

# Assembly of Cell Regulatory Systems Through Protein Interaction Domains

Tony Pawson<sup>1,2\*</sup> and Piers Nash<sup>1</sup>

The sequencing of complete genomes provides a list that includes the proteins responsible for cellular regulation. However, this does not immediately reveal what these proteins do, nor how they are assembled into the molecular machines and functional networks that control cellular behavior. The regulation of many different cellular processes requires the use of protein interaction domains to direct the association of polypeptides with one another and with phospholipids, small molecules, or nucleic acids. The modular nature of these domains, and the flexibility of their binding properties, have likely facilitated the evolution of cellular pathways. Conversely, aberrant interactions can induce abnormal cellular behavior and disease. The fundamental properties of protein interaction domains are discussed in this review and in detailed reviews on individual domains at *Science's* STKE at <http://www.sciencemag.org/cgi/content/full/300/5618/445/DC1>.

**R**egulatory proteins are frequently constructed in a cassette-like fashion from domains that mediate molecular interactions or have enzymatic activity. Interaction domains can target proteins to a specific subcellular location, provide a means for recognition of protein posttranslational modifications or chemical second messengers, nucleate the formation of multiprotein signaling complexes, and control the conformation, activity, and substrate specificity of enzymes (Fig. 1) (*1*). In signal transduction, enzymes (kinases, for example) often generate modified amino acids on their substrates that are then recognized by interaction modules. Thus, phosphotyrosine (pTyr) sites formed by the actions of tyrosine kinases bind effectors with pTyr recognition domains [i.e., Src homology 2 (SH2) or pTyr-binding (PTB)] (*2*), whereas phosphoinositides produced by phosphoinositide kinases recruit pleckstrin homology (PH), Phox homology (PX), and FYVE domains, among others (*3*). In this sense, catalytic and interaction domains work hand-in-glove to control the dynamic state of the cell.

## Modular Interaction Domains

Isolated interaction domains can usually fold independently, with their N and C termini juxtaposed in space (Fig. 2A), and are readily incorporated into a larger polypeptide in a

manner that leaves their ligand-binding surface available. They recognize exposed sites on their protein partners—including phosphorylated, proline-rich, or C-terminal motifs—or they bind the charged head groups of phospholipids in membranes, with dissociation constants in the low nanomolar to high micromolar range. Typically, a protein interaction domain recognizes a core determinant, with flanking or noncontiguous residues providing additional contacts and an element of selectivity. For example, the SH3 domain of the cytoplasmic tyrosine kinase Csk recognizes a core PXXP motif (P, proline; X, any amino acid) in the tail of the PEP tyrosine phosphatase, which adopts a polyproline type II helix typical of SH3-binding sites. However, the Csk SH3 domain also contacts two more C-terminal hydrophobic residues in PEP through a separate binding pocket to yield a selective association of the two proteins *in vivo* (*4*) (Fig. 2A). In some cases, the affinity of a single domain for a peptide motif appears sufficient for a specific interaction in cells. In addition, tertiary interactions, the subcellular localization and structural organization of interacting proteins, domain competition, and multidomain interactions also likely contribute *in vivo* to selectivity in signaling.

Signaling domains can be identified through their consensus amino acid sequences, allowing the binding properties and biological functions of a protein to be predicted on the basis of domain composition (*5*). Several interaction domains are present in hundreds of copies in the human proteome, and these are used repeatedly to regulate distinct aspects of cellular organization. For example, about 115 SH2 domains and 253 SH3 do-

main are encoded by the human genome (*6*). Some domains serve specific functions; SH2 domains, for example, generally require phosphotyrosine sites in their primary ligands and are therefore dedicated to tyrosine kinase signaling (*7*). Other domains can bind motifs found in a broader set of proteins and display a wider range of biological activities; SH3 domains, for example, regulate processes such as signal transduction, protein and vesicle trafficking, cytoskeletal organization, cell polarization, and organelle biogenesis (*7, 8*). The cell therefore uses a limited set of interaction domains (Fig. 1), which are joined together in diverse combinations, to direct the actions of regulatory systems.

The preferred binding motifs for individual domains can be identified through probes of degenerate peptide libraries or peptide arrays, phage display analysis, and other techniques; this information can then be used to explore the proteome for candidate binding partners (*9–11*). Complexity, however, can be introduced by the ability of a particular domain class to recognize distinct motifs, by the presence of separate ligand-binding sites within an individual domain, and by the importance of ligand conformation in domain recognition. Furthermore, the optimal binding motif for a domain is not necessarily the one best suited to a physiological interaction or to *in vivo* specificity (*12, 13*). Predictive data must therefore be validated by direct analysis of protein complexes from cells.

## Flexible Binding Properties of Interaction Domains

Interaction domains are remarkably versatile in their binding properties. An individual domain can engage several distinct ligands, either simultaneously or at successive stages of signaling. Type I receptor serine kinases (RSKs) for transforming growth factor- $\beta$  (TGF $\beta$ ) signal through R-SMAD signal transducer proteins, which contain an N-terminal DNA binding (MH1) domain and a C-terminal MH2 domain with diverse binding partners, including proteins with pSer-X-pSer motifs (where pSer is phosphoserine). The R-SMAD MH2 domain interacts with a scaffolding protein, SARA, that also associates with the receptor and therefore juxtaposes the receptor and its R-SMAD target (*14*). After TGF $\beta$  stimulation, the R-SMAD MH2 domain likely recognizes a phosphorylated Gly-

<sup>1</sup>Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada. <sup>2</sup>Department of Medical Genetics and Microbiology, University of Toronto, Toronto, Ontario M5S 1A8, Canada.

\*To whom correspondence should be addressed. E-mail: pawson@mshri.on.ca

and Ser-rich juxtamembrane region on the activated receptor, leading to phosphorylation of the R-SMAD within a C-terminal Ser-X-Ser motif (15). Phospho-R-SMAD then dissociates from the receptor and is recognized through both its pSer-X-pSer motif and additional contacts by the MH2 domain of other SMAD molecules, with which it forms an oligomeric complex that is retained in the nucleus (16). Once in the nucleus, the SMAD MH2 domain interacts with components of the transcriptional apparatus to stimulate or repress gene expression (17). In the course of these perambulations from the plasma membrane to the nucleus, the SMAD MH2 domain uses a large interaction surface to recruit multiple different binding partners.

In addition to their conventional phosphopeptide-binding site, some SH2 domains have a binding surface that can engage SH3 domains, and they can therefore act as mini-adaptors to link tyrosine-phosphorylated and SH3-containing proteins. As an example, the human SAP protein (also called SH2D1A and DSHP), which is composed almost entirely of a single SH2 domain, uses the SH2 phosphopeptide binding surface to engage a pTyr-based motif in the SLAM receptor of T cells; SAP also recognizes the SH3 domain of the Fyn tyrosine kinase through a distinct basic surface centered on Arg<sup>78</sup> of the SH2 domain. These interactions juxtapose Fyn and SLAM and stimulate SLAM phosphorylation, which recruits other SH2 proteins to phospho-SLAM. These latter proteins then regulate lymphoid cell responses to viral infection (18, 19) (Fig. 3A). In a similar fashion, the Crk SH2 domain can bind simultaneously to the SH3 domain of the cytoplasmic tyrosine kinase Abl through a unique proline-rich insert in a loop region, and to a pTyr peptide (20).

Different members of the same domain type can bind quite different motifs. For example, SH3 domains usually recognize core PXXP sequences, but a subset of SH3 domains (such as the SH3 C-terminal domain of the adaptor Gads in T cells) use the same binding surface to engage an RXXK motif (R, Arg; K, Lys) (21). Indeed, the same domain fold can be put to a variety of uses. PH, PTB, and EVH1 (Ena-Vasp homology 1) domains, subdomains of the FERM (band 4.1 protein and ERM homology) and BEACH (beige and CHS) modules, and a binding protein for the Ran guanosine triphosphatase (GTPase) RanBP2 have the same fold but engage a wide range of peptide, phosphopep-

ptide, and phospholipid ligands through distinct binding surfaces (22–24). The PH domain fold is therefore a malleable scaffold that can accommodate a wide variety of binding partners and has likely been selected in the course of evolution for its adaptive nature.

**Interaction Domains Assembled from Repeated Motifs**

A further means of building interaction surfaces is through the joining of repeated (up to ~50) copies of a small peptide motif, yielding a much larger structure with multifaceted binding properties. Such repeats include HEAT (huntingtin, elongation factor 3, A subunit of protein phosphatase 2A, and TOR1), TPR (tetratricopeptide), Arm (Armadillo), ankyrin, leucine-rich, and Pumilio-homology sequences, and the resulting tandem repeat proteins have a wide array of biological activities (25).

Typically, each Arm repeat is composed of three helices, with the second and third helices packing in an antiparallel fashion, similar to the helix packing in HEAT and TPR repeats. As a testament to the diverse interactions mediated by helical repeat proteins,  $\beta$ -catenin functions in multiple cellular compartments to regulate adhesion, Wnt signaling, and gene expression.  $\beta$ -Catenin has a central region composed of 12 Arm repeats that form a superhelix with an extended positively charged groove spanning repeats 1 to 10. This groove interacts variously with the cytoplasmic tail of E-cadherin at adherens

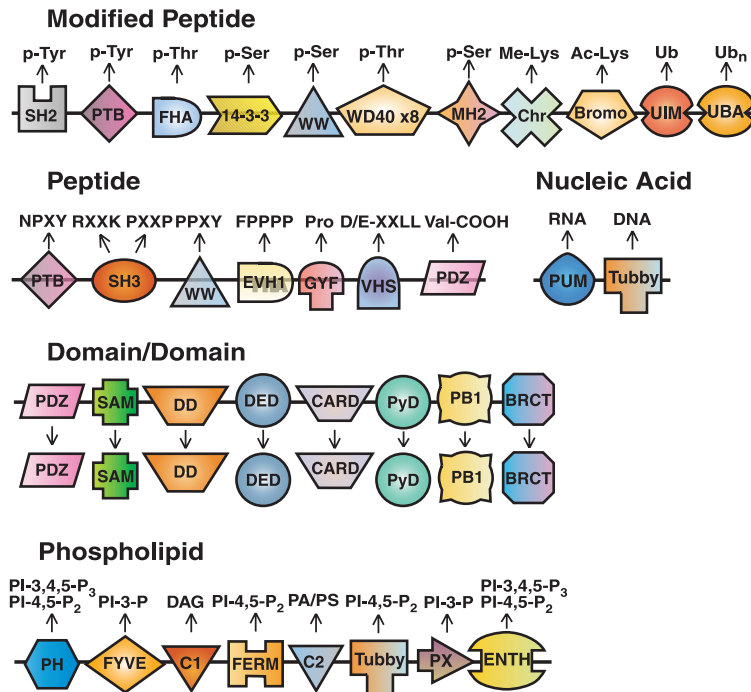
junctions, the cytoplasmic inhibitory protein APC (adenomatous polyposis coli), Tcf transcription factors in the nucleus, and the ICAT (inhibitor of  $\beta$ -catenin and Tcf) regulator that competes for Tcf binding (26–29). Each of these binding partners has a related core motif that binds as an extended strand within the recognition groove, whereas flanking sequences make somewhat distinct contacts with the Arm domain.

The flexible binding properties of repeat proteins are emphasized by *Drosophila* Pumilio, which binds the Nanos response elements in the 3' untranslated region (3'UTR) of *hunchback* RNA and recruits the Nanos and Brain Tumor proteins into a complex that suppresses *hunchback* mRNA translation. Pumilio has eight repeats of a ~36-residue trihelical motif, which assemble into an arc-shaped structure (30, 31) (Fig. 2C). This Pumilio homology domain has a concave surface that binds selectively to RNA motifs with a core UGU sequence (32). In effect, one base interacts with each Pum repeat through stacking interactions, with sequence-specific contacts provided by residues 12 and 16 of each repeat. This is somewhat reminiscent of the mechanism through which interaction domains bind peptide motifs; indeed, substitution of specific residues can modify the binding selectivity of the Pumilio homology domain for RNA, in much the same way that mutations in SH2 domains can alter their recognition properties for peptide motifs

(33). The binding sites for the Nanos and Brain Tumor proteins, which repress translation, are located on the outer surface of repeats 7 and 8 of the Pumilio homology domain. Interaction domains can therefore function as scaffolds for complexes that control basic cellular processes such as translation, and can recognize nucleic acids (34) in much the same way that they bind proteins.

**Domain-Domain Interactions**

A number of modular domains undergo homo- or heterotypic domain-domain interactions (Fig. 1) rather than binding short peptide motifs. Such domains frequently identify proteins involved in a common signaling process and then direct their coassembly into functional oligomeric complexes. Components of apoptotic or inflammatory signaling



**Fig. 1.** The building blocks—modular interaction domains in signal transduction. Interaction domains bind proteins, phospholipids, or nucleic acids. A subset of such domains is illustrated and their general binding functions are indicated. For more information, see [www.mshri.on.ca/pawson/research](http://www.mshri.on.ca/pawson/research) and <http://smart.embl-heidelberg.de>.

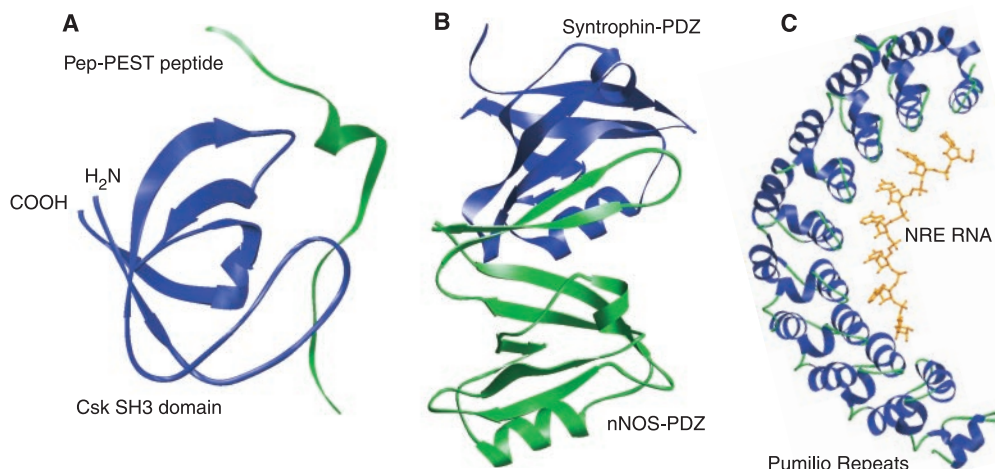
pathways are characterized by death domains or close structural relatives thereof [death effector domains, CARD (caspase recruitment) domains, Pyrin domains] that form heteromeric structures required for caspase dimerization and activation (35, 36) (Fig. 3B). In a variation on this theme, the SAM (sterile  $\alpha$  motif) domain of the *Drosophila* Polycomb group protein Polyhomeotic self-associates to form a helical polymeric structure, which appears important for maintaining chromatin in a repressed state during development (37).

The distinction between domains that bind peptide motifs and those that interact with other folded domain structures is by no means absolute. PDZ domains, for example, generally recognize short peptide motifs of  $\sim 4$  residues at the extreme C termini of their binding partners, but they can also mediate specific heterotypic PDZ-PDZ domain interactions (38–40) (Fig. 2B).

### Interaction Domains As Detectors of Posttranslational Modifications

Protein modifications frequently complete binding sites for interaction domains. Such domains must achieve a balance between inducibility and specificity, because much of their binding energy comes from recognition of the modified residue. SH2 domains have a conserved pTyr-binding pocket, and they also recognize residues C-terminal to the pTyr in a fashion that differs from one SH2 domain to another (41). However, the ability of most SH2 domains to discriminate between different phosphorylated sites is by no means absolute, given the limited interaction surface available to provide selectivity while maintaining pTyr-dependent binding (42). Indeed, the SAP SH2 domain (Fig. 3A), which is mutated in human X-linked lymphoproliferative syndrome, has an extended phosphopeptide-binding surface owing to its recognition of a threonine at the  $-2$  position (relative to the Tyr), and as a consequence binds an unphosphorylated peptide from the SLAM receptor with a dissociation constant of  $\sim 500$  nM. In this case, specificity is enhanced but recognition is not entirely dependent on phosphorylation of the binding target (43, 44).

The properties displayed by pTyr-binding domains are mirrored by modules that recognize pSer and pThr (45) (Fig. 1). FHA (Forkhead associated) domains, which have the same fold as SMAD MH2 domains, have a binding site for pThr and for more C-terminal residues on the target, notably the amino acid at the  $+3$  position (46). The N-terminal FHA domain of the yeast protein kinase Rad53, which is involved in DNA damage repair,



**Fig. 2.** Structural basis for three modes of modular protein interaction domain function. (A) Domain-peptide binding. The SH3 domain of Csk (blue) is shown bound to an extended peptide ligand derived from the C-terminal tail of the tyrosine phosphatase PEP containing a core PEST peptide motif (green) (4). (B) Domain-domain interaction. A dimer of the PDZ domains of syntrophin (blue) and neuronal nitric oxide synthase (nNOS) (green) is shown in a head-to-tail arrangement with a  $\beta$ -hairpin finger of nNOS docking into the peptide binding groove of syntrophin (40). (C) Repeat domains forming an extended binding face for nucleic acid. The helical repeats of human Pumilio1 (blue) are shown bound to a Nanos response element (NRE) in the 3'UTR of *hunchback* mRNA (orange) (32).

preferentially recognizes pTXXD motifs, whereas the FHA domain of the human protein kinase Chk2 binds pTXXI sequences (47). In yeast, activation of the protein kinase Mec1 after DNA damage induces multisite phosphorylation of the Rad9 protein. Phosphorylated Rad9 binds the FHA domains of the kinase Rad53, leading to Rad53 activation and consequent inhibition of mitotic exit and expression of genes in the DNA damage regulon (48). Similarly, short phosphopeptide motifs are recognized by 14-3-3 proteins (49), the Pin1 WW domain (50), or selected WD40 repeat domains (51).

The principles established for phosphorylation-dependent interactions have recently been extended to other forms of posttranslational modifications, because hydroxylation, acetylation, methylation, and ubiquitination of proteins can all function like phosphorylation to control modular protein interactions (Fig. 1). A case in point involves inducible binding of hydroxylproline (Hyp)-based peptide motifs in the transcription factor HIF1 $\alpha$  to the tumor suppressor VHL, the substrate binding component of an E3 protein ubiquitin ligase complex (52, 53). When cells are exposed to normal concentrations of oxygen, hydroxylation of HIF1 $\alpha$  on Pro<sup>402</sup> and Pro<sup>564</sup> nucleates a network of hydrogen bonds between HIF1 $\alpha$  and its binding partner VHL that increases binding affinity by three orders of magnitude, resulting in ubiquitination and degradation of HIF1 $\alpha$  (54, 55). As the oxygen tension falls, the hydroxylation of HIF1 $\alpha$  Pro<sup>402</sup> and Pro<sup>564</sup> (and thus binding to VHL) is lost, the protein is stabilized, and HIF1-regulated proangiogenic genes are expressed.

The acetylation and methylation of specific lysine residues in histones is important for

the organization of chromatin, and thus for the epigenetic regulation of gene expression (56). Acetylation of lysine residues in histones creates binding sites for bromodomains, which are embedded within proteins that induce an open chromatin state, such as histone acetyltransferases (HATs; for example, yeast Gcn5). Thus, histone acetylation induces further histone acetylation and perpetuates the altered configuration of chromatin. This is similar to the processive phosphorylation facilitated by the binding of SH2 domains within cytoplasmic tyrosine kinases to pTyr sites in substrates (57).

Acetylated peptides bind bromodomains as an extended strand, with the acetyl-lysine side chain protruding into a hydrophobic cavity (58, 59). Residues N- and C-terminal to the acetyl-lysine also interact with the bromodomain, as in the interaction between the HAT p300-CBP-associated factor (PCAF) and acetylated Tat protein from human immunodeficiency virus-1 (HIV-1), where Tat residues at the  $-3$ ,  $+3$ , and  $+4$  positions relative to acetylated Lys<sup>50</sup> engage the bromodomain (60).

A subset of chromodomains, in contrast, recognize methylated lysine motifs in histones and have been implicated in both gene silencing and activation (61). Analysis of the heterochromatin protein 1 (HP1) chromodomain bound to a histone H3 peptide with Lys<sup>9</sup> in the di- or trimethylated state shows that methyl-peptide binding induces a conformational change in the chromodomain. The resulting structure in which the peptide completes a three-stranded  $\beta$  sheet (62, 63) is reminiscent of peptide binding by PTB and PDZ domains.

REVIEW

The covalent modification of proteins is not limited to relatively small groups such as phosphate, but can involve the addition of large peptides such as ubiquitin. A conserved 20-amino acid motif (the ubiquitin interaction motif, UIM) found in a number of endocytic adaptor proteins, such as mammalian Eps15, epsin, and Hgs, recognizes ubiquitinated sites during the sorting process and can promote monoubiquitination (64). It appears that an important role of ubiquitination is to create binding sites for UIM domains, leading to protein sorting and signaling.

Modular Signaling Systems

Interaction domains mediate the association of cell surface receptors with their targets, as well as the formation of signaling complexes in the cytoplasm and nucleus. Similar domains function to regulate targeted proteolysis, endocytosis, vesicle and protein trafficking, cell polarity, cytoskeletal organization, and gene expression. These different regulatory systems therefore use

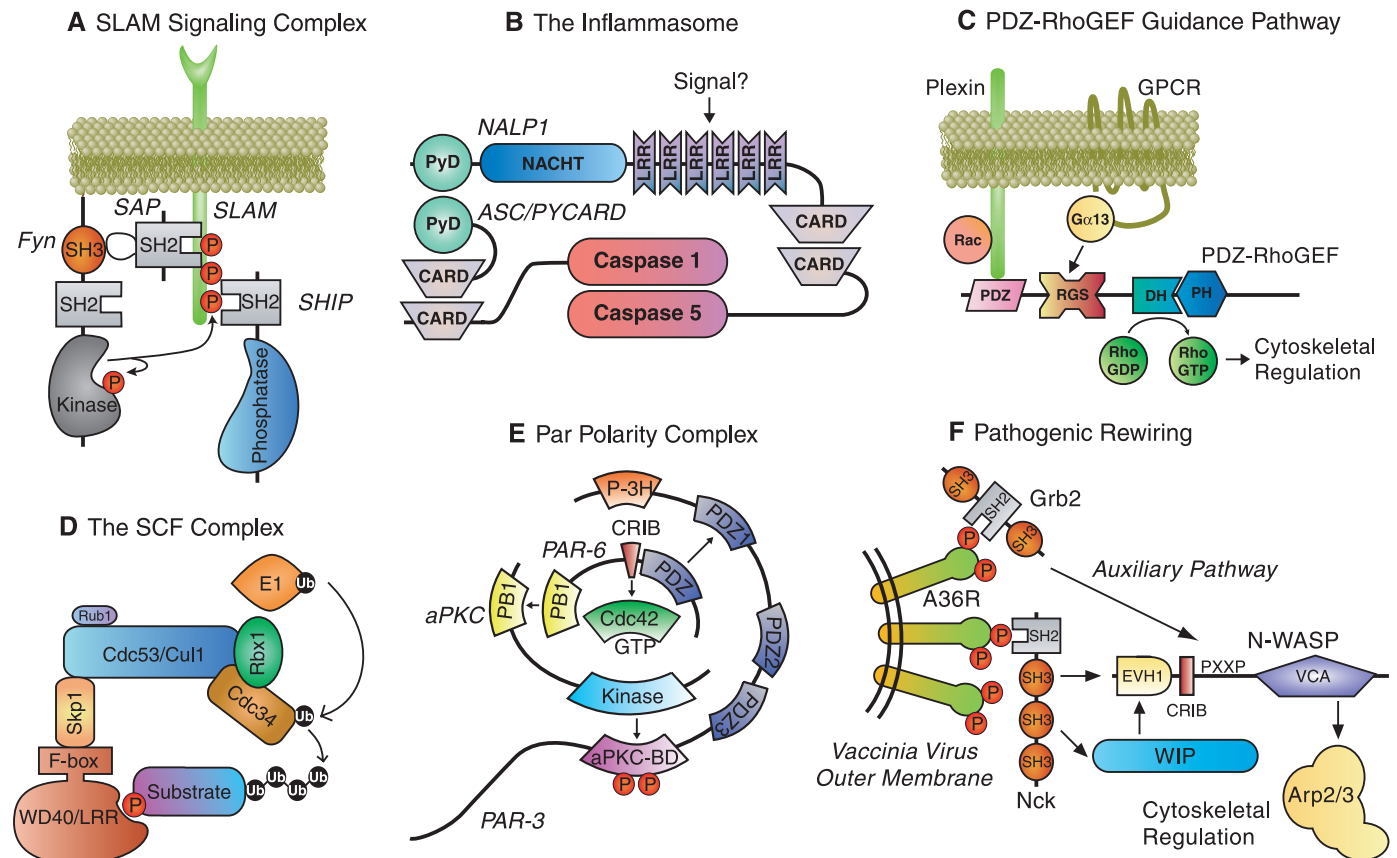
common strategies to assemble functional complexes, to compartmentalize molecular components, and to direct enzymes to their targets.

*Phosphorylation-dependent and -independent signaling from receptors.* Receptor tyrosine kinases (RTKs) and TGFβ RSKs both stimulate phosphorylation-dependent interactions that propagate specific signals, mediated by SH2 domains in the case of RTKs and by MH2 domains for RSKs. In addition, RTKs of the Eph family and type I TGFβ receptors apparently undergo a similar conformational change that coordinately allows activation of the kinase domain and the exposure of phosphorylated binding sites for cytoplasmic targets (15, 65).

Many other receptors lack kinase activity but nonetheless use interaction domains to recruit their cytoplasmic targets. Members of the tumor necrosis factor receptor (TNFR) family, such as Fas and TNFR1, have cytoplasmic death domains that heterodimerize with the death domains of adaptor and scaffolding proteins (66).

Other TNFR family members and the adaptor TRADD contain short peptide motifs that bind the C-terminal domains of TRAF (TNFR-associated factor) proteins (67–70), and Toll receptors have a cytoplasmic TIR (Toll-interleukin-1 receptor homology) domain that recruits the TIR domain of the adaptor MyD88 (71). The combined use of interaction modules (such as death, TRAF, and TIR domains) therefore couples receptors involved in innate immunity and cell death to intracellular targets, in much the same way that SH2 and MH2 domains link RTKs and RSKs to phosphorylation-dependent pathways.

A similar theme emerges from receptors that control cell movement and axon guidance, which bind the interaction domains of proteins that regulate the cytoskeleton. In *Drosophila*, Robo proteins are receptors that mediate the repulsive effects of the Slit protein, which acts as a guidance cue to control the movement of axons away from the midline and into specific longitudinally migrating tracts. The cytoplasmic regions of Robo receptors have conserved



**Fig. 3.** Reiterated use of interaction domains to build complex machines in signaling. (A) A signaling complex formed at the SLAM receptor in T cells. SHIP, SH2 inositol phosphatase. See text for details. (B) Activation of NALP1 results in the formation of the inflammasome complex that brings procaspases into close association and results in caspase activation (36). (C) Rho-GEFs of the PDZ-RhoGEF and LARG family use PDZ and RGS domains to integrate signals from plexin-B1 and G protein-coupled receptors (GPCRs). (D) An SCF E3 ubiquitin ligase complex. The SCF complex is composed of Skp1, a cullin protein (Cdc53), an E2 (Cdc34), an Rbx1 RING finger protein, and an F-box substrate adapter protein such as Cdc4, Skp2, or

β-TrCP (81). The substrate-binding region of the F-box protein recruits targets for ubiquitination; this interaction frequently requires phosphorylation of the target protein. (E) The PAR-3/PAR-6 complex in cellular polarity. A network of interactions creates a complex among PAR3, PAR6, aPKC, and Cdc42 required for the establishment and maintenance of cellular polarity. (F) Pathogens can coopt existing cellular machinery and “rewire” cellular signaling. The vaccinia virus A36R protein recruits Nck to the intracellular enveloped virus particle and thereby reorganizes the cytoskeleton (132). Additional detail on the modular protein interaction domains shown can be found at [www.mshri.on.ca/pawson/research](http://www.mshri.on.ca/pawson/research) and at *Science’s* STKE.

proline-rich motifs that bind the SH3 domains of proteins such as the Abl tyrosine kinase and srGAP1 (a Cdc42-selective GTPase-activating protein that antagonizes axonal outgrowth) (72–74), or the EVH1 domain of Ena, a modular protein that stimulates polymerization at the barbed ends of actin filaments (73, 75).

The receptor plexin-B1 mediates the repulsive effect of semaphorin 4D on axonal growth cones by inducing formation of actin stress fibers. The C terminus of plexin-B1 binds the PDZ domains of PDZ-RhoGEF or LARG (leukemia-associated RhoGEF), two closely related guanine nucleotide exchange factors (GEFs) that activate the Rho GTPase, which in turn promotes formation of actin stress fibers (76–78). This PDZ-binding motif is important for growth cone collapse induced by activated plexin-B1. Interestingly, these GEFs, together with their close relative p115-RhoGEF, also possess an RGS (regulator of G protein signaling) domain that binds to the G protein (heterotrimeric guanine nucleotide-binding protein)  $\alpha$  subunits  $G\alpha_{12}$  and  $G\alpha_{13}$ , which themselves are regulated by G protein-coupled receptors. Thus, in addition to providing a component of plexin-B1 signaling, these RhoGEFs are also effectors through which  $G\alpha$  subunits can regulate the cytoskeleton, and may use their PDZ and RGS domains to integrate signals from distinct classes of receptors (Fig. 3C).

**Ubiquitination, targeted proteolysis, and endocytosis.** Ubiquitin is passed from the E1 ubiquitin ligase to an E2 enzyme, which is then recruited into an E3 complex that binds the ubiquitination target. The Hect class of E3 proteins have a ubiquitin ligase domain, as well as interaction domains that bind the substrate. The Nedd4 and Smurf E3 families, as an example, have three or four N-terminal WW domains that bind proline-rich (PPXY) motifs in their targets (79). In this manner, Smurfs bind the inhibitory Smad7 protein, which in turn is recruited to the activated type I TGF $\beta$  RSK; these interactions target Smurf to ubiquitinate the receptor, which induces receptor and Smad7 degradation through proteasomal and lysosomal pathways (80).

A second class of E3 proteins possess RING fingers and can act as adaptors to recruit an E2 ubiquitin ligase into a larger structure, such as a so-called SCF (SKP1/cullin/F-box protein) complex (81). The substrate-binding subunits of SCF complexes contain an F-box, through which they are attached to the E3 ligase complex, and a C-terminal domain, often composed of WD40 or leucine-rich repeats, that binds the substrate for ubiquitination (51, 82) (Fig. 3D). This latter interaction can require phosphorylation of the substrate on Ser or Thr residues, as in binding of the I $\kappa$ B inhibitory subunit of the NF $\kappa$ B transcription factor or  $\beta$ -catenin to the  $\beta$ TrCP F-box protein component of an SCF ubiquitin ligase complex (83). In this way, proteins are targeted for ubiquitination by pSer- and

pThr-dependent interactions, thereby controlling signaling pathways or passage through the cell division cycle.

Tyrosine phosphorylation is linked to ubiquitination by the Cbl E3 ubiquitin ligase, which contains both SH2 and RING finger domains. The N-terminal SH2 domain binds specific pTyr sites on activated receptors, and is followed by a RING domain that recruits an E2 enzyme and thus induces RTK ubiquitination (84, 85). The ubiquitinated receptor can then be recognized by endocytic adaptor proteins with UIM domains, leading to its internalization. c-Cbl also binds the SH3 domain-containing protein CIN85, which in turn recruits both the SH3 domain of endophilin, a constituent of clathrin-coated vesicles with the potential to induce membrane invagination, and the  $\alpha$  subunit of the AP-2 clathrin adaptor (86, 87). In addition, CIN85 can be monoubiquitinated and may therefore recruit UIM-containing proteins. The RTK-Cbl-CIN85 complex can therefore establish a network of SH2-, SH3-, and ubiquitin-based interactions that control receptor endocytosis (88).

Interaction domains are also important in trafficking events at sites other than the plasma membrane. Three GGA (Golgi-localized  $\gamma$ -ear-containing adenosine diphosphate ribosylation factor-binding) proteins mediate the anterograde transport of the cation-independent mannose 6-phosphate receptor (MPR), and associated lysosomal enzymes bearing the mannose 6-phosphate marker, from the trans-Golgi network to endosomes, where the enzymes are released and transferred to lysosomes (89, 90). GGA proteins have an N-terminal VHS domain that binds specifically to an acidic-dileucine motif, DDS<sup>0</sup>EDLLH (D, Asp; S, Ser, E, Glu; L, Leu; H, His; D<sup>0</sup>, reference residue in the DXXLL motif) in the cytoplasmic tail of the MPR, enabling proper lysosomal enzyme sorting. The VHS domain forms a right-handed superhelix that binds the sorting motifs of the MPR as an extended strand, most notably binding the Asp at position 0 through a positively charged pocket and the +3 and +4 Leu residues through shallow hydrophobic pockets (91, 92). GGAs interact with other proteins, including the coat protein adaptor AP-1, through which they help sort the MPR into clathrin-coated vesicles. GGA1 and GGA3 have an internal motif (S-X-X-D-D/E-E-L-L/M, where M = Met) that binds the VHS domain, likely through an intramolecular interaction, to suppress recognition of the MPR. This autoinhibitory interaction with the VHS domain is dependent on phosphorylation of the -3 Ser by casein kinase 2 (93). Although the details are very different, there is a conceptual similarity between this phosphorylation-dependent mode of autoregulation proposed for GGA1 and GGA3, which may control binding to cargo, and the well-established autoinhibition of Src family

tyrosine kinases through the intramolecular association of the phosphorylated C-terminal tail with the SH2 domain (94).

**Cell polarity.** Polarization is crucial for the functions of epithelial and neuronal cells, which display distinct apical and basal surfaces, or synaptic and somatodendritic structures. Asymmetric cell division, a related process through which a cell distributes molecular determinants unequally to its two daughters, is essential for development of tissues (95). These events require the coordination of vesicle and protein trafficking, the formation of cell junctions, cytoskeletal organization, polarization of microtubules, and orientation of the mitotic spindle. A group of conserved proteins composed of interaction domains form a network through which external and intrinsic polarity cues are interpreted (96). The Par-3 and Par-6 proteins, originally identified for their roles in the asymmetric divisions of the *Caenorhabditis elegans* one- and two-cell embryos, anchor a conserved multiprotein complex with numerous functions in metazoans (97) (Fig. 3E). Par-6 has an N-terminal PB1 (Phox and Bem1) domain that heterodimerizes with the PB1 domain of atypical protein kinase C (aPKC  $\lambda$  and  $\zeta$ ), and a central CRIB (Cdc42-Rac interactive binding) motif that associates selectively with GTP-bound Cdc42, followed by a PDZ domain (98–100). The PDZ domain of Par-6 can itself associate with the first of three PDZ domains of Par-3 (known in flies as Bazooka). In this fashion, the combined use of interaction and catalytic domains assembles a complex that can be positioned at specific subcellular sites through PDZ-mediated interactions (for example, with the junctional protein JAM1) (101), receive signaling inputs through the Cdc42 GTPase, and transmit polarity signals through aPKC. In epithelial cells, a series of PDZ-based complexes establish and maintain apical-basal polarity, and these show both genetic and physical interactions suggestive of a larger network (102–106).

### Building Pathways and Networks

It is straightforward to envision how the successive use of interaction domains can form linear signaling pathways, as in the case of the Grb2 SH2-SH3 adaptor, which links pYXN motifs on RTKs to PXXP sites on Sos, a GEF for the Ras GTPase that stimulates the MAP (mitogen-activated protein) kinase pathway. However, such interactions can potentially generate more complex networks that may allow for a robust cellular response, generate crosstalk between pathways, and integrate signals from distinct receptors (107, 108). For example, Grb2 also binds through its C-terminal SH3 domain to an RXXX motif in the docking protein Gab1, which is consequently phosphorylated on tyrosine, creating binding sites for the SH2 domains of cytoplasmic signaling proteins such as the p85 subunit of phosphatidylinositol 3-kinase (PI3K), the tyrosine phospho-

## REVIEW

tase Shp2, and the adaptor Crk (109). PI3K, in turn, elicits a series of phospholipid- and pSer and pThr-dependent modular interactions that control cell survival and proliferation. Such data correspond with genetic arguments indicating that proteins such as Grb2 can have multiple distinct functions in embryos and the adult. One way to generate complexity may be to reuse the same adaptor proteins but in different cellular contexts, and potentially with distinct effectors.

Similar arguments can be made for other forms of regulation. Activation of the interferon  $\beta$  promoter in response to viral infection involves a succession of acetyl-Lys-bromodomain interactions, through which the HAT GCN5, the SWI-SNF remodeling complex, and the TFIID transcription complex are recruited to the promoter (110). This can perhaps be viewed as the equivalent of a linear signaling pathway. However, chromatin structure is controlled by a sophisticated interplay of histone modifications, including Lys acetylation, Lys and Arg methylation, ubiquitination, and Ser phosphorylation, with the potential to generate network properties similar to those proposed for RTK signaling (111, 112).

At a further level of complexity, the postsynaptic density (PSD) of neuronal synapses, which controls the dendritic response to neurotransmission, contains a supramolecular organization of interacting proteins, dominated by polypeptides with PDZ domains. These polypeptides ensure the appropriate trafficking and localization of glutamate receptors, and they control the activation of signaling proteins involved in synaptic responses (113, 114). For example, the PDZ protein PSD-95 acts as a nexus in the PSD through its interactions with ion channels [such as *N*-methyl-D-aspartate (NMDA) neurotransmitter receptors], signaling molecules including SynGAP and neuronal nitric oxide synthase, and docking proteins. Genetic inactivation of PSD-95 in the mouse affects synaptic function and spatial learning (115), and blocking the PDZ-mediated interaction between NMDA receptors and PSD-95 decreases ischemic brain damage in a rat model of stroke (116).

### Proteomic Analysis of Signaling Networks

By analyzing the *in vitro* binding specificities of interaction domains, and by directly analyzing protein complexes by techniques such as mass spectrometry (MS) and yeast two-hybrid analysis, it is in principle possible to assemble a wiring circuitry for cellular protein interactions. Recent advances in the use of MS to identify phosphorylated sites also raise the possibility of comprehensively following posttranslational modifications (117, 118), which can then be linked to the dynamic assembly of protein complexes. A start has

been made with the yeast *Saccharomyces cerevisiae*, which contains 28 SH3 proteins whose binding properties have been analyzed by both phage display and yeast two-hybrid techniques (10). Combining the results of these techniques has identified an SH3-mediated interaction network, containing complexes whose importance can be tested by both biochemical and genetic means. The relevant proteins are not equally connected; rather, the network is focused around a core of hub proteins (centered on the actin regulator and WAVE/WASP ortholog Las17), each of which makes at least six connections, with the remaining protein nodes being less highly connected. The biological relevance of such a scale-free network remains to be fully explored, although it has been argued that it would be relatively tolerant to loss of all but the most highly connected nodes (119). Protein complexes have also been analyzed in yeast by high-throughput MS (120, 121). Although individually these various approaches are prone to error, fail to capture the dynamic regulation and compartmentalization within the cell, and are still far from saturating, taken together they provide a new level of information (122, 123). These screens have emphasized the degree to which cellular proteins and signaling complexes are interconnected. As one example, analysis of proteins involved in DNA repair has connected previously identified complexes into a larger assembly that links components of the various DNA repair pathways to those of the DNA damage checkpoint (120).

### Evolution of Signaling Pathways

The reiterated use of interaction domains may have developed in part to facilitate the evolution of new cellular functions, because domains may be readily joined in new combinations to create novel connections and pathways within the cell. For example, coupling of protein phosphorylation to ubiquitination could have been achieved by simply linking a pSer-pThr- or pTyr- recognition module to a RING domain that binds components of the ubiquitination machinery. Conventional tyrosine kinases and SH2 domains are absent from yeast but make a coordinate appearance with the development of multicellular animals. An SH2 domain, by its design, can be inserted into preexisting proteins and thereby provide a common means of coupling entirely different proteins to tyrosine kinase signals. Clearly this does not exclude the subsequent elaboration of more sophisticated levels of control within signaling complexes.

The joining of separate domains can also create a new composite entity with more complex properties than either domain alone. Dystrophin and  $\beta$ -dystroglycan form part of a complex that couples the internal actin cytoskeleton to the extracellular basal lamina,

and is defective in Duchenne and Becker muscular dystrophies. The C-terminal region of the human dystrophin protein has a WW domain (which typically binds proline-rich motifs in a manner similar to SH3 domains) embedded within two EF-hand-like domains (124). This larger module forms a compound binding surface for an extended peptide motif from the cytoplasmic tail of  $\beta$ -dystroglycan. This extensive interaction may be necessary for the formation of a stable complex involved in organizing the muscle cytoskeleton, in contrast to the evanescent associations more typical of signal transduction pathways.

Interaction domains and motifs therefore provide a way to increase the connectivity of existing proteins, and thus to endow these proteins with new functions. This is likely one of several reasons that the apparent complexity of organisms can increase so markedly without a corresponding increase in gene number. An attribute of proteins encoded by the human genome is that they have a richer assembly of domains than do their counterparts in invertebrates or yeast (125, 126), and indeed the assortment of domains into novel combinations is likely an important aspect of genome divergence (127).

### Pathogenic Proteins Rewire Cellular Interaction Networks

Mutant cellular proteins that cause inherited disorders or malignancy can exert their effects through the loss of protein-protein interactions or, conversely, by the creation of aberrant protein complexes. In the former category, mutations in the PTB domain of the human ARH protein, which acts as an adaptor to link low-density lipoprotein receptors (LDLRs) to the endocytic machinery, result in a rare hypercholesterolemia similar to that associated with LDLR mutations, probably by causing defects in LDLR internalization (128, 129). In Noonan syndrome, an autosomal dominant disorder with pleiotropic developmental abnormalities, mutations in the Shp2 tyrosine phosphatase suppress an autoinhibitory interaction between the N-terminal SH2 domain and the catalytic domain, and therefore inappropriately activate the phosphatase (130).

Chromosomal rearrangements in cancer cells can result in the production of chimeric oncoproteins with the potential to bridge novel protein-protein interactions. The Bcr-Abl oncoprotein, for example, is oligomerized through its N-terminal Bcr region and is consequently autophosphorylated at Tyr<sup>177</sup> within the Bcr sequence; this site engages the Grb2 SH2 domain in leukemic cells and contributes to the oncogenicity and disease spectrum of Bcr-Abl in mouse models (131).

This ability to forge new interactions, and thus to reprogram cellular behavior, is also a strategy adopted by pathogenic microorganisms. The intracellular motility of vaccinia virus

particles is dependent in part on the ability of the viral protein A36R to induce the polymerization of actin filaments behind the viral particle. This activity of A36R requires its phosphorylation on Tyr<sup>112</sup> and Tyr<sup>132</sup>. Phosphorylated Tyr<sup>112</sup> recruits the SH2 domain of the Nck SH2-SH3 adaptor, which couples through its SH3 domains to a complex composed of N-WASP (N-Wiskott-Aldrich syndrome protein) and WIP (WASP interacting protein), thereby stimulating actin polymerization through the Arp (actin-related protein) 2/3 complex (Fig. 3F). These interactions provide vaccinia A36R with an ability to regulate actin formation; the Tyr<sup>132</sup> site plays an ancillary role through the recruitment of Grb2, which may stabilize the association with N-WASP (132). Intriguingly, the Tir protein of enteropathogenic *Escherichia coli* (EPEC) also couples to the cytoskeleton through Nck, using a very similar motif to that of vaccinia A36R (EHIpYDSVA and EHIpY-DEVA, respectively) (133). Thus, a virus and a bacterium have independently acquired the same motif to bind Nck and reorganize the host cytoskeleton.

### Experimental and Therapeutic Rewiring of Cell Signaling

These observations suggest that rewiring of protein-protein interactions could be used experimentally to alter cellular function—for example, by the creation of novel chimeric proteins that enforce unnatural interactions, as do some pathogenic polypeptides. In yeast, Ste11 is a MAP kinase kinase kinase (MAP-KKK) that functions in both the mating and osmosensing MAP kinase (MAPK) pathways and is directed to these alternate pathways by scaffolds (Ste5 or Pbs2, respectively). Ste11 signaling can, however, be routed selectively down one pathway or the other by its fusion to the relevant MAPK kinase or scaffold (134), and a chimeric Ste5-Pbs2 scaffold can channel an  $\alpha$ -factor mating pheromone signal to the osmosensing pathway (135). Notably, the requirement for association of the Ste11 or Ste7 kinases with the Ste5 scaffolding protein in the mating response can be overcome by fusing a PDZ domain to the relevant kinase and fusing a complementary heterodimerizing PDZ domain to Ste5 (or a distinct Ste5-associated kinase) (135). A precise spatial organization of the interacting proteins assembling on the Ste5 scaffold is therefore not essential for biological specificity, but may have evolved subsequently to provide enhanced selectivity, regulation, and efficiency in signaling. Understanding the network of cellular protein interactions should expand the scope for creating novel biological responses through engineered proteins or small molecules.

Small molecules can be used to modify protein-protein interactions in a variety of ways. Drugs such as FK506 and rapamycin inactivate

a cellular protein (the calcineurin phosphatase or TOR protein kinase, respectively) by nucleating a nonphysiological complex between the specific target and an immunophilin protein. These and other compounds therefore achieve a therapeutic effect by creating novel protein-protein interactions (136, 137). In a related fashion, the fungal toxin fusicoccin stabilizes a relatively weak interaction between a pThr-Val motif at the extreme C terminus of a plasma membrane proton pump and a 14-3-3 protein, causing constitutive activation of the pump and disease in infected plants (138). Such examples indicate that stabilizing or rewiring modular protein interactions is a promising route to drug design. Given the broad repertoire of protein-protein interactions involved in disease, the direct approach of inhibiting interactions is potentially of great value and has yielded lead compounds with *in vivo* activity (139–141). This latter method is more challenging because of considerations such as the relatively large surface areas involved in protein-protein interactions.

Drugs that target catalytic domains can also indirectly exploit modular protein interactions. Comparison of the autoinhibited structures of the Abl and Src cytoplasmic tyrosine kinases reveals that the Src SH2 domain undergoes an intramolecular interaction with a pTyr site in the enzyme's C-terminal tail, whereas the Abl SH2 domain docks directly with helices in the large lobe of the kinase domain (142). These different modes of SH2-kinase interaction impose distinct structural constraints on the Abl and Src autoinhibited kinase domains, which are exploited by Gleevec (also known as STI-571 and imatinib), a drug that inhibits the aberrant tyrosine kinase activity of the Bcr-Abl oncoprotein and therefore has a therapeutic effect in the treatment of chronic myelogenous leukemia (143). As a consequence of the distinct kinase conformations enforced by the Abl and Src SH2 domains, Gleevec selectively inhibits Abl kinase activity, even though it interacts with residues that are conserved between the two kinases (142, 144).

### Signaling Kinetics

Protein-protein interactions can affect signaling kinetics, for example by recruiting positive or negative regulators involved in feed-forward or feedback control. In addition, if a protein-protein interaction is dependent on multisite phosphorylation, this may create a switch-like response as the activity of the relevant kinase rises above a set threshold. Such appears to be the case for the phosphorylation-dependent binding of a yeast cyclin-dependent kinase (CDK) inhibitor, Sic1, to the WD40 domain of the F-box protein Cdc4, a component of an SCF E3 ubiquitin ligase complex. Sic1 represses the activity of the S-phase CDK, and must therefore be eliminated to allow for the onset of DNA replication. This is achieved through the phospho-

rylation of Sic1 on Ser and Thr residues by the G1 CDK, which in turn results in binding of Sic1 to Cdc4, and consequent Sic1 ubiquitination and degradation. Recruitment of Sic1 to Cdc4 requires that Sic1 be phosphorylated on six or more sites, a mechanism that may provide a timing device in passage through the G<sub>1</sub> phase of the cell cycle and an ultrasensitive switch for entry into S phase (13, 51).

The tethering of distinct signaling proteins into a single complex may also affect the kinetics of the response to a stimulus. The muscle-specific scaffolding protein mAKAP binds both to the RII regulatory subunit of adenosine 3',5'-monophosphate (cAMP)-dependent protein kinase (PKA) and to the N terminus of a type 4 phosphodiesterase (PDE4D3) (145). In unstimulated cells, the tonic activity of PDE4D3 keeps local concentrations of cAMP, and thus the activity of PKA, to a minimum. However, hormonal stimulation increases cAMP concentrations to a level that overwhelms the suppressive effect of the phosphodiesterase, thereby activating PKA. PKA phosphorylates, among other targets, the adjacent PDE4D3 at Ser<sup>13</sup> and Ser<sup>54</sup>, inducing a factor of 3 increase in the  $V_{\max}$  of the phosphodiesterase and attenuating cAMP signaling. This complex may therefore turn an otherwise weak but prolonged signal into a sharply defined pulse of PKA activity. In a similar vein, the binding of the GSK-3 $\beta$  protein kinase and its substrate  $\beta$ -catenin to the scaffolding protein axin increases  $\beta$ -catenin phosphorylation by a factor of 20,000 relative to the rate of the unscaffolded reaction (146).

The kinetics of signaling through MAP kinase pathways have likewise been of great interest, because cells respond very differently to transient or prolonged MAPK activation. One of the targets of the Erk MAPK is the c-Fos transcription factor. When MAPK activity is low, c-Fos expression is induced but the protein is unstable. However, more sustained MAPK signaling phosphorylates Ser<sup>362</sup> and Ser<sup>374</sup> in the C terminus of c-Fos, and this exposes a docking motif (FTYP) for Erk, leading to a more stable complex between the kinase and transcription factor, the phosphorylation of additional N-terminal threonine residues on c-Fos, and stabilization of the AP-1 transcription complex. Thus, sustained MAP kinase signaling induces a physical interaction between c-Fos and Erk that is important for the expression of immediate early genes (147). Such examples illustrate how rather simple interactions can be exploited to generate more complex cellular responses in signaling and cell cycle control.

### Conclusions

Interaction domains play a pervasive role in regulating the dynamic organization of eukary-

## REVIEW

otic cells, and indeed this principle extends to prokaryotes, as modules such as FHA domains are common in bacteria (148). Although such domains are superficially rather simple in their binding properties, increasing evidence suggests that interaction domains have been selected for their flexibility, their ability to assemble multiprotein machines, and their potential to mediate sophisticated biological functions. The modular nature of cell regulatory proteins has likely been a driving force in the evolution of increasingly complex and specialized cellular activities, and it is interesting to contemplate the possibility of endowing cells with new properties by using protein interactions to respecify signaling pathways.

### References and Notes

1. T. Pawson, P. Nash, *Genes Dev.* **14**, 1027 (2000).
2. M. B. Yaffe, *Nature Rev. Mol. Cell Biol.* **3**, 177 (2002).
3. P. J. Cullen, G. E. Cozier, G. Banting, H. Mellor, *Curr. Biol.* **11**, R882 (2001).
4. R. Ghose, A. Shekhtman, M. J. Goger, H. Ji, D. Cowburn, *Nature Struct. Biol.* **8**, 998 (2001).
5. R. R. Copley, T. Doerks, I. Letunic, P. Bork, *FEBS Lett.* **513**, 129 (2002).
6. P. Nash, T. Pawson, unpublished data.
7. T. Pawson, J. D. Scott, *Science* **278**, 2075 (1997).
8. B. J. Mayer, *J. Cell Sci.* **114**, 1253 (2001).
9. Z. Songyang et al., *Cell* **72**, 767 (1993).
10. A. H. Tong et al., *Science* **295**, 321 (2002).
11. M. B. Yaffe et al., *Nature Biotechnol.* **19**, 348 (2001).
12. H. W. Kessels, A. C. Ward, T. N. Schumacher, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8524 (2002).
13. P. Nash et al., *Nature* **414**, 516 (2001).
14. T. Tsukazaki, T. A. Chiang, A. F. Davison, L. Attisano, J. L. Wrana, *Cell* **95**, 779 (1998).
15. M. Huse et al., *Mol. Cell* **8**, 481 (2001).
16. J. W. Wu et al., *Mol. Cell* **8**, 1277 (2001).
17. J. Massague, *Nature Rev. Mol. Cell Biol.* **1**, 169 (2000).
18. B. Chan et al., *Nature Cell Biol.* **5**, 155 (2003).
19. S. Latour et al., *Nature Cell Biol.* **5**, 149 (2003).
20. L. W. Donaldson, G. Gish, T. Pawson, L. E. Kay, J. D. Forman-Kay, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14053 (2002).
21. Q. Liu et al., *Mol. Cell* **11**, 471 (2003).
22. N. Blomberg, E. Baraldi, M. Nilgels, M. Saraste, *Trends Biochem. Sci.* **24**, 441 (1999).
23. G. Jogl et al., *EMBO J.* **21**, 4785 (2002).
24. K. Hamada et al., *EMBO J.* **22**, 502 (2003).
25. M. R. Groves, D. Barford, *Curr. Opin. Struct. Biol.* **9**, 383 (1999).
26. A. H. Huber, W. L. Weis, *Cell* **105**, 391 (2001).
27. K. Eklof Spink, S. G. Fridman, W. L. Weis, *EMBO J.* **20**, 6203 (2001).
28. F. Poy, M. Lepourcelet, R. A. Shivdasani, M. J. Eck, *Nature Struct. Biol.* **8**, 1053 (2001).
29. T. A. Graham, W. K. Clements, D. Kimelman, W. Xu, *Mol. Cell* **10**, 563 (2002).
30. X. Wang, P. D. Zamore, T. M. Hall, *Mol. Cell* **7**, 855 (2001).
31. T. A. Edwards, S. E. Pyle, R. P. Wharton, A. K. Aggarwal, *Cell* **105**, 281 (2001).
32. X. Wang, J. McLachlan, P. D. Zamore, T. M. Hall, *Cell* **110**, 501 (2002).
33. M. S. Kimber, J. Nachman, G. Gish, T. Pawson, E. Pai, *Mol. Cell* **5**, 1043 (2000).
34. P. Lasko, Gene regulation at the RNA layer: RNA-binding proteins in intercellular signaling networks. *Sci. STKE* **2003**, re6 (2003).
35. K. Hofmann, *Cell. Mol. Life Sci.* **55**, 1113 (1999).
36. F. Martinon, K. Burns, J. Tschopp, *Mol. Cell* **10**, 417 (2002).
37. C. A. Kim, M. Gingery, R. M. Pilpa, J. U. Bowie, *Nature Struct. Biol.* **9**, 453 (2002).
38. Z. Songyang et al., *Science* **275**, 73 (1997).
39. B. J. Hillier, K. S. Christopherson, K. E. Prehoda, D. S. Bredt, W. A. Lim, *Science* **284**, 812 (1999).
40. C. Nourry, S. G. N. Grant, J.-P. Borg, PDZ domain proteins: Plug and play! *Sci. STKE* **2003**, re7 (2003).
41. J. Kuriyan, D. Cowburn, *Annu. Rev. Biophys. Biomol. Struct.* **26**, 259 (1997).
42. J. M. Bradshaw, V. Mitaxov, G. Waksman, *J. Mol. Biol.* **299**, 521 (2000).
43. J. Sayos et al., *Nature* **395**, 462 (1998).
44. S.-C. Li et al., *Curr. Biol.* **9**, 1355 (1999).
45. M. B. Yaffe, A. E. Elia, *Curr. Opin. Cell Biol.* **13**, 131 (2001).
46. D. Durocher et al., *Mol. Cell* **6**, 1169 (2000).
47. J. Li et al., *Mol. Cell* **9**, 1045 (2002).
48. M. F. Schwartz et al., *Mol. Cell* **9**, 1055 (2002).
49. M. B. Yaffe et al., *Cell* **91**, 961 (1997).
50. M. A. Verdecia, M. E. Bowman, K. P. Lu, T. Hunter, J. P. Noel, *Nature Struct. Biol.* **7**, 639 (2000).
51. S. Orlicky, X. Tang, A. Willems, M. Tyers, F. Sicheri, *Cell* **112**, 243 (2003).
52. P. Jaakkola et al., *Science* **292**, 468 (2001).
53. M. Ivan et al., *Science* **292**, 464 (2001).
54. J. H. Min et al., *Science* **296**, 1886 (2002).
55. W. C. Hon et al., *Nature* **417**, 975 (2002).
56. T. Jenwein, C. D. Allis, *Science* **203**, 1074 (2001).
57. B. J. Mayer, H. Hirai, R. Sakai, *Curr. Biol.* **5**, 296 (1995).
58. D. J. Owen et al., *EMBO J.* **19**, 6141 (2000).
59. R. H. Jacobson, A. G. Ladurner, D. S. King, R. Tjian, *Science* **288**, 1422 (2000).
60. S. Mujtata et al., *Mol. Cell* **9**, 575 (2002).
61. A. J. Bannister et al., *Nature* **410**, 120 (2001).
62. P. R. Nielsen et al., *Nature* **416**, 103 (2002).
63. S. A. Jacobs, S. Khorasanizadeh, *Science* **295**, 2080 (2002).
64. S. Polo et al., *Nature* **416**, 451 (2002).
65. L. Wybenga-Groot et al., *Cell* **106**, 745 (2001).
66. A. Ashkenazi, *Science* **281**, 1305 (1998).
67. H. Ye, Y. C. Park, M. Kreishman, E. Kieff, H. Wu, *Mol. Cell* **4**, 321 (1999).
68. J. R. Bradley, J. S. Pober, *Oncogene* **20**, 6482 (2001).
69. H. Wajant, F. Henkler, P. Scheurich, *Cell Signal.* **13**, 389 (2001).
70. Y. C. Park et al., *Cell* **101**, 777 (2000).
71. K. Takeda, T. Kaisho, S. Akira, *Annu. Rev. Immunol.* **21**, 335 (2003).
72. K. Wong et al., *Cell* **107**, 209 (2001).
73. G. J. Bashaw, T. Kidd, D. Murray, T. Pawson, C. S. Goodman, *Cell* **101**, 703 (2000).
74. J. Rhee et al., *Nature Cell Biol.* **4**, 798 (2002).
75. J. E. Bear et al., *Cell* **109**, 509 (2002).
76. J. M. Swiercz, R. Kuner, J. Behrens, S. Offermanns, *Neuron* **35**, 51 (2002).
77. V. Perrot, J. Vazquez-Prado, J. S. Gutkind, *J. Biol. Chem.* **277**, 43115 (2002).
78. J. Abrandt, H. G. Vikis, J. S. Gutkind, N. Ahn, K. L. Guan, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12085 (2002).
79. A. Zarrinpar, R. P. Bhattacharyya, W. A. Lim, The structure and function of proline recognition domains. *Sci. STKE* **2003**, re8 (2003).
80. P. Kavsak et al., *Mol. Cell* **6**, 1365 (2000).
81. A. R. Willems et al., *Philos. Trans. R. Soc. London Ser. B* **354**, 1533 (1999).
82. N. Zheng et al., *Nature* **416**, 703 (2002).
83. M. Karin, Y. Ben-Neriah, *Annu. Rev. Immunol.* **18**, 621 (2000).
84. G. Levkowitz et al., *Mol. Cell* **4**, 1029 (1999).
85. C. A. Joazeiro et al., *Science* **286**, 309 (1999).
86. P. Soubeyran, K. Kowanetz, I. Szymkiewicz, W. Y. Langdon, I. Dikic, *Nature* **416**, 183 (2002).
87. A. Petrelli et al., *Nature* **416**, 187 (2002).
88. I. Dikic, *FEBS Lett.* **529**, 110 (2002).
89. Y. Zhu, B. Doray, A. Poussu, V. P. Lehto, S. Kornfeld, *Science* **292**, 1663 (2001).
90. R. Puertollano, R. C. Aguilar, I. Gorshkova, R. J. Crouch, J. S. Bonifacio, *Science* **292**, 1663 (2001).
91. S. Misra, R. Puertollano, Y. Kato, J. S. Bonifacio, J. H. Hurley, *Nature* **415**, 933 (2002).
92. T. Shiba et al., *Nature* **415**, 937 (2002).
93. B. Doray, K. Bruns, P. Ghosh, S. A. Kornfeld, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8072 (2002).
94. F. Sicheri, I. Moarefi, J. Kuriyan, *Nature* **385**, 582 (1997).
95. J. A. Knoblich, *Nature Rev. Mol. Cell Biol.* **2**, 11 (2001).
96. S. Ohno, *Curr. Opin. Cell Biol.* **13**, 641 (2001).
97. K. Kemphues, *Cell* **101**, 345 (2000).
98. D. Lin et al., *Nature Cell Biol.* **2**, 540 (2000).
99. G. Joberty, C. Peterson, L. Gao, I. G. Macara, *Nature Cell Biol.* **2**, 531 (2000).
100. R.-G. Qiu, A. Arbo, G. S. Martin, *Curr. Biol.* **10**, 697 (2000).
101. K. Ebnet et al., *EMBO J.* **20**, 3738 (2001).
102. T. W. Hurd, L. Gao, M. H. Roh, I. G. Macara, B. Margolis, *Nature Cell Biol.* **5**, 137 (2003).
103. J. Betschinger, K. Mechtler, J. A. Knoblich, *Nature* **422**, 326 (2003).
104. P. Plant et al., *Nature Cell Biol.* **5**, 301 (2003).
105. D. Bilder, M. Schober, N. Perrimon, *Nature Cell Biol.* **5**, 53 (2003).
106. G. Tanentzapf, U. Tepass, *Nature Cell Biol.* **5**, 46 (2003).
107. T. Pawson, *Nature* **373**, 573 (1995).
108. D. Fambrough, K. McClure, A. Kazlauskas, E. S. Lander, *Cell* **97**, 727 (1999).
109. L. S. Lock, I. Royal, M. A. Naujokas, M. Park, *J. Biol. Chem.* **275**, 31536 (2000).
110. T. Agalioti, G. Chen, D. Thanos, *Cell* **111**, 285 (2002).
111. A. H. Hassan et al., *Cell* **111**, 369 (2002).
112. S. L. Schreiber, B. E. Bernstein, *Cell* **111**, 771 (2002).
113. M. Sheng, C. Sala, *Annu. Rev. Neurosci.* **24**, 1 (2001).
114. M. Sheng, M. J. Kim, *Science* **298**, 776 (2002).
115. M. Migaud et al., *Nature* **396**, 433 (1998).
116. M. Aarts et al., *Science* **298**, 846 (2002).
117. S. B. Ficarro et al., *Nature Biotechnol.* **20**, 301 (2002).
118. A. R. Salomon et al., *Proc. Natl. Acad. Sci. U.S.A.* **100**, 443 (2003).
119. R. Albert, H. Jeong, A. L. Barabasi, *Nature* **406**, 378 (2000).
120. Y. Ho et al., *Nature* **415**, 180 (2002).
121. A. C. Gavin et al., *Nature* **415**, 141 (2002).
122. G. D. Bader, C. W. Hogue, *Nature Biotechnol.* **20**, 991 (2002).
123. C. von Mering et al., *Nature* **417**, 399 (2002).
124. X. Huang et al., *Nature Struct. Biol.* **7**, 634 (2000).
125. J. C. Venter et al., *Science* **291**, 1304 (2001).
126. E. S. Lander et al., *Nature* **409**, 860 (2001).
127. G. Apic, J. Gough, S. A. Teichmann, *Bioinformatics* **17**, 583 (2001).
128. C. K. Garcia et al., *Science* **292**, 1394 (2001).
129. G. He et al., *J. Biol. Chem.* **277**, 44044 (2002).
130. M. Tartaglia et al., *Nature Genet.* **29**, 465 (2001).
131. Y. He et al., *Blood* **99**, 2957 (2002).
132. N. Scaplehorn et al., *Curr. Biol.* **12**, 740 (2002).
133. S. Gruenheid et al., *Nature Cell Biol.* **3**, 856 (2001).
134. K. Harris et al., *Curr. Biol.* **11**, 1815 (2001).
135. S. H. Park, A. Zarrinpar, W. A. Lim, *Science* **299**, 1061 (2003).
136. J. Choi, J. Chen, S. L. Schreiber, J. Clardy, *Science* **273**, 239 (1996).
137. J. P. Griffith et al., *Cell* **82**, 507 (1995).
138. M. Wurtele, C. Jelic-Ottmann, A. Wittinghofer, C. Oecking, *EMBO J.* **22**, 987 (2003).
139. A. G. Cochran, *Chem. Biol.* **7**, R85 (2000).
140. J. T. Nguyen, C. W. Turck, F. E. Cohen, R. N. Zuckermann, W. A. Lim, *Science* **282**, 2088 (1998).
141. T. K. Sawyer et al., *Mini Rev. Med. Chem.* **2**, 475 (2002).
142. B. Nagar et al., *Cell* **112**, 859 (2003).
143. B. J. Druker et al., *N. Engl. J. Med.* **344**, 1038 (2001).
144. T. Schindler et al., *Science* **289**, 1938 (2000).
145. K. L. Dodge et al., *EMBO J.* **20**, 1921 (2001).
146. R. Djajani et al., *EMBO J.* **22**, 494 (2003).
147. L. O. Murphy, S. Smith, R. H. Chen, D. C. Fingar, J. Blenis, *Nature Cell Biol.* **4**, 556 (2002).
148. M. Pallen, R. Chaudhuri, A. Khan, *Trends Microbiol.* **10**, 556 (2002).
149. We thank J. Scott, D. Durocher, J. Forman-Kay, J. Dennis, M. Toone, B. Seet, and J. Wrana for comments on the manuscript. We apologize to authors of the many important concepts and papers not discussed here for want of space. T.P. is a Distinguished Scientist of the Canadian Institutes of Health Research (CIHR). P.N. is a Postdoctoral Senior Research Fellow of the CIHR.