

Phosphotyrosine signaling through modular protein-protein interactions

Tony Pawson

The dynamic organization of mammalian cells is controlled by a vast array of post-translational modifications, such as protein phosphorylation. External signals in the form of growth factors, cytokines, antigens and metabolic hormones, as well as intracellular cues such as DNA damage and passage through the cell cycle, exert their effects by activating protein kinases, that in turn phosphorylate specific targets. For example, growth factors commonly bind the extracellular regions of transmembrane receptors with cytoplasmic protein-tyrosine domains. Growth factor-binding induces receptor dimerization, and consequent intermolecular autophosphorylation of clustered receptor chains. Oncogenic mutations in the genes for receptor tyrosine kinases (RTK) typically induce constitutive receptor clustering and activation.

Phosphorylation can alter the activity of RTKs in two distinct ways. Phosphorylation of tyrosine residues in the activation loop of the kinase domain results in a conformational reorganization which relieves the inhibitory effect of the activation segment on the active site. However, tyrosine phosphorylation sites that lie in non-catalytic regions of the receptor (for example in the juxtamembrane region or in the C-terminal tail) provide docking sites for cytoplasmic targets with Src homology 2 (SH2) or phosphotyrosine (pTyr)-binding (PTB) domains. SH2 domains are sequences of approximately 100 amino acids, which possess a modular structure such that the N- and C-termini of an SH2 domain are closely juxtaposed in space. SH2 domains selectively recognize short phosphorylated motifs of 5-9 amino acids in proteins such as activated RTKs. SH2 domains have a conserved pTyr binding pocket, and more variable surfaces that engage the peptide residues N- or C-terminal to the phosphotyrosine. Different SH2 domains therefore bind preferentially to distinct tyrosine phosphorylated sites. As a result, the sequences flanking a receptor's autophosphorylation sites dictate the spectrum of intracellular targets that are recruited to the activated receptor, and the signaling pathways that are stimulated within the cell. The human proteome has 120 SH2 domains contained in 110 proteins. These polypeptides have diverse biochemical functions and

molecular architectures, and can be divided into a number of classes; 1) Proteins with intrinsic catalytic activity, 2) adaptor proteins composed exclusively of interaction domains, that couple RTKs to downstream targets, 3) transcription factors, 4) regulators that enhance or attenuate signaling by specific tyrosine kinases. In aggregate, these SH2 domain proteins directly regulate a variety of critical effectors and cellular processes, including Ras, Rho, Arf and Rab family GTPases and their targets, phospholipid-dependent signaling, transcriptional regulation, and cytoplasmic protein kinases and phosphatases. Signaling by RTKs, as well as multi-chain receptors such as antigen and cytokine receptors and integrins that utilize cytoplasmic tyrosine kinases, therefore depends on pTyr-dependent protein-protein interactions.

SH2 domains provide the prototype for a large number of different domain families that mediate specific molecular interactions. A number of these interaction domains recognize post-translational modifications, including protein phosphorylation (on serine/threonine as well as tyrosine), acetylation of lysine residues, methylation of lysine or arginine, prolyl hydroxylation, ubiquitination and sumoylation. Indeed a major function of post-translational modifications is to create binding sites for specific modular interaction domains, which are themselves usually components of larger multi-domain proteins. Other interaction domains recognize unmodified peptide motifs such as proline-rich sequences in the case of SH3 domains or C-terminal sites for PDZ domains. Interaction domains can also undergo heterotypic domain-domain interactions, in a fashion that depends on the folded structures of both domains (typified by death domains which control the activation of caspases involved in apoptosis). Furthermore, interaction domains can bind non-peptide ligands. For example several domain families, typified by PH domains, bind specific phosphoinositides, and thereby localize proteins to membranes enriched in a particular phospholipid. Interaction domains can also engage small molecules and metabolites (i.e. cAMP) or bind in a sequence-specific manner to nucleic acids (for example KH domains, a subset of SAM domains). Approximately 100 different classes of interaction domains can be found in human proteins, with each type of domain being present in many copies (10-300). Strikingly, approximately 70% of human proteins have a recognizable domain-based structure. Interaction domains therefore represent a

prominent feature of the proteome, and a dominant mechanism for organizing cellular regulatory circuits.

The modular construction of interaction domains and their binding motifs may have facilitated the evolution of signaling pathways important for multicellular animals. According to this scheme, a protein could acquire a new interaction domain (such as an SH2 domain) during the course of evolution, and thus gain a new function (such as coupling to pTyr signaling). This hypothesis is supported by a variety of data. First, it is possible to artificially create an entirely new signaling pathway by joining interaction domains together in a novel combination, for example by linking an SH2 domain to a domain that stimulates apoptosis, a death effector domain, and thus link an oncogenic RTK to the induction of cell death. Second, experimental data derived from mutagenesis studies suggest that the ancestral interaction sequences of the mouse ShcA docking protein are broadly required for its biological activity, whereas more recently acquired motifs have a more specialized role. Finally, chimeric oncoproteins, and the products of viral and bacterial pathogens, can mediate novel sets of protein-protein interactions that are important for their effects on cellular behaviour. For example, the oncogenic Bcr-Abl tyrosine kinase, contains an N-terminal sequence encoded by the Bcr gene, and a C-terminal sequence derived from the Abl cytoplasmic tyrosine kinase, but these elements are normally present on two entirely different proteins. In addition to the ability of the N-terminal region of Bcr to tetramerize, and thus promote Abl autophosphorylation, the Bcr sequence of Bcr-Abl becomes phosphorylated on Tyr-177, creating a binding site for the Grb2 adaptor. Tyrosine kinase oncoproteins can therefore induce malignant transformation by creating aberrant SH2 domain-mediated protein-protein interactions. Tyrosine kinase inhibitors, by suppressing the enzymatic activity of oncogenic tyrosine kinases, uncouple the ability of these pTyr-dependent complexes to drive malignant transformation. Furthermore, the SH2 and SH3 domains of cytoplasmic tyrosine kinases such as Abl and Src play a critical autoregulatory role which is exploited by inhibitors such as Gleevec to maintain the inactive conformation of the kinase domain.

References

1. Seet et al. 2006. Reading protein modifications with interaction domains. *Nature Rev Mol Cell Biol.* 7:473-483.
2. Liu et al. 2006. The human and mouse complement of SH2 domain proteins – establishing the boundaries of phosphotyrosine signaling. *Mol Cell* 22:851-868
3. Pawson and Nash. 2003. Assembly of cell regulatory systems through protein interaction domains. *Science* 300:445-452.