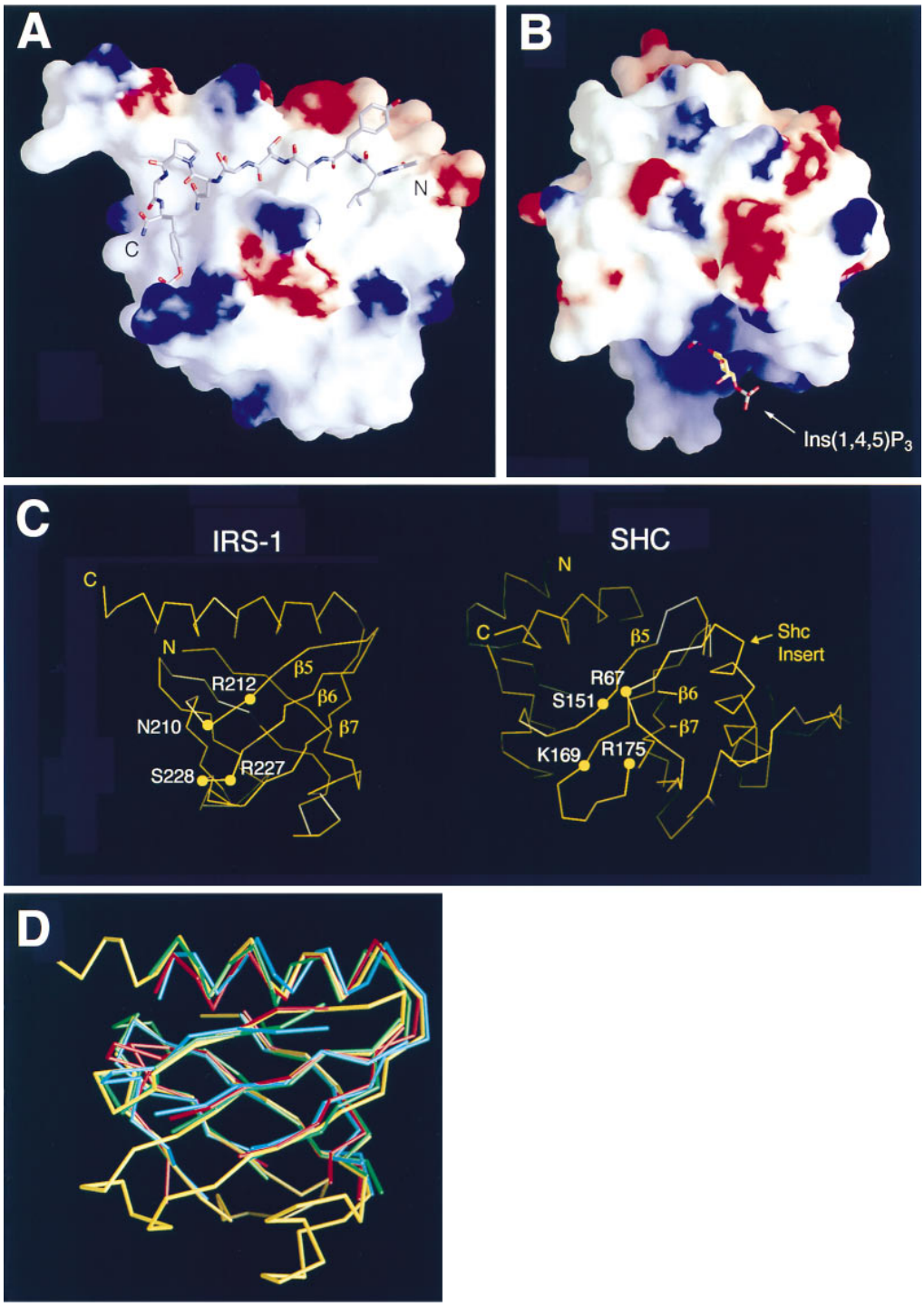

Figure 3. Comparisons of PTB and PH Domains

Molecular surface representation of (A) the IRS-1 PTB domain and (B) the phospholipase Cδ PH domain (Ferguson et al., 1995), shown roughly in the same orientation as in Figure 2. The molecular surfaces were calculated and shaded according to electrostatic potential (−5kt/e, red to +7kt/e, blue) using the program GRASP (Nicholls et al., 1991). (A) The insulin receptor peptide (gray) fills an L-shaped cleft on the surface of the PTB domain. The binding surface is largely uncharged outside of the phosphotyrosine-binding pocket. (B) Inositol 1,4,5-trisphosphate (red) binds in the center of a large positively charged region at the base of the PH domain. This electrostatic polarization is a feature of all PH domains of known structure but is not observed in the IRS-1 PTB domain. (C) Divergent recognition of phosphotyrosine by IRS-1 and Shc PTB domains. Cα tracings of the IRS-1 (left) and Shc (right) PTB domains are shown. Key residues involved in phosphotyrosine coordination (yellow balls) are in similar spatial positions, yet arise from different regions in the primary sequences and distinct elements of secondary structure. (D) PTB and PH domains share a common fold. Cα carbons of the phospholipase Cδ (green), dynamin (blue), and spectrin (red) PH domains are superimposed on the IRS-1 PTB domain (yellow). For clarity, selected loops from the PH domains have been omitted. Cα atoms of these core residues superimpose with an RMS deviation of ~1.0 Å.

(pY−8) on the peptide. Serine, the pY−5 residue, forms a hydrogen bond with His-250 on the carboxy-terminal helix of the domain, and it is also packed against Leu-254 in the next turn of the helix. Bulky hydrophobic side chains are present at pY−5 in the optimal Shc PTB domain recognition motifs found in growth factor receptors and transforming proteins (Dikic et al., 1995; Kavanaugh et al., 1995; Trüb et al., 1995; Wolf et al., 1995). There is insufficient space in the IRS-1 PTB domain complex to accommodate large, hydrophobic side chains at peptide position pY−5. It is therefore clear from our structure why IRS-1 does not bind to growth factor receptors. This discrimination may in turn explain why growth factor receptors do not phosphorylate IRS-1. The hydroxyl group of serine at pY−4 extends away from the surface of the domain.

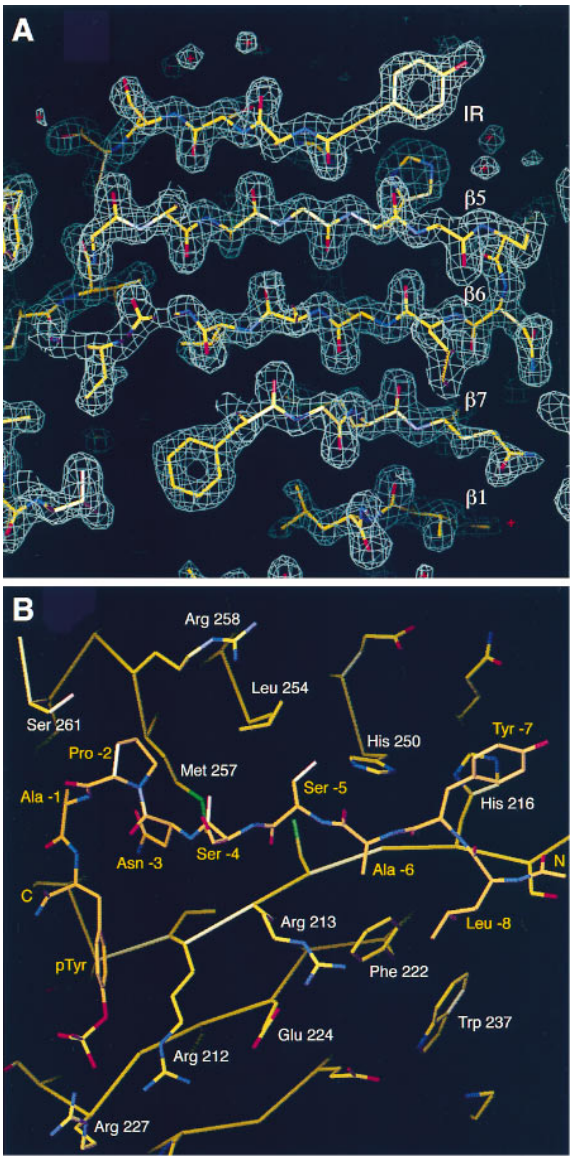

Figure 4. Structure and Coordination of the Insulin Receptor Peptide

(A) The bound insulin receptor peptide (IR) forms an additional strand in the β sandwich. Bound water molecules are represented as red crosses. The 1.8 Å 2Fo-Fc electron density map is contoured at 1.5 σ.

(B) Interactions between the insulin receptor and IRS-1. Cα trace and selected side chains of the PTB domain (yellow) and bound peptide (tan) are shown.

A common consensus motif NPXpY appears in sites recognized by both IRS-1 and Shc PTB domains, suggesting that the sequence and conformation of positions pY−3, pY−2, and pY−1 are particularly critical for binding. In our structure, these residues adopt a type I β-turn conformation, thereby defining the overall shape of the bound peptide (Figures 2A and 5B). In addition, they interact extensively with the protein. Asparagine is particularly critical. Its side chain oxygen accepts an intrapeptide hydrogen bond from the amide nitrogen of Ala-1, an i to i+2 interaction that often stabilizes Asn–Pro-containing turns. Moreover, the side chain NH_2 group donates hydrogen bonds to the carbonyl oxygens of Leu-208 and Ile-211 in the 3_{10} turn. The bound peptide buries these carbonyls, which coordinate water molecules in the unliganded structure. From the structure it would seem likely that Asn is essential at this position, as it is unique among amino acids in its ability to satisfy all of the observed hydrogen bonding constraints.

The 3_{10} turn and the C-terminal residues of helix α2 form the binding surface for the NPApY sequence. The carbonyls in the 3_{10} turn accept hydrogen bonds from the peptide: the carbonyls of Leu-208 and Ile-211, from Asn (pY−3), as just described, and the carbonyl of Met-209 from the terminal amide (and probably from the main-chain amide of residue pY+1 in the intact insulin receptor). The rest of the binding surface is hydrophobic. The clustering of three methionines is striking—Met-257 and Met-260 in the helix and Met-209 in the 3_{10} turn ring the "NPXpY"-binding pocket (Figure 5A). Additional hydrophobic interactions are contributed by methylene groups from the side chains of Arg-258 and Ser-261, which contact proline at pY−2. Alanine appears to be an optimal residue at the pY−1 position. Larger hydrophobic side chains could not be accommodated in the shallow pocket. Larger polar side chains, including the glutamic acid found naturally at this position in the insulin receptor, could turn and extend into solvent.

Phosphotyrosine Recognition

The pY side chain lies in an open pocket created by the 3_{10} turn and residues at the end of strands β5 and β6. An extensive network of hydrogen bond and ionic interactions coordinate the phosphate oxygens (Figure 5B), consistent with the observation that phosphorylation of the tyrosine is necessary for detectable peptide binding (Wolf et al., 1995). Two arginine residues line the binding pocket. Arg-212 in strand β5 is in particularly intimate contact with phosphotyrosine. Its aliphatic side chain extends along the hydrophobic surface of the aromatic ring, and its guanidinium group forms bidentate hydrogen bonds with phosphate oxygens. Glu-224 forms a long hydrogen bond with Arg-212, which may help orient its guanidinium group for phosphotyrosine binding; in the unliganded structure Arg-212 is poorly ordered. The side chain of Arg-227 extends from the loop connecting strands β6 and β7 and also donates two hydrogen bonds

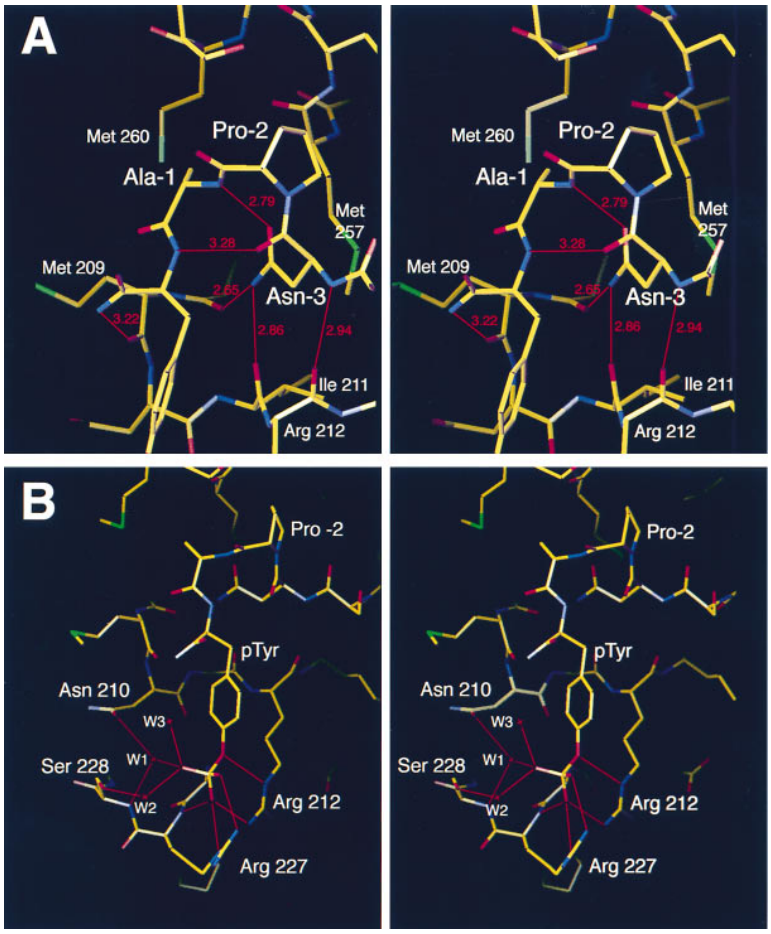

Figure 5. Recognition of the NPApY Sequence

(A) Detail of the NPApY β turn. Intrapeptide and intermolecular hydrogen bonds and distances are shown in red. Three methionines form a hydrophobic patch on the PTB domain, and contact the peptide.

(B) Phosphotyrosine is recognized primarily by domain arginines 212 and 227. Ser-228 and Asn-210 coordinate water molecules (red crosses) that make additional interactions with the phosphate group.

to the phosphate oxygens. The main-chain amide of Arg-227 contributes a third hydrogen bond to the phosphate.

Two additional residues, Asn-210 and Ser-228, interact with the phosphotyrosine by coordinating "bridged" water molecules. The side chain of Asn-210 and the backbone amide of Ser-228 hydrogen bond with water W1, which is also found in the apodomain structure. The hydroxyl group of Ser-228 coordinates a second bridging water molecule (W2).

Comparison with the Shc Domain

The phosphotyrosine-binding domains in Shc and IRS-1 share little sequence homology, and these binding domains were independently recognized and characterized. A search of sequence databases using the sequences of Shc and its close homolog Sck identified several possible PTB domains (Bork and Margolis, 1995), at least one of which was subsequently shown to bind protein (Fiore et al., 1995). This search did not identify the IRS-1 domain. Because of their functional similarities, a structural relationship between the Shc and IRS-1 domains has been suspected, but a previous attempt to align IRS and Shc family PTB domains is not consistent with the three-dimensional structures (Sun et al., 1995).

Comparison of the IRS-1 PTB domain with the recently determined nuclear magnetic resonance structure of the Shc domain shows that both share the core PH domain fold (Zhou et al., 1995). The longer (~170 residue) Shc domain has a short N-terminal helix not observed in our structures and a large insertion between strands β1 and β2, which forms a helix and two long connecting loops (Figure 1B). The two domains also share a common overall mode of peptide binding—the general conformation and orientation of the phosphopeptide in complex with the Shc domain is the same as in the IRS-1 domain. There is, however, a surprising divergence in the mechanism of phosphotyrosine recognition. None of the residues that coordinate the phosphotyrosine appear to be conserved between the two domains (Figures 1B and 3C). The Shc structure shows four residues likely to interact with phosphotyrosine: Arg-67, Ser-151, Lys-169, and Arg-175. Site-directed mutations of Arg-67 and Ser-151 both decrease binding to phosphopeptides, and mutation of Arg-175 to Gln or Lys completely eliminates it (Zhou et al., 1995). None of these residues are identical to their counterparts in IRS-1. Arg-175, which appears to be most critical for binding in Shc, corresponds to proline 233 of IRS-1, while Arg-67 of Shc is in a long insertion not present at all in IRS-1. Likewise, Arg-212, the residue we predict to be most crucial for phosphotyrosine binding to the IRS-1 domain, corresponds to Ser-151 in Shc. Of the four residues in the Shc-binding site, Lys-169 has the closest homology in

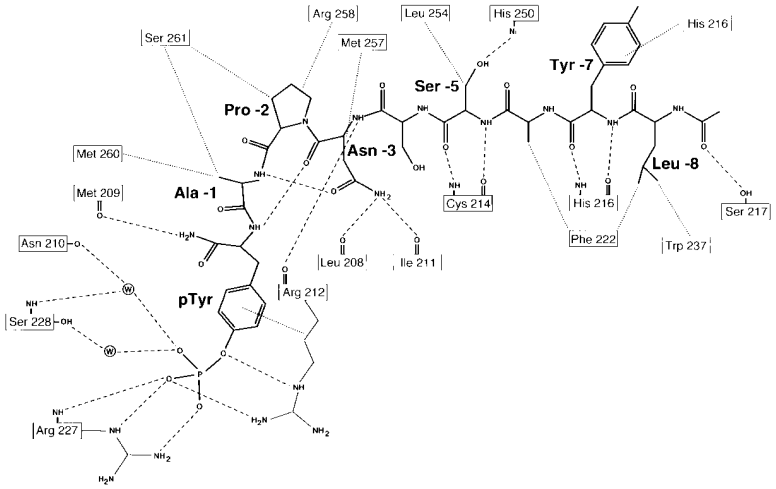

Figure 6. Schematic Representation of Interactions between the Insulin Receptor Peptide and the IRS-1 PTB Domain

Hydrogen bonds are shown as dashed lines, hydrophobic interactions as thin dotted lines.

IRS-1, being in roughly the same position as Arg-227. This divergence contrasts sharply with the conservation of residues in the phosphotyrosine-binding pockets of SH2 domains (Waksman et al., 1992; Eck et al., 1993). The βB5 or "FLVRES" arginine is found in all SH2 domains. A second arginine at αA2 is found in all but a few, and when absent, it does not appear to be replaced by other basic residues (Lee et al., 1994; Eck et al., 1996).

The extent to which the network of intra- and intermolecular interactions involving Asn-3 are conserved in Shc is not yet clear. This residue is retained in all known PTB recognition sites, and given its observed role in stabilizing the conformation of the NPApY β turn and in pairing with buried carbonyl oxygens in the 3_{10} turn in our structure, we would expect similar interactions to be a feature of all PTB domains. The pY−3 asparagine side chain appears to be poorly defined by nuclear Overhauser effects in the Shc structure. Interactions with the i+2 amide nitrogen and with the turn connecting strands β4 and β5 (the location of the 3_{10} turn in IRS-1) are not described, though the peptide adopts a grossly similar β-turn conformation. Interactions of the Asn side chain NH_2 with the aromatic ring and carbonyl of Phe-198 in the C-terminal helix of Shc may replace those that we observe in IRS-1.

Comparison with PH Domains

The structural similarity between PH and PTB domains was first revealed in a comparison of the Shc structure with PH domains (Zhou et al., 1995). The structural similarity of the IRS-1 domain with PH domains is even more striking. The α-carbon trace of the IRS-1 domain is shown superimposed with the spectrin (Hyvönen et al., 1995), phospholipase Cδ (Ferguson et al., 1995), and dynamin (Ferguson et al., 1994) PH domain crystal structures in Figure 3D. The RMS deviation between ∼60 superimposed core residues in these PH domains and IRS-1 is ∼1 Å, no greater than the deviation among the PH domains themselves. The structures diverge markedly in the length and conformation of loops connecting elements of secondary structure, but the IRS domain does not stand out in comparison to the PH domains. The C-terminal helix is longer in the liganded IRS-1 domain than in PH domains of known structure, but a longer helix was not observed in the Shc PTB domain, so this is not a defining feature of PTB domains. A conserved tryptophan present in the C-terminal helices of the PH domains is not present in the IRS-1 PTB domain. However, the side chain of Trp-164, which arises from the β1 strand, occupies a similar position within the hydrophobic core of the IRS-1 domain.

The common fold and similarities we observe in hydrophobic packing between the PH and PTB domains, together with their related roles in signal transduction, argue for a common evolutionary origin. The function of PH domains is poorly understood, but at least a few of these domains appear to mediate membrane localization by stereospecific binding to phosphoinositide head groups. Structures of the spectrin and phospholipase Cδ PH domains have been determined in complex with inositol-(1,4,5)-trisphosphate (Ferguson et al., 1995; Hyvönen et al., 1995). PH domains of known structure are electrostatically polarized—the end of the β sandwich opposite the C-terminal helix contains a cluster of basic residues and a net positive charge. This polarization has been suggested to play a role in membrane attachment, and indeed the inositol trisphosphate in the PLCδ structure binds in the center of this positively charged region (Figure 3B). No such polarization is observed in the IRS PTB domain, which is shown side-by-side with the PLCδ PH domain in Figures 3A and 3B. Notably, the IRS-1 PTB domain does not bind inositide polyphosphates with high affinity (Zhou et al., 1996; L. Rameh and L. Cantley, personal communication).

The comparison also makes clear that despite a similarity in function (binding phosphate groups) and conserved secondary structure, the locations of ligand-binding sites are quite different on PH and PTB domains. The use of a conserved, modular fold for a variety of functions is a common theme in protein evolution, so it is not surprising to see divergence in function between PH and PTB domains. More remarkable is the divergence in the structure/function relationship between the Shc and IRS-1 domains, because conserved structure

and conserved function usually implies a closely related mechanism. Although IRS-1 and Shc bind similar ligands with relatively high affinity and in the same general mode, the two PTB domains apparently use divergent sets of residues to accomplish this task.

Comparison with SH2 Domains

There is no structural relationship between SH2 domains and PTB domains. Both domains specifically bind phosphotyrosine-containing peptides, but they differ markedly in the way in which this specificity is achieved. In both domains, two arginine residues coordinate phosphate oxygens, but the geometry of the interactions is distinct, because the arginines are presented by a different protein architecture. The hydrophobic interaction of Arg-212 with the phosphotyrosine ring and the hydrogen bonds to the phosphate group are reminiscent of the interaction of Lys βD6 seen in some Src family SH2 structures (Waksman et al., 1992; Eck et al., 1993)

PTB and SH2 domains differ more dramatically in their recognition of sequences flanking phosphotyrosine. Structural features of the SH2 domain allow marked variation in sequence specificity and sequence context. Insertions and sequence divergence in the EF and BG loops of SH2 domains, and more modest substitutions in the βD5 strand, allow for selectivity in binding residues carboxy-terminal to the phosphotyrosine without perturbation of the SH2 domain fold (Eck et al. 1993; Waksman et al. 1993; Lee et al. 1994; Nolte et al., 1996). In contrast, PTB domains bind residues N-terminal to phosphotyrosine using elements of secondary structure — strand β5 and the C-terminal helix. Divergence in the peptide sequences recognized by PTB domains is therefore restrained by the requirement for an asparagine-directed turn and for an extended β strand. In SH2 domains, backbone conformations of bound peptides vary, but hydrogen bond interactions to the domain are maintained by compensating changes in SH2 domain structure. These structural differences suggest a more limited ability of the PTB domain to diverge in its recognition of peptide sequences, which may be consistent with the functions of PTB and SH2 domains. PTB domains are found in substrates of receptor tyrosine kinases, while SH2 domains are typically associated with downstream enzymes. Perhaps a more limited repertoire of recognition motifs is suitable at this higher point in diverging pathways.

Implications

Three-dimensional structures are now known for two PTB domains. Although these domains share the general PH domain fold, their mechanisms for phosphopeptide recognition are distinct. High resolution, liganded structures for the PLCδ and spectrin PH domains have been determined, as well. Inositide polyphosphate groups bind at different sites on the two PH domains, at the base of the β sandwich and distant from the surfaces used by the PTB domains for phosphate recognition. Thus the PH domain fold appears to represent a versatile scaffold for achieving diverse functions. There are nearly 90 related domains identified by sequence homology, and more likely exist. It will be interesting to learn what ligands bind these domains and how recognition occurs.

The domain architecture of IRS proteins facilitates a segregation of distinct functions, as is common to signaling molecules. The PTB domain is necessary for efficient phosphorylation, probably by increasing effective concentrations of substrate at the kinase active site and by allowing it to be phosphorylated at multiple sites. Flanking regions of IRS proteins may assist in this function. The PH domain N-terminal to the PTB domain and binding regions carboxy-terminal to it may participate by binding adjacent phospholipids or peptide sequences (Gustafson et al., 1995; Myers et al., 1995; Voliovitch et al., 1995; He et al., 1996; Sawka-Verhelle et al., 1996). In this context, the PTB and PH domains, and possibly additional domains in IRS proteins, may function together to enhance affinity and specificity, much as has been seen previously with tandem SH2 and SH3 domains (Eck et al., 1994, 1996; Hatada et al., 1995).

Experimental Procedures

Protein Expression and Binding Studies

BamHI and XhoI sites were introduced by polymerase chain reaction immediately up- and downstream of the regions of the human IRS-1 cDNA (Araki et al., 1993) encoding residues 144–250, 144–278, 144–306, 144–316, 155–316, and 162–316. The DNA fragments were subcloned into the corresponding sites in a pGEX-4T-1 (Pharmacia) plasmid and used to transform E. coli strain XL-1 blue (Stratagene). Following induction of protein expression with IPTG (isopropyl thio β-D-galactoside) and cell collection and lysis, the protein was purified by affinity chromatography using an immobilized glutathione agarose column (Molecular Probes). For binding studies, GST fusion proteins were eluted with 50 mM glutathione and dialyzed against 100 mM ammonium bicarbonate containing 1.0 mM dithiothreitol (DTT). Each protein was analyzed by direct binding with a radiolabeled insulin receptor–derived peptide, [^{125}I]LYASSNPApYLSASDV, as described (Trüb et al., 1995; Wolf et al., 1995).

Crystallization of the IRS-1 PTB

For crystallization, the GST/IRS-1 144–316 fusion protein bound to a glutathione column was cleaved with bovine thrombin (10 U/mL; Sigma) in 25 mM Tris–HCl buffer (pH 8.5) containing 150 mM NaCl and 5 mM $CaCl_2$. This procedure yields free IRS-1 144–267, owing to cleavage at the engineered site in the vector and an additional highly susceptible thrombin site at residue 267 within IRS-1. Two additional residues (Gly–Ser) at the amino terminus of the cleaved protein arise from the vector. Recovered protein was further purified using a Mono-Q HR10/10 ion-exchange column eluted with a linear NaCl gradient (0–200 mM in 180 mL at 2.0 mL/min) in 25 mM Tris–HCl (pH 8.5) containing 1 mM DTT. The protein was concentrated in a Centricon-10 device to 20 mg/ml. Apodomain crystals were grown using the hanging drop vapor diffusion method by combining 2 μL of protein solution (3 mg/ml, 25 mM Tris–HCl [pH 8.5], 0.1 M NaCl, 10 mM DTT) with 2 μL of well solution (1.7 M ammonium sulfate, 0.1 M sodium citrate [pH 5.6], 0.3 M NaCl, 20 mM DTT). The orthorhombic crystals ($P2_12_12_1$ a = 35.07 Å, b = 50.71 Å, c = 56.26 Å) grew to about 0.15 × 0.15 × 0.2 mm over several days at 22°C. Crystals were harvested into a synthetic mother liquor containing 2.0 M ammonium sulfate, 0.3 M NaCl, 0.1 M sodium citrate [pH 5.6]. Heavy atom derivatives were prepared by soaking crystals for 1–2 days in this harvest buffer plus 0.2 mM methyl mercury nitrate, 0.25 mM mercuric chloride, or 1.0 mM potassium platinum tetrachloride.

Crystals of the PTB domain in complex with the insulin receptor peptide were grown by the same method using 2 μL protein solution (10 mg/ml, 1 mM IR peptide, 25 mM Tris–HCl [pH 8.5], 0.1 M NaCl, 10 mM DTT) with 2 μL of well solution (1.7 M ammonium sulfate, 0.1 M NaCl, 2% glycerol, 0.1 M PIPES [pH 6.5], 20 mM DTT). The insulin receptor phosphopeptide acetyl-LYASSNPApY-NH_2 was synthesized following an N^{α}-Fmoc protecting group strategy and

Table 1. Structure Determination and Refinement

Summary of MIR Phase Determination

Derivative	Resolution Limit	Completeness (%)	R_{merge}	Isomorphous Difference	Phasing Power		Number of Sites
					Acentric	Centric	
Native	1.95	98.1	4.3	–	–	–	–
$MeHgNO_3$	2.5	86.6	5.9	30.5	1.28	0.96	1
$MeHgNO_3$	2.35	98.5	5.1	20.3	1.32	1.05	1
$HgCl_2$	2.35	88.4	6.8	29.8	0.76	0.56	1
K_2PtCl_4	2.5	96.4	5.1	25.3	0.65	0.54	1
K_2PtCl_4	2.5	99.1	6.5	21.6	0.66	0.56	1

Data Collection and Refinement Statistics

	Complex	Unliganded
Space group	P422	$P2_12_12_1$
Unit cell	a = 68.37, c = 57.52	a = 35.07, b = 50.71, c = 56.26
Reflections		
Observed/unique	72,265/12,453	32,748/7589
R_{merge} (%)	6.57	4.29
Completeness (%)	94.9	98.2
I/σI (last shell)	2.93	6.23
Refinement		
Resolution	20.0–1.8 Å	20.0–1.95 Å
R_{cryst}/R_{free} ($I > 2\sigma$)	17.1/20.6	18.6/25.2
R_{cryst}/R_{free} (all data)	19.6/23.1	20.3/27.3
RMS deviation		
Bond length (Å)	0.008	0.011
Bond angles (deg)	1.17	1.1

purified by high pressure liquid chromatography (Piccione et al., 1993). Tetragonal crystals (P422; a = b = 68.37 Å, c = 57.52 Å) grew to about 0.1 × 0.1 × 0.3 mm over several days at 22°C.

Data Collection and Structure Determination

All diffraction data were recorded at room temperature using an 18 or 30 cm MarResearch image plate scanner on an Elliot GX-13 rotating anode source with double-focusing mirror optics. Oscillation images (1°) were processed using a Mar version of the program XDS (Kabsch, 1988). Diffraction intensities were scaled and merged with the program MARSCALE (Kabsch, 1988). For heavy atom derivative data sets, local scaling to the Wilson-scaled native data was carried out with MAXSCALE (M. Rould).

The CCP4 program suite (Collaborative Computational Project Number 4, 1994) was employed for structure determination. Heavy atom occupancies were initially refined with the program HEAVY (Terwilliger and Eisenberg, 1988). Final phasing calculations were carried out with MLPHARE (Otwinowski, 1991). The structure of the apodomain was determined by multiple isomorphous replacement (MIR) using methyl mercury nitrate, mercuric chloride, and potassium platinum tetrachloride derivatives. Inclusion of two methyl mercury and two platinum derivative data sets improved the quality of the phases even though the duplicate data sets had identical sites and similar occupancy. Anomalous data from all derivatives were included in the phase calculations. The anomalous signal from the methyl mercury derivative was particularly strong and was used to determine the hand of the heavy atom constellation. Phasing statistics are presented in Table 1. The clarity of the electron density map was improved with real space density modification using the program DM (Cowtan, 1994). Experimental phases were extended from 2.7 to 2.0 Å with solvent flattening, histogram matching, and Sayre's equation algorithms in DM. The complete protein main chain and ~80% of the side chains were readily constructed in the modified MIR map using the program O (Jones et al., 1989). A map calculated with model phases after an initial round of crystallographic refinement allowed interpretation of the remaining side chains. The structure was refined with iterative cycles of manual fitting and simulated annealing and positional refinement in X-plor (Brunger, 1992). Isotropic temperature factors were refined. Incorporation of a solvent correction model allowed inclusion of low resolution data in refinement. Ordered water molecules were positioned with the aid of the program ARP (Lamzin and Wilson, 1993). The final model, including residues 161 to 261 and 119 solvent molecules, was refined to an R value of 18.6% (R_{free} = 25.2%) with good stereochemistry (Table 1).

The structure of the PTB domain/peptide complex was determined by molecular replacement using the partially refined apodomain structure as a search model. Unambiguous rotation and translation solutions were obtained using data between 8.0 and 3.0 Å resolution with the program AMoRe (Navaza, 1992). After rigid body and simulated annealing refinement of the properly positioned search model, 2Fo-Fc maps revealed clearly interpretable electron density for the entire phosphopeptide ligand and for most of the 20 protein side chains not present in the search model. As in the unliganded structure, no electron density was observed for the 19 residues amino-terminal to Lys-161 (144–160). Ordered water molecules were positioned with the aid of the program ARP. The structure was refined using simulated annealing and positional refinement protocols in X-plor to a crystallographic R value of 17.1% (R_{free} = 20.6%). The R_{free} value for the complex is significantly lower than that for the unliganded structure, which is likely attributable to the lower Wilson temperature factor for data recorded from crystals of the complex. A solvent correction model allowed inclusion of low resolution data in refinement. Isotropic temperature factors were refined with tight restraints. The final model included residues 161–265 of IRS-1, the 9-residue peptide, and 120 water molecules. Data collection and refinement statistics are summarized in Table 1.

Acknowledgments

The authors thank Masaya Miyazaki, Lone Hansen, and Gert Wolf for assistance in peptide synthesis, obtaining PTB domain proteins, and assessing PTB domain function, respectively; Kazuyasu Sakaguchi (National Cancer Institute/National Institutes of Health) for electrospray mass spectrometric analyses; and Andrea Musacchio for helpful discussions. Special thanks are given to Stephen Harrison (Howard Hughes Medical Institute, Children's Hospital, Boston) for continued support and advice given throughout this project. Financial support for these studies was provided by grants from the Lucille P. Markey Charitable Trust to Children's Hospital (M. J. E.),

the Juvenile Diabetes Foundation (S. E. S.), and NIH (DK43123 to S. E. S.). M. J. E. is the recipient of a Burroughs-Wellcome Fund Career Award in the Biomedical Sciences. S. E. S. is the recipient of a Burroughs-Wellcome Fund Scholar Award in Experimental Therapeutics. The Joslin Diabetes Center biochemistry facility is supported by a Diabetes and Endocrinology Research Center grant from the NIH.

Received March 6, 1996; revised April 10, 1996.

References

Araki, E., Sun, X.-J., Haag, B.L.I., Chuang, L.-M., Zhang, Y., Yang-Feng, T.L., White, M.F., and Kahn, C.R. (1993). Human skeletal muscle insulin receptor substrate-1. Diabetes *42*, 1041–1054.

Blaikie, P., Immanuel, D., Wu, J., Li, N., Yajnik, V., and Margolis, B. (1994). A region in Shc distinct from the SH2 domain can bind tyrosine-phosphorylated growth factor receptors. J. Biol. Chem. *269*, 32031–32034.

Bork, P., and Margolis, B. (1995). A phosphotyrosine interaction domain. Cell *80*, 693–694.

Brunger, A.T. (1992). X-PLOR version 3.1: a system for x-ray crystallography and NMR (New Haven, Connecticut: Yale University Press).

Carson, M. (1991). Ribbons 2.0. J. Appl. Crystallogr. *24*, 958–961.

Collaborative Computational Project Number 4 (1994). The CCP4 suite: Programs for protein crystallography. Acta Crystallogr. *D50*, 760–776.

Cowtan, K. (1994). "DM": an automated procedure for phase improvement by density modification. Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography *31*, 34–38.

Dikic, I., Batzer, A.G., Blaikie, P., Obermeier, A., Ullrich, A., Schlessinger, J., and Margolis, B. (1995). Shc binding to nerve growth factor receptor is mediated by phosphotyrosine interaction domain. J. Biol. Chem. *270*, 15125–15129.

Eck, M.J., Shoelson, S.E., and Harrison, S.C. (1993). Recognition of a high-affinity phosphotyrosyl peptide by the Src homology-2 domain of p56^{lck}. Nature *362*, 87–91.

Eck, M.J., Atwell, S.K., Shoelson, S.E., and Harrison, S.C. (1994). Crystal structure of the regulatory domains of the Src-family tyrosine kinase p56^{lck}. Nature *368*, 764–769.

Eck, M.J., Pluskey, S., Trüb, T., Harrison, S.C., and Shoelson, S.E. (1996). Spatial constraints on the recognition of phosphoproteins by the tandem SH2 domains of the phosphatase SH-PTP2. Nature *379*, 277–280.

Ferguson, K., Lemmon, M., Schlessinger, J., and Sigler, P. (1994). Crystal structure at 2.2 Å resolution of the pleckstrin homology domain from human dynamin. Cell *79*, 199–209.

Ferguson, K.M., Lemmon, M.A., Schlessinger, J., and Sigler, P.B. (1995). Structure of the high affinity complex of inositol triphosphate with a PLC pleckstrin homology domain. Cell *83*, 1037–1046.

Fiore, F., Zambrano, N., Minopoli, G., Donini, V., Duilio, A., and Russo, T. (1995). The regions of the Fe65 protein to the phosphtyrosine interaction/phosphotyrosine binding domain of Shc bind the intracellular domain of the Alzheimer's amyloid precursor protein. J. Biol. Chem. *270*, 30853–30856.

Gustafson, T.A., He, W., Craparo, A., Schaub, C.D., and O'Neill, T.J. (1995). Phosphotyrosine-dependent interaction of Shc and IRS-1 with the NPEY motif of the insulin receptor via a novel non-SH2 domain. Mol. Cell. Biol. *15*, 2500–2508.

Hatada, M.H., Lu, X., Laird, E.R., Green, J., Morgenstern, J.P., Lou, M., Marr, C.S., Phillips, T.B., Ram, M.K., Theriault, K., Zoller, M.J., and Karas, J.L. (1995). Molecular basis for the interaction of the protein tyrosine kinase ZAP-70 with the T-cell receptor. Nature *377*, 32–38.

He, W., O'Neill, T.J., and Gustafson, T.A. (1995). Distinct modes of interaction of SHC and insulin receptor substrate-1 with the insulin receptor NPEY region via non-SH2 domains. J. Biol. Chem. *270*, 23258–23262.

He, W., Craparo, A., Zhu, Y., O'Neill, T.J., Wang, L.M., Pierce, J.H., and Gustafson, T.A. (1996). Interaction of insulin receptor substrate-2 with the insulin and IGF-1 receptors: evidence for two distinct phosphotyrosine-dependent interaction domains within IRS-2. J. Biol. Chem. *271*, 11641–11645.

Hubbard, S.R., Wei, L., Ellis, L., and Hendrickson, W.A. (1994). Crystal structure of the tryosine kinase domain of the human insulin receptor. Nature *372*, 746–754.

Hyvönen, M., Macias, M.J., Nilges, M., Oschkinat, H., Saraste, M., and Wilmanns, M. (1995). Structure of the binding site for inositol phosphates in a PH domain. EMBO J. *14*, 4676–4685.

Jones, T.A., Bergdoll, M., and Kjeldgaard, M. (1989). Crystallographic Computing and Modeling Methods in Molecular Design, C. Bugg and S. Ealick, eds. (New York: Springer).

Kabsch, W.J. (1988). Evaluation of single crystal diffraction data from a position sensitive detector. J. Appl. Crystallogr. *21*, 916–924.

Kavanaugh, W.M., and Williams, L.T. (1994). An alternative to SH2 domains for binding tyrosine-phosphorylated proteins. Science *266*, 1862–1865.

Kavanaugh, W.M., Turck, C.W., and Williams, L.T. (1995). PTB domain binding to signaling proteins through a sequence motif containing phosphotyrosine. Science *268*, 1177–1179.

Lamzin, V.S., and Wilson, K.S. (1993). Acta Crystallogr. *D49*, 129–147.

Lee, C.H., Kominos, D., Jaques, S., Margolis, B., Schlessinger, J., Shoelson, S.E., and Kuriyan, J. (1994). Crystal structures of peptide complexes of the N-terminal SH2 domain of the Syp tyrosine phosphatase. Structure *2*, 423–438.

Musacchio, A., Gibson, T., Rice, P., Thompson, J., and Saraste, M. (1993). The PH domain: a common piece in a pathwork of signalling proteins. Trends Biochem. Sci. *18*, 343–348.

Myers, M.G., Grammer, T.C., Brooks, J., Glasheen, E.M., Wang, L.-M., Sun, X.J., Blenis, J., Pierce, J.H., and White, M.F. (1995). The pleckstrin homology domain in insulin receptor substrate-1 sensitizes to insulin signaling. J. Biol. Chem. *270*, 11715–11718.

Navaza, J. (1992). Molecular Replacement: Proceedings of the CCP4 Study Weekend, E.J. Dodson, S. Gover, and W. Wolf, eds. (Daresbury, UK: SERC), pp. 87–90.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

Nolte, R., Eck, M., Schlessinger, J., Shoelson, S., and Harrison, S. (1996). Crystal structure of the PI 3-kinase p85 amino-terminal SH2 domain and its phosphopeptide complexes. Nat. Struct. Biol. *3*, 364–373.

O'Neill, T.J., Craparo, A., and Gustafson, T.A. (1994). Characterization of an interaction between insulin receptor substrate 1 and the insulin receptor by using the two-hybrid system. Mol. Cell. Biol. *14*, 6433–6442.

Otwinowski, Z. (1991). Maximum likelihood refinement of heavy atom parameters. In Isomorphous Replacement and Anomalous Scattering, W. Wolf, P.R. Evans, and A.G.W. Leslie, eds. (Warrington, United Kingdom: Science and Engineering Council/Daresbury Laboratory), pp. 80–86

Piccione, E., Case, R.D., Domchek, S.M., Hu, P., Chaudhuri, M., Backer, J.M., Schlessinger, J., and Shoelson, S.E. (1993). PI 3-kinase p85 SH2 domains specificity defined by direct phosphopeptide/SH2 domain binding. Biochemistry *32*, 3197–3202.

Sawka-Verhelle, D., Tartare-Deckert, S., White, M.F., and Van Obberghen, E. (1996). IRS-2 binds to the insulin receptor through its PTB domain and through a newly identified domain comprising amino acids 591–786. J. Biol. Chem. *271*, 5980–5983.

Sun, X.J., Rothenberg, P., Kahn, C.R., Backer, J.M., Araki, E., Wilden, P.A., Cahill, D.A., Goldstein, B.J., and White, M.F. (1991). The structure of the insulin receptor substrate IRS-1 defines a unique signal transduction protein. Nature *352*, 73–77.

Sun, X.J., Crimmins, D.L., Myers, M.G., Jr., Miralpeix, M., and White, M.F. (1993). Pleiotropic insulin signals are engaged by multisite phosphorylation of IRS-1. Mol. Cell. Biol. *13*, 7418–7428.

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). Role of IRS-2 in insulin and cytokine signalling. Nature *377*, 173–177.

Terwilliger, T.C., and Eisenberg, D. (1988). Isomorphous replacement: effects of errors on the phase probability distribution. Acta Crystallogr. *A43*, 6–13.

Trüb, T., Choi, W.E., Wolf, G., Ottinger, E., Chen, Y., Weiss, M.A., and Shoelson, S.E. (1995). Specificity of the PTB domain of Shc for β turn–forming pentapeptide motifs amino-terminal to phosphotyrosine. J. Biol. Chem. *270*, 18205–18208.

van der Geer, P., Wiley, S., Lai, V.K.-M., Olivier, J.P., Gish, G.D., Stephens, R., Kaplan, D., Shoelson, S.E., and Pawson, T. (1995). A conserved amino-terminal Shc domain binds to phosphotyrosine motifs in activated growth factor receptors and phosphopeptides. Curr. Biol. *5*, 404–412.

Voliovitch, H., Schindler, D.G., Hadari, Y.R., Taylor, S.I., Accili, D., and Zick, Y. (1995). Tyrosine phosphorylation of insulin receptor substrate-1 in vivo depends upon the presence of its pleckstrin homology region. J. Biol. Chem. *270*, 18083–18087.

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). Crystal structure of the phosphotyrosine recognition domain SH2 of v-*src* complexed with tyrosine phosphorylated peptides. Nature *358*, 646–653

Waksman, G., Shoelson, S.E., Pant, N., Cowburn, D., and Kuriyan, J. (1993). Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. Cell *72*, 779–790.

White, M.F., Livingston, J.N., Backer, J.M., Lauris, V., Dull, T.J., Ullrich, A., and Kahn, C.R. (1988a). Mutation of the insulin receptor at tyrosine 960 inhibits signal transmission but does not affect tyrosine kinase activity. Cell *54*, 641–649.

White, M.F., Shoelson, S.E., Keutmann, H., and Kahn, C.R. (1988b). A cascade of tyrosine autophosphorylation in the β-subunit activates the phosphotransferase of the insulin receptor. J. Biol. Chem. *263*, 2969–2980.

Wolf, G., Trüb, T., Ottinger, E., Groninga, L., Lynch, A., White, M., Miyazaki, M., Lee, J., and Shoelson, S.E. (1995). The PTB domains of IRS-1 and Shc have distinct but overlapping binding specificities. J. Biol. Chem. *270*, 27407–27410.

Zhou, M.-M., Ravichandran, K.S., Olejniczak, E.T., Petros, A.M., Meadows, R.P., Harlan, J.E., Wade, W.S., Burakoff, S.J., and Fesik, S.W. (1995). Structure and ligand recognition of the phosphotyrosine binding domain of Shc. Nature *92*, 7784–7788.

Zhou, M.-M., Huang, B., Olejniczak, E.T., Meadows, R.P., Shuker, S.B., Miyazaki, M., Trüb, T., Shoelson, S.E., and Fesik, S.W. (1996). Structure basis for IL-4 receptor phosphopeptide recognition by the IRS-1 PTB domain. Nat. Struct. Biol. *3*, 388–393.