

Peptide–Surface Association: The Case of PDZ and PTB Domains

Minireview

Stephen C. Harrison
Howard Hughes Medical Institute
Children's Hospital
320 Longwood Avenue
Boston, Massachusetts 02115
Harvard University
7 Divinity Avenue
Cambridge, Massachusetts 02138

Regulated protein associations govern signal transduction pathways. Distinct, small domains of much larger proteins are often responsible for the interactions. In thinking about the molecular–structural basis for specificity, it is useful to distinguish “surface–surface association” from “peptide–surface association.” The former mode is familiar from protomer contacts in oligomeric enzymes: the apposition of complementary surfaces from two folded subunits. Each surface usually contains residues from more than one segment of polypeptide chain. Small G proteins probably recognize their effectors in this manner, as suggested by the structure of Rap1 complexed with a binding domain from Raf (Nassar et al., 1995). The peptide–surface mode of association is illustrated by complexes of SH2 domains with phosphotyrosine-containing proteins (Waksman et al., 1993; Eck et al., 1993). One partner, the SH2 domain, is a folded structure; the other, the phosphotyrosine-bearing segment, is a peptide-like loop or appendage, which docks onto the domain. Association involves a disorder-to-order transition for this peptide segment, an event formally similar to a step in protein folding. That is, the bound segment has a unique conformation, but the unbound segment has multiple conformations. Two further examples of peptide–surface recognition in signal transduction have recently been described in three-dimensional detail—the “PDZ domain” (named for three of the proteins containing it), which recognizes C-terminal -X-Ser/Thr-X-Val sequences (Doyle et al., 1996), and the “PTB” (phosphotyrosine binding) domain, which recognizes -Asn-Pro-X-pTyr- (Zhou et al., 1995; Eck et al., 1996). The PDZ and PTB domains turn out to have some surprising similarities in the way that they bind peptide segments from other proteins (Figure 1). The domains themselves are small β -barrels with related but distinct arrangements of strands. In both cases, the bound peptide forms an additional β -strand, which augments one of the sheets in the domain. The C-terminal part of the peptide interacts with residues in a shallow specificity pocket.

PDZ Domains

PDZ domains were first noticed as repeated sequences in the brain-specific protein PSD-95, the *Drosophila* septate junction protein Disks-large, and the epithelial tight-junction protein ZO1 (Cho et al., 1992; reviewed by Gomperts, 1996). In these proteins, there are three such domains, of about 100 residues each, followed in the sequence by a Src homology 3 (SH3) domain and a yeast guanlylate kinase homology domain. PDZ domains,

which have also been called “GLGF repeats” and “Disks-large binding repeats” (DHRs), appear in different contexts in other proteins. In PSD-95 and its relatives, they have been shown to bind the X-Ser/Thr-X-Val motif found as the C-terminal residues in certain ion-channel subunits, and clustering of channels appears to be one function of the concatenated domains (Kim et al., 1995; Kornau et al., 1995; Niethammer et al., 1996; Gomperts, 1996). MacKinnon and colleagues have recently determined crystal structures of the third PDZ domain from PSD-95, in the presence and absence of a bound peptide (Doyle et al., 1996). They show that four ordered peptide residues interact with the protein. Two of them form antiparallel β -sheet hydrogen bonds with at the edge of one sheet, thus adding an extra strand. The carboxylate of the terminal valine faces the “carboxylate binding loop,” which contains the GLGF sequence used to identify PDZ domains (Figure 1). The peptide amide groups of the two glycines and of the phenylalanine donate hydrogen bonds to the carboxylate, which is also linked through a water to the guanidinium group of an arginine. The valyl side chain fits into a small pocket, defined by four conserved hydrophobic residues, including the leucine and phenylalanine of GLGF. The extended peptide lies against an α -helix, which contributes to the hydrophobic pocket for the valine and to the recognition of the threonine two residues before it.

PTB Domains

PTB domains are regions of 100–150 residues in the insulin receptor substrates 1 and 2 (IRS-1 and IRS-2) and in the adaptor protein Shc (Kavanaugh and Williams, 1994). These proteins bind to autophosphorylation sites on the insulin and EGF receptors, respectively. The amino-acid sequences of the IRS-1 and IRS-2 PTB domains are closely related to each other, but they have no identifiable similarity to the sequence of the domain from Shc. Nonetheless, as shown by recent NMR (Zhou et al., 1995; Zhou et al., 1996) and crystallographic (Eck et al., 1996) studies, both the Shc and IRS-1 domains have essentially identical 7-stranded β -sandwich frameworks, capped by C-terminal helices. Moreover, the β -strands of their folded structures superimpose as well on the framework of a pleckstrin homology (PH) domain as various PH domains do on each other. Both Shc and IRS-1 PTB domains recognize peptides containing phosphotyrosine at the end of an NPXpY sequence (Wolf et al., 1995). In addition, the IRS-1 PTB domain requires a hydrophobic residue at pY-8; the Shc domain, a hydrophobic at pY-5. By contrast, PH domains (of which a number of structures have now been determined) are thought to interact with phosphoinositide head groups rather than with peptides or other folded proteins (Ferguson et al., 1995). As in the PDZ complexes, the peptides bound to Shc and IRS-1 PTB domains form antiparallel β -strand hydrogen bonds with the edge of a sheet (Figure 1). They also lie against the C-terminal α -helix. The NPXpY motif at the C-terminal end of the PTB-bound peptide is, of course, much more extensive than the simple carboxylate at the C-terminus of the PDZ-bound peptide, and the conformation of the motif is

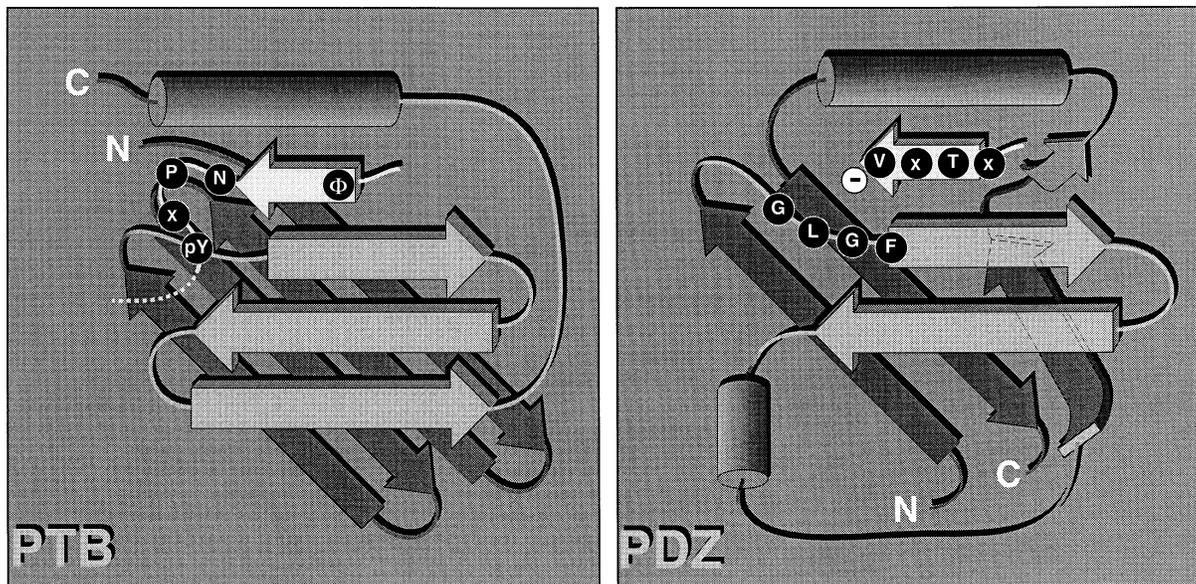


Figure 1. Peptide Binding by β -Augmentation of PTB and PDZ Domains

Diagrams are modified from Doyle et al. (1996). Arrows represent β -strands; cylinders, α -helices. The bound peptides are in white; the domains are shaded. The segment of peptide bound to the PTB domain adopts a strand-turn conformation. The phosphotyrosine (pY) is the last ordered residue but need not be the C-terminus of the bound peptide. The residue at pY-5 (Shc) or pY-8 (IRS-1) is hydrophobic (Φ); the turn begins with an obligatory asparagine (N) and generally contains proline (P) in an NPXpY sequence. The segment of peptide bound to the PDZ domain is a simple, C-terminal strand. The carboxylate (minus sign) of the terminal valine (V) is bound by the Gly-Leu-Gly-Phe (GLGF) loop; the threonine (T) is specified by interactions with residues in the α -helix.

such that the bound polypeptide could continue beyond the phosphotyrosine. In both cases, however, specificity results significantly from interactions with the loop preceding the strand with which the bound peptide has its β -sheet hydrogen bonds. In the PTB complexes, residues in this loop participate in an elaborate network of hydrogen bonds that anchor, quite elegantly, a β -turn formed by the NPXpY residues. An NPXY (not phosphorylated) sequence is found in the cytoplasmic tail of the LDL receptor, where it appears to serve as a signal for endocytosis by clathrin coated pits, and NMR studies have shown that it has some tendency to form β -turns on its own (Bansal and Gierasch, 1991). The same might also be true of this sequence with a phosphorylated tyrosine, but the extensive set of hydrogen bonds between the IRS-1 PTB domain and the NPXpY segment of the bound peptide suggest that the turn is in any case strongly stabilized by interaction with the domain.

Comparison of the IRS-1 (Eck et al., 1996; Zhou et al., 1996) and Shc (Zhou et al., 1995) PTB complexes shows that the β -strand and turn interactions are conserved, while those with the phosphotyrosyl side chain are not. That is, in the two structures, arginine side chains emanating from quite different parts of the domain interact with the pY residue. Thus, the essential common features of peptide recognition by the two PTB domains are augmentation of the β -sheet by the bound peptide and recognition (or imposition) of the β -turn. Differential specificity appears to depend on the presence of pockets for binding hydrophobic residues at pY-5 (Shc) or pY-8 (IRS-1). The IRS-1 PTB domain could not accommodate a bulky residue at pY-5, and this interference may prevent it from binding to growth factor receptors.

β -Sheet Augmentation and β -Clamps

The most striking feature that the PDZ and PTB complexes have in common is also the augmentation of a β -sheet by addition of an antiparallel peptide strand (Figure 1). Such β -augmentation is a noteworthy characteristic of assembly interactions in certain viral capsids. In SV40 and polyomavirus, for example, pentamers of the major capsid protein VP1 associate with each other through contacts made by the extended, peptide-like C-terminal arms of the VP1 polypeptide chains (Liddington et al., 1991). The core of the VP1 subunit is a jelly-roll β -sandwich, and in the assembled particle the C-terminal arm of a VP1 subunit from another pentamer augments one of the sheets by forming an anti-parallel β -strand at one edge. This strand is then clamped in place by addition of a further strand from the N-terminal arm of the target subunit itself. Both the "invading" C-terminal arm and the "clamping" N-terminal arm are disordered in unassembled VP1. A somewhat similar set of interactions stabilizes picornavirus capsids. The virus structures illustrate a property of β -sheets that is particularly suitable for assembly interactions: the free peptide groups at the edge of a sheet invite augmentation, and the complex can be stabilized if necessary by further β -strand additions. Could this clamping happen in PDZ or PTB domain complexes? For example, could another region of the PTB- or PDZ-containing protein add β -strands to the distal edge of the bound peptide segment? Or could the bound strand serve not only to associate the proteins (such as IRS-1 and the insulin receptor) but also to nucleate addition of yet another protein to the assembly? The structures themselves do not rule out such a possibility. That is, the domains do not have

elements that would interfere with further extension of the sheet—for example, by association of another peptide “tail” or even of the edge of another β -sandwich domain.

A peptide-surface association mechanism offers much greater flexibility than surface-surface association, since the way the peptide enters and leaves the region of contact can vary. Moreover, because the peptide is not part of an independently folded structure, its sequence is constrained only by the requirements of recognition by the appropriate domain. Thus, evolution can insert a short association sequence into a protein without otherwise affecting its activity. These properties may help to account for the widespread occurrence of peptide-surface association in signalling pathways. Related assembly motifs are likely to appear in other intracellular complexes, such as those involved in transcriptional initiation or membrane traffic.

Selected Reading

- Bansal, A., and Gierasch, L.M. (1991). *Cell* 67, 1195–1201.
- Cho, K.O., Hunt, C.A., and Kennedy, M.B. (1992). *Neuron* 9, 929–942.
- Doyle, D.A., Lee, A., Lewis, J., Kim, E., Sheng, M., and MacKinnon, R. (1996). *Cell* 85, 1067–1076.
- Eck, M.J., Pluskey, S., Trüb, T., and Harrison, S.C. (1996). *Nature* 379, 277–280.
- Eck, M.J., Shoelson, S.E., and Harrison, S.C. (1993). *Nature* 362, 87–91.
- Ferguson, K.M., Lemmon, M.A., Schlessinger, J., and Sigler, P.B. (1995). *Cell* 83, 1037–1046.
- Gomperts, S.A. (1996). *Cell* 84, 659–662.
- Kavanaugh, W.M., and Williams, L.T. (1994). *Science* 266, 1862–1865.
- Kim, E., Niethammer, M., Rothschild, A., Jan, Y.N., and Sheng, M. (1995). *Nature* 378, 85–88.
- Kornau, H.C., Schenker, L.T., Kennedy, M.B., and Seeburg, P.H. (1995). *Science* 269, 1737–1740.
- Liddington, R.C., Yan, Y., Zhao, H.C., Sahli, R., Benjamin, T. L., and Harrison, S.C. (1991). *Nature* 354, 278–284.
- Nassar, N., Gudrun, H., Herrmann, C., Scherer, A., McCormick, F., and Wittinghofer, A. (1995). *Nature* 375, 554–560.
- Niethammer, M., Kim, E., and Sheng, M. (1996). *J. Neurosci.* 16, 2157–2163.
- Waksman, G., Shoelson, S.E., Pant, N., Cowburn, D., and Kuriyan, J. (1993). *Cell* 72, 779–790.
- Wolf, G., Trüb, T., Ottinger, E., Groninga, L., Lynch, A., White, M., Miyazaki, M., Lee, J., and Shoelson, S.E. (1995). *J. Biol. Chem.* 270, 27407–27410.
- 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). *Nature Struct. Biol.* 3, 388–393.
- 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). *Nature* 92, 7784–7788.