Domains, Motifs, and Scaffolds: The Role of Modular Interactions in the Evolution and Wiring of Cell Signaling Circuits

Roby P. Bhattacharyya, Attila Reményi, Brian J. Yeh, and Wendell A. Lim

Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94143; email: lim@cmp.ucsf.edu

Annu. Rev. Biochem.
2006. 75:655–80

First published online as a Review in Advance on March 17, 2006

The *Annual Review of Biochemistry* is online at biochem.annualreviews.org

doi: 10.1146/annurev.biochem.75.103004.142710

0066-4154/06/0707-0655$20.00

Key Words

signal transduction, modularity, evolvability, docking, synthetic biology

Abstract

Living cells display complex signal processing behaviors, many of which are mediated by networks of proteins specialized for signal transduction. Here we focus on the question of how the remarkably diverse array of eukaryotic signaling circuits may have evolved. Many of the mechanisms that connect signaling proteins into networks are highly modular: The core catalytic activity of a signaling protein is physically and functionally separable from molecular domains or motifs that determine its linkage to both inputs and outputs. This high degree of modularity may make these systems more evolvable—in principle, novel circuits, and therefore highly innovative regulatory behaviors, can arise from relatively simple genetic events such as recombination, deletion, or insertion. In support of this hypothesis, recent studies show that such modular systems can be exploited to engineer nonnatural signaling proteins and pathways with novel behavior.

Contents

INTRODUCTION

Living cells must constantly monitor and respond to their environment and internal conditions. In metazoans, individual cells must communicate with and respond to other cells in the organism. Thus, cells require a remarkable array of sophisticated signal processing behaviors that rivals or surpasses that of modern computers. Many of these responses involve processing by networks of cytoplasmic signaling proteins. Here we review recent advances in our understanding of the fundamental design principles underlying the structure and mechanism of eukaryotic signaling proteins, focusing particularly on how they are functionally linked to one another to form complex circuits capable of information processing. We discuss how the modular organization of the polypeptides that participate in signaling may help facilitate the evolution of innovative circuitry and corresponding phenotypes, providing increased fitness in a competitive and changing environment.

EVOLVABILITY OF CELL CIRCUITRY: MAKING NEW CONNECTIONS

How have the incredibly diverse and complex phenotypes observed in modern eukaryotic organisms evolved? A growing body of work suggests new phenotypes rarely arise

through the evolution of radically new proteins (1). Rather, innovation is thought to occur through the establishment of novel connectivities between existing or duplicated proteins to generate new regulatory circuits and thereby new regulatory behaviors (**Figure 1*a***). This model is consistent with the surprisingly small number of protein-coding genes in even very complex organisms and thus the limited number of protein or domain types observed (2–4). Phenotypic diversity and complexity appear to arise from new combinations of proteins and/or protein domains working as a network, not from the generation of completely new protein functions. This strategy is similar to that of electronic circuits—a huge variety of circuits can be built from a finite set of electronic components by wiring them together in different ways. Thus, a critical question is how new input-output connections can be established between biological components.

Connecting Transcriptional Nodes: Structural and Functional Modularity

Although this review focuses on protein-based signaling circuits, it is instructive to consider briefly how new connectivities are generated in transcriptional networks, a different class of biological regulatory networks (**Figure 1*b***). Transcriptional control is mediated by promoters that respond to signals provided by upstream transcription factors and convert this input into gene expression. Transcriptional nodes are highly modular (1, 5–6a). First, they display structural modularity: The output region, the coding sequence to be transcribed, is physically separable from the input regions, the *cis*-acting elements that regulate expression. Second, and perhaps more importantly, they display functional modularity: Input and output components still function when separated and can be recombined to yield new input-output connectivities. For example, insertion of a new *cis*-acting element into a promoter can place a gene under the control of a new input pathway (6a, 7). Alternatively, insertion of a new gene behind a promoter can result in a radically new output in response to the same input signal. Even linking input and output elements that have had no previous physiological relationship will often work, in large part because gene expression is controlled by standardized general transcription machinery. Thus, the highly modular structure of promoters allows the input and output elements to be easily transferred to yield novel connectivities. Transcriptional nodes are therefore thought to provide a highly evolvable system (6a, 8). Recombination of transcriptional input and output components is thought to be a major source of phenotypic variation during evolution (1).

Classical Regulatory Proteins

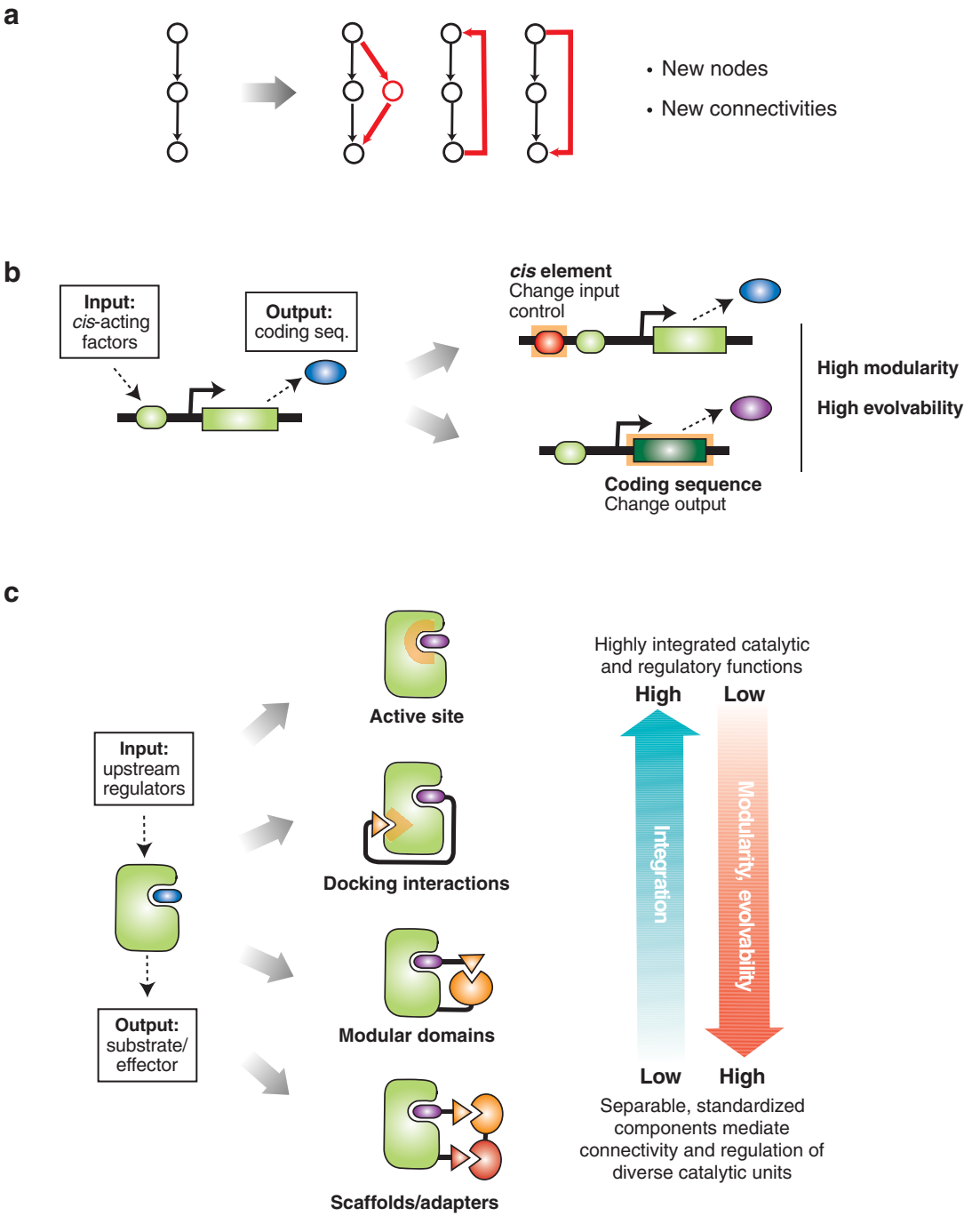
Historically, the best-studied regulatory proteins are enzymes involved in metabolic pathways, which lack the modularity of transcriptional nodes and therefore present several fundamental problems with respect to generating new input-output connectivities. The output of an enzyme—the reaction it catalyzes and the products generated—is dependent on precise stereochemical requirements; thus, enzymes cannot easily undergo radical changes in output without compromising catalytic activity. Input control of enzymes can be mediated by allosteric effectors; binding of these effectors at allosteric sites is coupled to specific conformational changes at the active site (9, 10). The intimate and subtle coupling between allosteric sites and the catalytic center limits the possibility of radically modifying allosteric input control without concomitantly compromising function or stability of the active site. In summary, such metabolic enzymes rarely show structural or functional modularity; the elements that mediate input and output are often found within a single cooperatively folding unit and therefore cannot easily be independently modified. Such systems, which we refer to as being tightly functionally integrated, have less readily

Node: simplest element of a signaling network that can translate input into output

Structural modularity: the ability of a large molecule or system to be physically separated into multiple, structurally independent domains

Functional modularity: a domain's ability to function independent of context, allowing transfer of function between diverse molecular systems

Functional integration: the consolidation of multiple functions, e.g., input and output, into a single, nondecomposable structural module



transferable elements and thus are not as evolvable as functionally modular systems.

Modularity of Eukaryotic Signaling Proteins

Signaling pathways involve enzymes that catalyze reactions such as phosphorylation, dephosphorylation, and nucleotide exchange. The input control of such enzymes determines when, where, and by what they are activated. The output control determines what downstream partners these enzymes act upon once activated.

Signal transduction enzymes utilize far more modular mechanisms to determine their input-output connectivities than do classical metabolic enzymes (11). Over the last decade, our understanding of the design principles of signaling enzymes has increased dramatically as a result of mechanistic and structural studies as well as the sequencing of multiple eukaryotic genomes. Signaling enzymes often contain, in addition to their core catalytic function, multiple independently folding domains or motifs that mediate connectivity by interacting with other signaling elements. These modules are found in different combinations with diverse catalytic functions, suggesting insertion and recombination of modules may be a common mechanism of the evolution of new proteins and connections (2–4).

Eukaryotic signaling proteins appear to have developed a range of modular strategies for controlling their input and output connectivities, all of which involve increased functional separation between core catalytic elements and connectivity elements (**Figure 1*c***). Here we review three basic mechanisms by which the catalytic activity of kinases and other signaling functions are directed and regulated in a modular manner: the use of peripheral docking sites, modular interaction domains, and scaffolding and adapter proteins. Each of these mechanisms can be used to select functional upstream and downstream partners as well as, in many cases, to allosterically regulate catalytic activity. These mechanisms represent a continuum of increasing structural modularity in which catalytic function is separated from the elements that determine its wiring (e.g., scaffolds or adapter proteins represent a separation of catalysis and input control into separate gene products).

We explore the hypothesis that the increasing modularity observed in signaling proteins correlates with higher evolvability: This framework may promote the formation of diverse linkages between catalytic functions via generic, standardized connecting elements. These modular connecting elements may facilitate the evolution of more complex phenotypes, much as standardized components facilitate the design of diverse and complex devices in engineering. We also review an

Module: an independently folding domain that can carry out a simple function

Docking: interaction between a catalytic domain and a partner protein that does not involve the active site

Scaffold: a protein that binds and colocalizes three or more members of a catalytic pathway

Adapter: protein that binds and colocalizes two functionally interacting members of a catalytic pathway

Evolvability: the ability of a system to generate new heritable traits or behaviors through genetic changes

Figure 1

Modularity and evolvability of cellular regulatory circuits and nodes. (*a*) Evolution of new regulatory pathways and responses. A simple linear pathway (*black*) can be converted to a more complex one through the addition of novel nodes that introduce branch points or by the generation of novel functional linkages between existing components, such as the feedback or feedforward circuits depicted. New components and connections are shown in red. (*b*) New connectivity with transcriptional nodes. Transcriptional circuits exemplify a highly modular network, as simple recombination events can alter input-output relationships. Introduction of new *cis*-acting elements such as promoters and enhancers can alter input control, and insertion of a new coding sequence downstream of an existing set of *cis*-acting elements can impose an existing mode of regulation upon expression of a different gene. (*c*) New connectivity with protein/enzyme nodes. Four means of mediating connections between protein nodes are depicted: active site recognition, docking interactions, recognition through modular domain/ligand pairs, and interactions mediated by organizing factors such as scaffolds or adapters. These connection strategies fall on a continuum of modularity versus integration; greater separation between catalytic functions and interactions that mediate connections lends itself to greater evolvability of the signaling network.

emerging body of work that demonstrates recombination of modular components can be used to rewire signaling pathways in nonnative ways, supporting this hypothesis. Because protein phosphorylation is an important currency of information in a cell and protein kinases are among the best studied of the signaling enzymes (12), much of our review focuses on the diversity of ways that protein kinases are integrated into signaling pathways.

DOCKING INTERACTIONS: RECOGNITION BEYOND THE ACTIVE SITE

Since Fischer (13) formulated his lock-and-key hypothesis for enzymes at the end of the nineteenth century, biochemists have generally assumed that the substrate specificity of an enzyme was determined primarily by stereochemical complementarity with its active site. Recently, however, a number of signaling enzymes have been characterized in which surfaces distinct from the active site play an equally important role in mediating substrate or partner recognition (**Figure 2*a***). For example, many proteases have secondary substrate recognition sites referred to as exosites (14, 15). Similarly, many protein kinases have secondary partner recognition sites referred to as docking sites (16, 17). Here we focus on kinase docking sites, as these are well understood and most relevant to our focus on intracellular signaling.

As demonstrated over 30 years ago, many protein kinases display clear preferences for the amino acid sequence immediately surrounding the phosphorylated residue in the substrate (18). Such preferences can now be identified by peptide library–based phosphorylation studies (19, 19a). However, in many

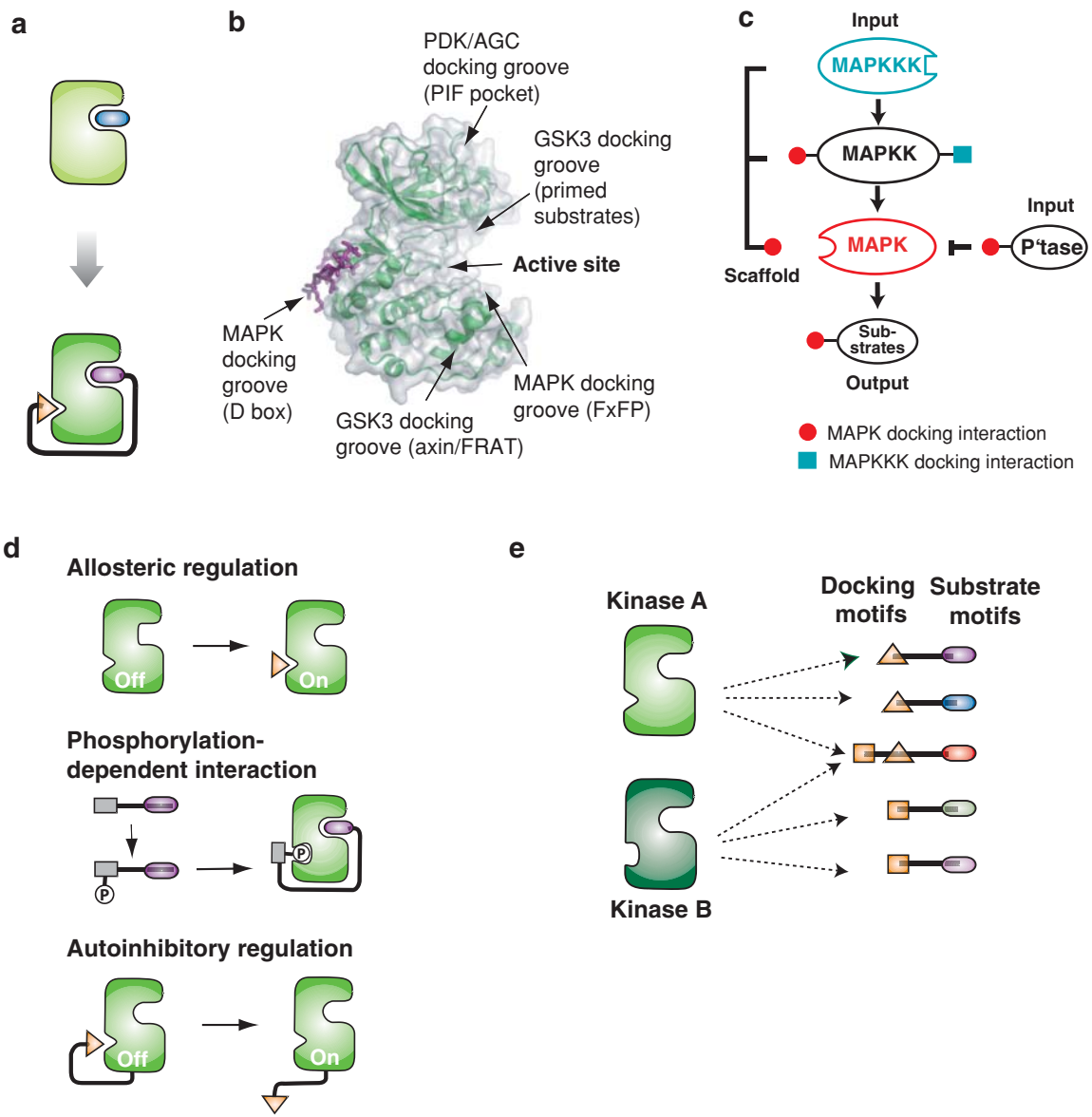
Figure 2

Docking grooves can mediate connectivity and regulation of serine/threonine kinases. (*a*) Docking involves interactions between an enzyme and its substrate that take place away from the active site of the enzyme. Such interactions contribute to substrate selection and catalytic efficiency. (*b*) Docking grooves are found at various surfaces on an enzyme. The structure shown is of the budding yeast MAPK Fus3 but is meant to represent a generic kinase fold for the purposes of illustrating the different possible binding surfaces. Several examples are shown of kinase docking interactions. The mitogen-activated protein kinase (MAPK) docking groove for D-box ligands is on the back side of the kinase opposite the active site; the figure illustrates a Fus3/Ste7 D-box peptide complex crystal structure (26). The 3-phosphoinositide-dependent kinase (PDK)/AGC docking groove, also known as the PDK1 interaction fragment (PIF) pocket, mediates interactions at the N-terminal lobe of the kinase (39). The glycogen synthase kinase-3 (GSK3) docking groove for binding primed substrates is located on the N-terminal lobe adjacent to the active site (41), whereas the interacting surface for another ligand, axin/FRAT, is in the C-terminal lobe (140, 141). The MAPK DEF docking groove for FxFP ligands is adjacent to the active site on the C-terminal lobe (33). Thus, the highly conserved kinase structure has been exploited on several different surfaces for diverse types of docking interactions. (*c*) Docking interactions mediate many different types of connectivities within an MAP kinase cascade. MAPKs have docking grooves that interact with cognate docking motifs in activators [MAP kinase kinases (MAPKKs)], inactivators (MAPK phosphatases), substrates, and other pathway modulators such as scaffolds. In addition, MAP kinase kinase kinases (MAPKKKs) have docking grooves on their kinase domains that interact with DVD (domain for versatile docking) motifs on their MAPKK substrates. (*d*) Some docking interactions regulate enzyme activity in more complex ways than simple localization. Certain docking motifs alter the efficiency of an enzyme (k_{cat}) through classical allosteric effects, repositioning residues involved in catalysis. Others are involved in regulated interactions with substrates, in which covalent modifications such as phosphorylation can either promote or inhibit a docking interaction. Finally, some enzymes contain weak intramolecular docking motifs that autoinhibit their own activity; such an enzyme can then be activated by displacement of the intramolecular docking interaction by an external docking site on a substrate or other effector. (*e*) Docking motifs can direct the specificity of enzymes. Whereas a relatively large array of substrates may fit the stereochemical requirements for catalysis at the active site, those with appropriate docking motifs will be selectively used by kinases with cognate docking grooves.

cases, these substrate motif preferences are not sufficient to predict functional connectivity of kinases: Some ideal motifs do not appear to be endogenous substrates, and conversely, some known endogenous substrates do not match ideal profiles (16, 20, 20a). In addition, certain kinases appear to be quite promiscuous for minimal peptide substrates (16, 20).

Use of Distributed Surfaces for Recognition

The use of docking site interactions has emerged as a common mechanism used by certain serine (Ser)/threonine (Thr) kinases to achieve both selectivity and regulation (16, 17). Docking interactions involve a docking groove on the kinase that is distinct from the active site. The docking groove recognizes



MAPK: mitogen-activated protein kinase

JNK: c-Jun N-terminal kinase

ERK: extracellular signal regulated kinase

a peptide docking motif, which is distinct from the actual phosphoacceptor substrate motif but on the same molecule. Docking interactions appear to function as extended recognition surfaces that can increase enzyme-substrate encounters (reduce K_m) and confer higher specificity than can be achieved by interactions between the active site and substrate motif alone. Moreover, such increases in efficiency and specificity can be achieved without alteration and compromise of active site function.

Docking grooves are found in several Ser/Thr kinase families; here we focus on the mitogen-activated protein kinases (MAPKs) (21, 22). The best-characterized MAPK docking motif is referred to as the D-box, which is recognized by a conserved groove on the MAPK (23). The structures of several D-box docking complexes have been solved (**Figure 2*b***) (24–26), revealing the docking groove is on the opposite surface from the active site. Mutation of either the docking groove on MAPKs or of the docking motif on substrates disrupts proper signal transmission (26–28).

Many MAPKs have analogous D-box interacting sites, including the mammalian MAPKs p38, c-Jun N-terminal kinase (JNK), extracellular signal regulated kinase (ERK), and the yeast kinases Fus3 and Kss1 in *Saccharomyces cerevisiae* and Spc1 in *Schizosaccharomyces pombe* (22, 23, 28–31a). However, many of these kinases show distinct motif-sequence preferences (26, 32). Presumably the distinct docking and active site specificities work together to increase overall selectivity of kinase-substrate interactions.

Docking grooves can also be found at other locations on the surface of certain MAPKs, such as the groove that recognizes the consensus motif FxFP. This docking groove on the MAPK ERK, referred to as the DEF site (docking groove for ERK, FxFP), has been mapped by hydrogen-exchange studies to lie on the large domain of the kinase, below the active site (33). Interestingly, the positional relationship of the phosphoacceptor sites and the MAPK docking motifs within substrates can vary. Whereas D-box motifs are located variably with respect to the phosphoacceptor site, FxFP motifs are almost always 10 residues C-terminal to the phosphoacceptor site. Thus, such motifs can play a role in specifically directing which sites are effectively phosphorylated in a substrate bearing multiple potential phosphorylation sites (34).

Docking grooves have been identified in several families of Ser/Thr kinases, in addition to MAPKs (16). These docking grooves are distributed across the surface of the kinase domain (**Figure 2*b***), illustrating how much of the kinase surface can potentially be tapped for this type of additional recognition function. The spatial relationship between the docking groove and active site on the kinase may set the distance constraints between the docking and phosphoacceptor sites in substrates.

Versatility of Docking Interactions in Organizing Kinase Connectivity

Studies of MAPK pathways reveal the importance and versatility of docking interactions in guiding many circuit connections (**Figure 2*c***). Not only are docking motifs found in MAPK substrates, such as downstream transcription factors, but they are also found in upstream kinases [MAP kinase kinases (MAPKKs)], downregulatory phosphatases, and other regulatory partners, such as scaffold proteins (22, 26, 28, 29, 34a). More recently, docking interactions have been found to play an important role at a different level in MAPK cascades: Several MAP kinase kinase kinases (MAPKKKs) have been found to recognize peptide docking motifs found in their specific MAPKK downstream partners (35, 36). Such motifs have been found in yeast and mammalian systems. The motifs appear to bind directly to the kinase domain of the MAPKKK and to play a critical role in determining MAPKKK → MAPKK specificity.

Regulation via Docking Interactions

In most cases, docking interactions appear to play a relatively passive role as modular specificity control elements: They presumably increase the likelihood of enzyme-substrate encounter. However, in some cases, these interactions appear to regulate kinase function directly (**Figure 2*d***). For example, there are now several reported cases in which peptide binding at the docking groove can allosterically activate kinase function. Certain D-box docking site peptides can stimulate MAPK catalytic activity or autophosphorylation (24), whereas others may inhibit activity (25). FxFP motif binding to ERK appears to be coupled to the positioning of the ERK activation loop (33). In addition, 3-phosphoinositide-dependent kinase-1 (PDK1) interacts with downstream substrate kinases that contain a conserved docking motif known as the PDK1 interaction fragment (PIF). Binding of PIF motifs to PDK1 increases kinase activity (37, 38).

Another way in which docking motifs can act as regulatory elements is when the docking interactions are themselves phosphorylation dependent. For example, PIF motifs must be phosphorylated before they bind effectively to the PIF pocket and activate PDK1 (PIF motif: Phe-X-X-Phe-pSer/pThr-Phe/Tyr). Thus, downstream substrates must be subjected to a priming phosphorylation prior to the interaction with and phosphorylation by PDK1 (38, 39). A similar priming event is required for phosphorylation of some substrates by glycogen synthase kinase-3 (GSK3), which is part of the insulin signaling pathway. GSK3 substrates must be phosphorylated on a residue that is C-terminal to the Ser/Thr site to be modified by GSK3 (40). This priming phosphorylation motif binds to a phospho-recognition docking groove adjacent to the active site (41) (**Figure 2*b***). The priming phosphorylation scheme observed in GSK3 and PDK1 pathways provides a mechanism for making signal processing dependent on a sequence of catalytically distinct phosphorylation events, thereby increasing the specificity and complexity of control.

Finally, docking interactions, because they are critical for proper substrate recognition, can be used as targets for autoinhibition. For example, GSK3 can be inactivated by kinases that phosphorylate its N terminus. This phosphorylation event creates an intramolecular motif that mimics a docking site sequence, binding at the priming phosphate docking groove and occluding downstream substrate recognition (41, 42).

PDK1: 3-phosphoinositide-dependent kinase-1

PIF: PDK1 interaction fragment

GSK3: glycogen synthase kinase-3

Evolvability of Kinase Circuits Using Docking Interactions

The development of substrate recognition sites distinct from the actual phosphoacceptor sequence dramatically increases the modularity of kinase interactions and connectivities. Related kinases can develop slightly different docking grooves, thus allowing them to have distinct specificities without evolutionarily taxing the structure and efficiency of the active site. For instance, the closely related yeast MAPKs, Fus3 and Kss1, which function in the mating and invasive growth pathways, respectively, retain docking grooves that equivalently recognize docking motifs on interacting partners shared by the two kinases, such as the MAPKK Ste7, which functions in both pathways (42a). However, they have evolved some degree of discrimination in binding to substrates specific to one pathway: Fus3 binds the docking motif from the mating pathway effector Far1 more tightly than does Kss1 (26), explaining its selectivity toward this substrate (43). These short docking peptides that mediate specific recognition can be spliced into potential substrates to mediate a new, specific connection (**Figure 2*e***).

Nonetheless, docking motifs are limited in their degree of modularity and evolvability. The docking grooves are intimately tied to the core catalytic module, in this case the catalytic Ser/Thr kinase domain. Thus, although

SH2: Src homology 2

SH3: Src homology 3

docking motifs can be easily transferred to new substrates, the docking grooves cannot be dramatically altered or transferred to unrelated catalytic activities. Docking grooves are a step toward the separation of recognition and catalysis, but they do not employ generic interactions that could be transferred to new functions. Thus, docking grooves may represent a more ancestral solution to achieving modular connectivities. Interestingly, although docking interactions are prevalent in many Ser/Thr kinases (the more ancient eukaryotic protein kinases), similar docking interactions have not been identified in the more recently evolved tyrosine kinases. Instead, as discussed below, many other catalytic functions utilize structurally independent recognition domains to mediate connectivity—a further step toward more standardized circuit connectivity.

MODULAR RECOGNITION DOMAINS: STRUCTURAL SEPARATION OF CONNECTIVITY AND CATALYSIS

The evolution of metazoans appears to have coincided with an explosion in the use of modular protein domains, including many recognition domains that play a major role in diverse cell signaling processes (2–4) (**Table 1**). These include, for example, domains that recognize peptides [e.g., Src homology 3 (SH3) domains], phosphopeptides (e.g., SH2 domains), and phospholipids [e.g., pleckstrin homology (PH) domains]. The detailed functions of these diverse domains are reviewed elsewhere (44–52). Compared with the more specialized Ser/Thr kinase docking sites, such domains represent an even more complete physical separation between elements that

Table 1 Abundance of selected modular domains (and proteins containing them) in commonly studied eukaryotes

	Homo sapiens	*Mus musculus*	*Drosophila melanogaster*	*Caenorhabditis elegans*	*Saccharomyces cerevisiae*
SH3[a]	223 (180)[b]	124 (92)	113 (76)	83 (68)	26 (22)
WW	91 (49)	27 (17)	21 (14)	40 (22)	9 (6)
PDZ	234 (126)	119 (78)	98 (71)	106 (79)	3 (2)
SH2	112 (98)	73 (67)	33 (30)	67 (66)	1 (1)
PTB	34 (30)	14 (12)	7 (7)	23 (20)	0 (0)
14-3-3	8 (8)	4 (4)	4 (4)	2 (2)	1 (1)
BRCT	39 (20)	23 (12)	28 (16)	44 (29)	9 (6)
FHA	16 (16)	9 (9)	17 (17)	12 (12)	13 (12)
C2	149 (99)	94 (63)	51 (36)	93 (64)	22 (11)
Total genes[c]	30,000	30,000	14,000	19,000	6,300

[a]Abbreviations and descriptions of domains in table: SH3 = Src homology 3 domain, binds PxxP peptide ligands (52); WW = PxxP binding domain named after two conserved Trp residues (52); PDZ = domain from PSD-95, Dlg, ZO-1, binds C-terminal peptide ligands (47); SH2 = Src homology 2 domain, binds phospho-Tyr peptide ligands (50); PTB = phospho-Tyr binding domain (50); 14-3-3 = phospho-Ser/Thr binding domain (44); BRCT = breast cancer susceptibility gene, C-terminal domain, binds phospho-Ser/Thr peptide ligands (46); FHA = forkhead-associated domain, binds phospho-Ser/Thr peptide ligands (45); C2 = domain from protein kinase C, binds phospholipids and occasionally phospho-Tyr peptide ligands (51).

[b]These data are gathered from the SMART (Simple Modular Architecture Research Tool) database (**http://smart.embl-heidelberg.de**) in Genomic Mode in May 2006 and reflect our current best estimates of the domain contents of the genomes of these organisms; however, since our knowledge of some of these genomes is less than total, some redundancies may exist, leading to artificially inflated domain counts in some cases (61, 62).

[c]Source: Human Genome Project Information, Functional and Comparative Genomics Fact Sheet (**http://www.ornl.gov/sci/techresources/Human_Genome/faq/compgen.shtml**).

mediate connectivity from those that mediate catalytic functions.

Increased Recombinational Possibilities

From a genetic perspective, modular interactions offer more flexibility than docking interactions: Both the peptide motifs and their cognate domains can be transferred through simple genetic exchanges such as recombination and insertion. Thus, both an enzyme and its substrate can make new connections by incorporating a relevant recognition domain or motif (**Figure 3*a***). Circumstantial evidence for this higher degree of transferability can be found by comparing metazoan genomes. Increasing phenotypic complexity appears to correlate not with the development of new domains (only 7% of human protein families are vertebrate specific), but rather with an increase in the type and number of new domain combinations: Humans have 1.8-fold more distinct protein architectures (arrangements of domains in primary sequence) than do worms and flies (2). An example of domain mixing and matching is shown in **Figure 3*b***, illustrating how specific regulatory and catalytic domains can be found in many combinations to yield proteins, and therefore pathways, with highly diverse input-output relationships.

Regulation by Modular Domains

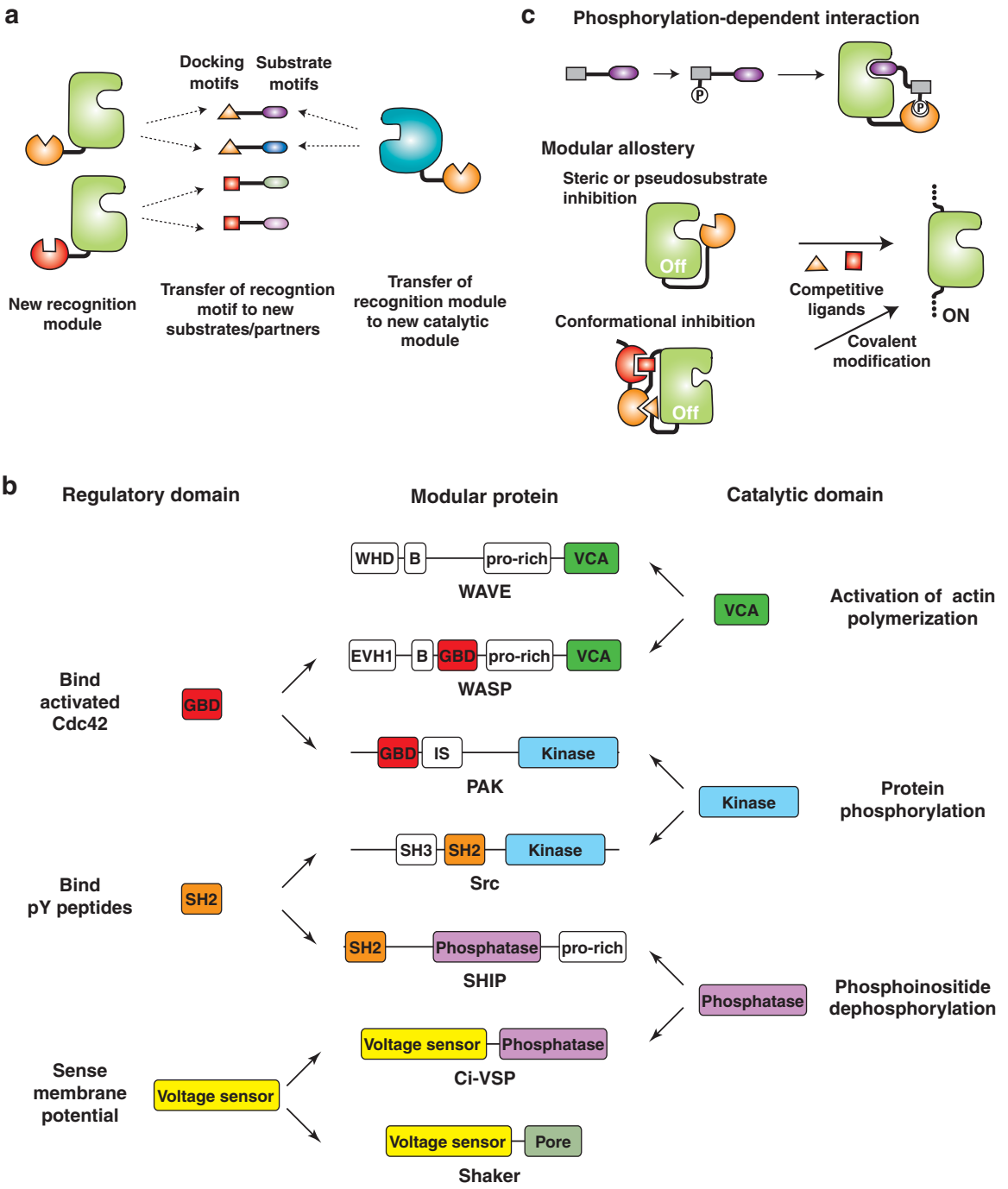
Similar to docking sites, modular domains can be used not only to physically link partner proteins but also as regulatory elements (**Figure 3*c***). Several classes of interaction domains display conditional recognition. These include phosphopeptide recognition domains such as SH2 domains, for which the linkage of a catalytic domain to its partners depends on a prior phosphorylation event (53). Similarly, regulated membrane localization can be achieved with lipid recognition modules that bind to rare phosphoinositide species such as phosphoinositol-(3,4,5)-trisphosphate that are only produced upon activation of phosphoinositide 3-kinase (48).

Modular recognition domains can also play more sophisticated roles in achieving allosteric regulation, most commonly through autoinhibitory mechanisms. Domains can interact in an intramolecular fashion with catalytic domains, either acting as pseudosubstrates or sterically occluding accessibility of the active site (54). The catalytic function can be specifically switched on by the binding of competitive ligands or by covalent modification events that disrupt the autoinhibitory interaction. In other cases, domains can interact with cognate motifs in a manner that conformationally disrupts catalytic function. In some cases, such as the Src family kinases or the actin regulator neuronal Wiskott-Aldrich syndrome protein (N-WASP), multiple domains function together to stabilize an inactive state of their respective catalytic output domains (55–60). In these cases, the proteins can act as sophisticated switches that are able to respond in complex ways to multiple inputs. For example, a protein might approximate an AND gate if two intramolecular interactions must both be disrupted to release the autoinhibited catalytic function. Interestingly, these modular allosteric switches show behavior similar to more conventional allosteric proteins: Switching involves preferential stabilization of a high-activity state by a ligand. However, in the case of modular switches, there is a clear physical and functional separation between the regions of the protein that mediate input regulation and those that mediate output catalytic activity. Not only does this architecture lend itself to increased transferability of function, but modularity may also allow the incremental construction of switch proteins with multiple layers of input control.

N-WASP: neuronal Wiskott-Aldrich syndrome protein

The Problem of Domain Discrimination

Although the use of modular domains may allow the rapid generation of new signaling input-output relationships, the expansion of



domain families presents a new problem: How can repeated domains in a proteome encode specific information in the context of many related family members (**Table 1**)? For example, let us consider the SH3 domain family, which in most cases binds to proline-rich peptides containing the core motif PxxP: The Simple Modular Architecture Research Tool (SMART) database predicts there are 31 SH3 domains in yeast, 132 in *Caenorhabditis elegans*, 273 in *Drosophila*, and 894 in humans (61, 62). How can an ordered array of component connectivities be maintained by such a large set of related domains?

Recent studies suggest several strategies have evolved for maintaining domain discrimination. First, domains can diverge so far from other family members that they display distinct, noncanonical recognition profiles (**Figure 4*a***). For example, some SH3 domains have diverged to no longer recognize PxxP motifs: The C-terminal SH3 domain of the T-cell adapter protein Gads (Grb2-related adapter downstream of Shc) instead recognizes RxxK motifs (63). This recognition event occurs on a surface distinct from the canonical proline binding pocket (64, 65). One Gads ligand, hematopoietic progenitor kinase-1, binds primarily through an RxxK motif, but its binding is augmented by a weak secondary PxxP motif, thus illustrating the versatility of this divergent domain (66). A pair of SH3 domains in p47phox has been found to act as a single unit, using the surface between the two domains to recognize a novel motif (67). Similarly, noncanonical domains have been found in many other domain families, including the SH2 domain from the protein SAP (also called SH2D1A) that binds unphosphorylated motifs (68) and the C2

Figure 3

Modular interaction domains can mediate new connectivity. (*a*) Transferability of modular recognition and catalytic functions. Modular domains facilitate the formation of new connections between proteins, as standardized recognition domains or their ligands can be swapped onto catalytic modules or substrates via recombination events, opening a new set of possible enzyme-substrate interactions. (*b*) Evidence of evolutionary input-output transfer. Naturally occurring examples are depicted in which domains are reused in various combinations to mediate distinct connections between catalytic activities and target molecules. The VCA (verprolin homology, cofilin homology, acidic) domain, which activates actin polymerization, is common to the actin regulatory proteins WAVE and WASP (142), but it is covalently linked to a different set of interaction domains in each case, contributing to distinct modes of deployment of this output activity. Of these interaction domains, the GTPase binding domain (GBD) of WASP is also found in p21-activated kinase (PAK) (143) and is used to direct its binding partner, activated Cdc42, to each of these two diverse proteins. The kinase domain from PAK is reused in many different contexts. The classical example of Src is depicted, in which the kinase domain is joined with several protein interaction domains, including the Src homology 2 (SH2) domain, which regulates the activation state of the kinase and mediates its interaction with phosphotyrosine-containing peptides (55, 58–60). This SH2 domain is, likewise, reused in many signaling components, such as the SHIP phosphatase (144). The phosphatase domain found in SHIP is reused in multitudes of signaling proteins as well, including the recently described voltage-sensing phosphatase from *Ciona intestinalis*, Ci-VSP. As a result of a fusion of the phosphatase domain with a voltage-sensing domain more traditionally found in voltage-gated channels such as Shaker, Ci-VSP exhibits regulation of its phosphatase activity by membrane potential (145, 146). Thus, many complex signaling proteins are built from a relatively small toolkit of standardized components that are combinatorially connected. (*c*) Enzyme regulation by modular domains. Some modular domains only recognize their ligands after covalent modification resulting from other cellular signaling processes, thus linking the connectivity of proteins containing these domains to the regulation of these other pathways. In addition, modular domains are often used to regulate enzyme activity more directly. These domains can participate in interactions that inhibit catalysis, either by sterically blocking access to the catalytic site or by preferentially stabilizing an inactive conformation of the catalytic domain. These inactive states can then be reversed upon exposure to competing ligands that bind to the domains or by covalent modification of the domains or ligands. Abbreviation: pro-rich, proline-rich peptides.

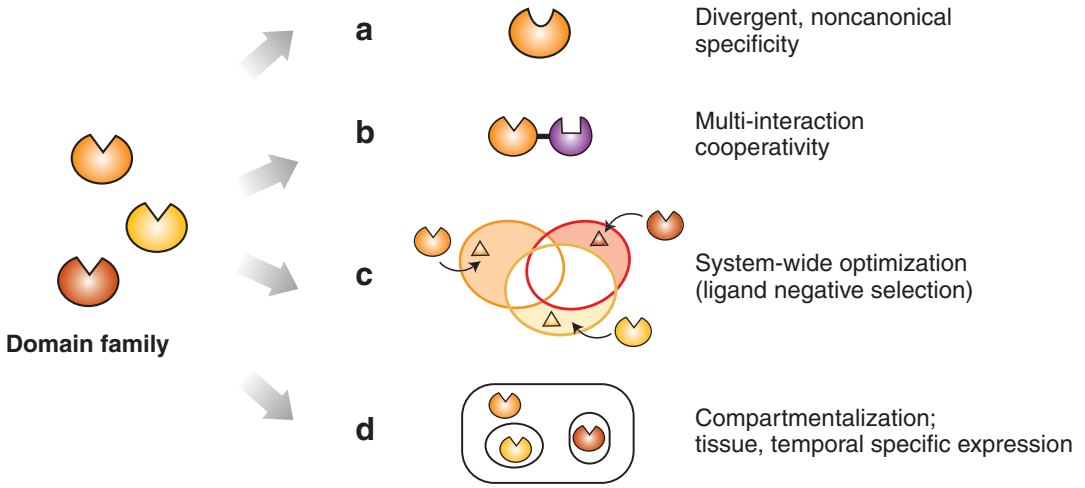


Figure 4

Mechanisms of domain discrimination. (*a*) Domains can evolve divergent ligand-binding pockets that recognize sequences that stray from the consensus for the domain family. (*b*) Multiple domains can be used in combination to generate a combinatorial increase in selectivity and/or affinity compared with the individual recognition events alone. (*c*) Domains and ligands within an organism can coevolve to occupy regions of recognition space with an acceptably low level of cross-recognition. (*d*) Domains and ligands can be segregated in space and time so they are more likely to be coexpressed with genuine interacting partners than with spurious cross-reactive partners.

domain from protein kinase C-δ, from a class of domains that normally binds phospholipids or unphosphorylated peptides, that recognizes phosphotyrosine motifs (51, 69).

A second mechanism for increasing domain-mediated specificity is to use multiple domains to recognize dual ligands in a cooperative manner (70–72) (**Figure 4*b***). A third mechanism is to use system-wide optimization of the domain interaction network (**Figure 4*c***). Recent studies in yeast have shown that although many of the ~30 SH3 domains have overlapping specificity as determined by peptide libraries, there appears to be some level of negative selection against sequences that interact in a highly promiscuous manner (73). Many physiological partner peptides are optimized for specificity not only by positive selection for binding to the proper domain, but also by negative selection against interaction with competing domains in that genome. Thus, in many cases, individual SH3 domains are only observed to interact with a handful of more than 1500 potential PxxP partners within the genome (74). Finally, a fourth way to achieve specificity is to segregate domains either through subcellular compartmentalization, differential temporal expression, or tissue-specific expression (**Figure 4*d***) (73). Thus, domains with highly overlapping recognition properties might never have to compete for the same targets.

Nonetheless, even with these mechanisms, there likely comes a point at which the information-encoding capability of a domain family is saturated, and increasing signaling complexity may require the development of orthogonal domains. For example, SH2 domains are generally only found in metazoans, and the development of SH2 domains and tyrosine phosphorylation-based signaling in general may have been a prerequisite for the evolution of multicellularity, given its need for increased signaling bandwidth (cell-cell signaling in addition to cell-environment signaling). Interestingly, SH2 domains and receptor tyrosine kinases have recently been identified in choanoflagellates, the closest single-celled

eukaryotes to the evolutionary branch point of multicelluarity (75).

SCAFFOLDS AND ADAPTERS: GENETICALLY INDEPENDENT WIRING ELEMENTS

From an evolutionary perspective, the ultimate separation of catalytic and connectivity elements would be to segregate such functions into distinct proteins, each of which is genetically independent. Such a separation is achieved with scaffold and adapter proteins, which act as organizing platforms that recruit specific catalytic elements and their upstream and/or downstream partners to the same complex. In general, adapters are defined as organizing molecules that link together two partners, whereas scaffolds, in general, are defined as organizing molecules that link together more than two partners (**Figure 5*a***).

Organization of Signaling Complexes by Scaffold and Adapter Proteins

Diverse scaffold- or adapter-mediated complexes are observed in eukaryotes (76). For example, cyclins can be thought of as adapters that change the substrate recruitment properties of associated cyclin-dependent kinases (77). Thus, changes in cyclin expression alter the set of target cyclin-dependent kinase substrates (78). Many MAPK cascades and protein kinase A (PKA) response pathways are coordinated by scaffold proteins that organize sequentially acting members of a pathway together (79–84). In some cases, multiple proteins work together to organize specific pathways (85). From an information-transfer perspective, the most important aspect of scaffolding is that it allows catalytic proteins such as kinases to play several distinct roles depending on the complex into which they are assembled (**Figure 5*b***). For example, the yeast MAPKKK Ste11 is used in three distinct MAPK cascades, each of which responds to distinct inputs and yields distinct outputs: the mating, invasive growth, and high-osmolarity response pathways. Ste11 can be used for all of these functions because, in at least two of these pathways, scaffold proteins wire Ste11 such that it retains information about what input activated it and is directed to phosphorylate the appropriate substrates (85–89b). Thus, the subpopulations of Ste11 molecules that participate in different complexes have distinct functions.

Scaffolds and adapters can be built from various interaction components. For example, cyclins interact with both kinase and substrates using highly specialized binding surfaces (77). However, other organizing factors utilize more modular interaction components. The JNK scaffold JIP (JNK interacting protein) uses a canonical MAPK docking motif to bind to JNK (25). Some scaffold interactions are mediated by modular domains, with the scaffold bearing either a modular interaction domain or a motif recognized by such a domain (90, 91).

Regulation by Scaffolds and Adapters

Scaffolds provide many possibilities for complex pathway regulation (**Figure 5*c***). For example, differential expression of a scaffold can determine if a pathway will function in a particular cell type (92). Moreover, in some cases, splice variants of scaffolds lacking specific interaction or localization modules have been identified; thus, temporal or tissue-specific control of splicing could alter pathway wiring (92–94). In addition, some scaffold-mediated interactions can be subject to independent regulation. For example, receptor tyrosine kinases and the immune signaling scaffolds LAT (linker for the activation of T-cells) and SLP-76 (SH2 domain-containing leukocyte phosphoprotein of 76 kDa) contain multiple protein recruitment sites that must first be tyrosine phosphorylated before they organize a complex of SH2 domain-containing proteins (71, 95, 96). In another case, phosphorylation of the mammalian Ras/MAPK scaffold kinase suppressor of Ras (KSR) by the Cdc25-associated kinase is used to

PKA: protein kinase A

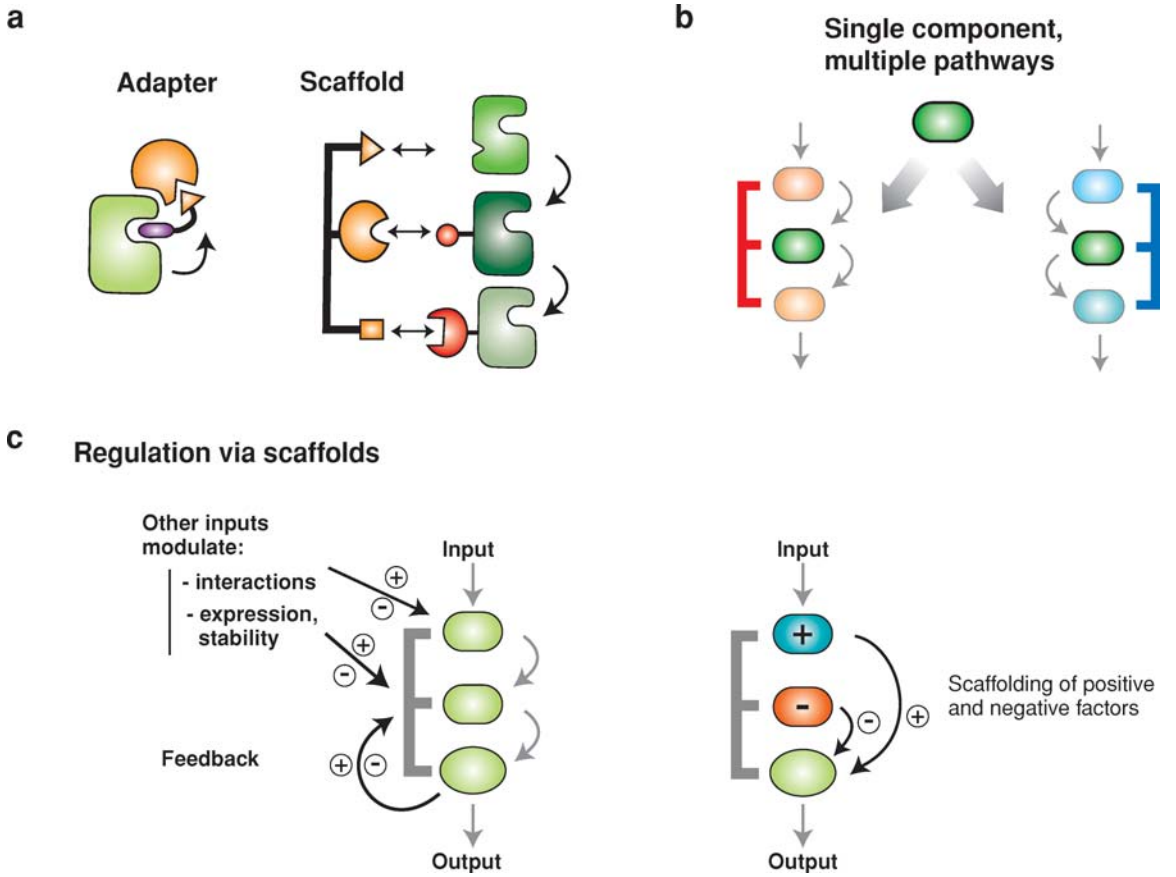


Figure 5

Scaffolds and adapters as mediators of new connectivity. (*a*) Adapters, such as the rounded molecule shown in orange on the left, link two components together. Scaffolds, such as the black and orange molecule on the right, link three or more components of a signaling pathway together. Both classes of molecules separate the catalytic functions of a signaling pathway from the recruitment functions. Either class of molecules may use standardized modular domains or more specific protein-interaction motifs to assemble their associated signaling complex. (*b*) Identical signaling molecules can signal through multiple distinct pathways, responding to different inputs and yielding different outputs despite being activated in chemically indistinguishable ways, by virtue of their recruitment to pathway-dedicated scaffold proteins. These scaffolds act to insulate the shared signaling component in the appropriate complex, encouraging the appropriate interactions for the pathway in question. Activation is often coupled to scaffold localization in such cases, ensuring the fidelity of signal transmission. (*c*) As elements that govern the assembly of components of a signaling complex, scaffolds or adapters can contribute in a number of ways to pathway regulation in addition to passive colocalization. First, they can themselves be the target of modifications, either from outside inputs or from elements of the pathway itself in instances of feedback regulation. Such modifications can alter the way in which interactions occur on these molecular platforms, or they can alter the expression or stability of these critical assembly factors. Second, they can contribute to pathway regulation by recruiting both positive and negative regulators, sometimes in temporally restricted ways, to alter pathway dynamics.

control its interactions and recruitment (97–99). KSR function can also be regulated by a novel protein called impedes mitogenic signal propagation (IMP), which negatively regulates KSR function by inducing its hyperphosphorylation and sequestration into detergent-insoluble domains, preventing it from assembling the normal Ras/MAPK signaling complex (100). The upstream input Ras apparently has two effects. First, activated Ras disrupts the IMP/KSR interaction and induces IMP degradation, thus freeing KSR. Second, Ras interacts with Raf and KSR to promote positive signaling in the pathway. In all of these cases, modulation of the scaffold or its interactions allows regulation of one specific pathway in a manner that does not necessarily affect other unrelated functions of the same catalytic proteins.

Scaffolds can also be used to precisely shape pathway behavior, in addition to simply determining pathway linkages. For example, pathway output, such as terminal kinase activation, might lead to phosphorylation of the scaffold or other components, leading to positive or negative feedback (101, 101a). In the well-studied mating pathway in budding yeast, the scaffold Ste5 allosterically activates one of its binding partners, the MAPK Fus3, initiating a negative feedback loop that regulates pathway output (101a). Such scaffold-mediated feedback loops could have strong effects on the temporal activation profile of the pathway as well as its quantitative dose-response behavior.

In several cases, scaffolds have been found to recruit not only positive acting factors, but also antagonistic negative factors, thus contributing in more complex ways to the shape of the overall signaling response. Some scaffolds recruit both an activating kinase and a deactivating phosphatase. For example, the scaffold JIP-1 (Jnk interacting protein-1) not only binds the MAPK JNK and upstream activators (MAPKKs and MAPKKKs) (102), it also binds the JNK phosphatase MKP7 (MAP kinase phosphatase-7), targeting it to dephosphorylate JNK (103). Another JNK scaffold, β-arrestin 2, binds MKP7 as well, but the phosphatase transiently dissociates from the scaffold upon pathway stimulation, rebinding after 30–60 min (104). Thus, β-arrestin 2 encodes a more sophisticated mechanism for time-dependent pathway activation and inactivation. In addition to MAPK scaffolds, several members of the A kinase anchoring protein family of PKA scaffolds recruit phosphatases as well as PKA (105–107). Some scaffolds use other recruitment strategies to achieve this sort of dynamic pattern of activation and inactivation. For example, a number of A kinase anchoring proteins recruit phosphodiesterases such as PDE4, which break down cyclic AMP, locally inactivating PKA in a spatially restricted negative feedback loop that alters pathway response and may contribute to cyclic AMP homeostasis (108, 109).

The multi-PDZ domain protein InaD acts as a scaffold that mediates *Drosophila* photoreceptor signaling. InaD assembles multiple members of the G-protein coupled phototransduction cascade, including the G-protein effector phospholipase C, the ion channel Trp (transient receptor potential), and an isoform of protein kinase C involved in downregulation of the response (110–112). Thus, InaD appears to play a role not only in accelerating and amplifying the phototransduction response, but also in limiting its timing.

REWIRING SIGNALING PATHWAYS: SYNTHETIC BIOLOGY

Many of the mechanisms described in this review are presented as elements that could facilitate the evolution of new pathway linkages and phenotypes. One approach to testing the evolvability of modular signaling components is to attempt to mimic evolution by using them to create new, synthetic pathways, an approach that is part of the new field of synthetic biology (54, 113, 114).

Synthetic biology: the discipline concerned with engineering biological systems to produce novel biological behaviors

Synthetic Scaffolds, Adapters, and Docking Interactions

Several studies have shown that chimeric or synthetic scaffolds and adapters can be used to generate novel, nonphysiological pathways. For example, a chimeric scaffold made from components of the yeast mating and high osmolarity response MAPK pathways can be used to generate a novel pathway that, in vivo, leads to the osmolarity response when cells are stimulated with mating pheromone (**Figure 6*a***) (115). Even the simple step of transferring an osmoresponse-specific MAPKKK docking motif from the osmolarity MAPKK (Pbs2) to the mating MAPKK (Ste7) is sufficient to induce cross talk to the mating response when cells are stimulated with high osmolarity (36). In mammalian cells, chimeric adapter proteins that link growth factor receptors to apoptotic signaling proteins can be used to convert proliferative signals into death signals (**Figure 6*b***) (116).

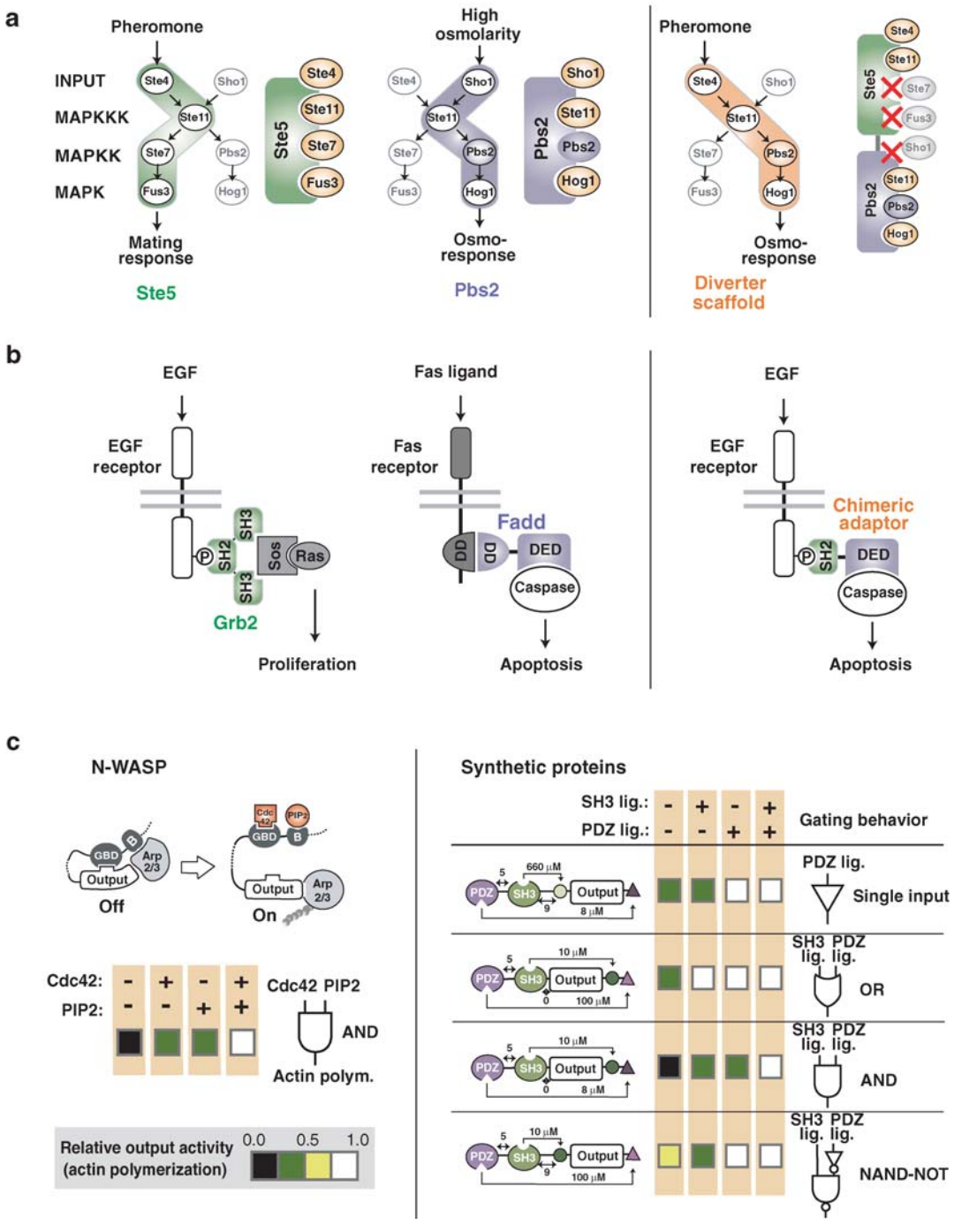
Synthetic Signaling Switches

Modular recombination can be used to reprogram the input control of a signaling protein. The actin regulatory factor N-WASP is normally allosterically regulated; it stimulates actin-related protein (Arp)2/3 complex-mediated actin polymerization when bound to the GTPase Cdc42 and the phospholipid phosphoinositol-(4,5)-bisphosphate (56, 57). This regulation is achieved through modular autoinhibitory interactions. Dueber et al. (117) deleted the regulatory domains of N-WASP and replaced them with heterologous modular interaction domains and their cognate motifs. Many of the resulting synthetic proteins showed allosteric regulation: They were basally repressed but could be activated by the addition of competing ligands (**Figure 6*c***). In many cases, when multiple modular domains were used, the proteins displayed complex signal integration, such as AND-gate behavior requiring the presence of two ligands for potent activation.

Interestingly, the relative ease of generating novel phenotypes from these modular connections can also contribute to the phenotypic manipulation of signaling systems by pathogens or even in stochastic mutations that lead to the development of cancer (49). For instance, the *Yersinia* virulence factor YopM acts as an adapter for the human kinases RSK1 and PRK2, directing a nonnative phosphorylation

Figure 6

Synthetic biology: rewiring modular signaling systems. (*a*) Park et al. (115) constructed a synthetic "diverter" scaffold by combining elements from the yeast mating and high osmolarity scaffolds such that the shared mitogen-activated protein kinase kinase kinase (MAPKKK) Ste11 could be activated by a pheromone on the diverter scaffold but could only transmit this signal to components of the high osmolarity pathway. In this manner, a pheromone input was transduced to the output of the high osmolarity response pathway. (*b*) Howard et al. (116) took advantage of the modular design of mammalian signaling components to construct a chimeric adapter protein that generated a novel input-output linkage. By fusing an Src homology 2 (SH2) domain that recognizes a phosphotyrosine on a growth factor-responsive receptor tyrosine kinase to a death-effector domain (DED) that recruits a caspase involved in apoptosis, the authors redirected a proliferative signal input into an output that favors cell death. (*c*) Dueber et al. (117) rebuilt the naturally occurring modular allosteric switch neuronal Wiskott-Aldrich syndrome protein (N-WASP) by replacing its normal regulatory domains with heterologous modular domains. In this manner, the authors generated variants of N-WASP whose activity was gated in different ways by nonnative inputs, the ligands for the appended modular domains. This simple strategy of recombining two modular domain-ligand pairs onto a catalytic domain yielded a set of proteins with surprisingly diverse gating behaviors depending on subtle variations in parameters such as linker length and binding affinity. Abbreviations: Arp, actin-related protein; EGF, epidermal growth factor; GBD, GTPase binding domain; PIP2, phosphoinositol-(4,5)-bisphosphate. Modified and reprinted from Reference 54, copyright 2004, with permission from Elsevier.



MODULAR DESIGN OF SYNTHETIC DNA MODIFYING ENZYMES

Recent studies have pushed the boundaries of modular design principles in the engineering of synthetic DNA binding proteins. Several groups (126–130a) have constructed libraries of synthetic or naturally occurring zinc finger domains and selected for those that bind to particular 3-bp DNA sequences. If these elements are modular, then multiple zinc fingers could be fused to form multidomain proteins capable of recognizing longer DNA sequences (131, 132). This strategy has succeeded in a number of cases, with as many as 6 zinc fingers fused to form 18-bp recognition elements (126, 132, 133). These extended DNA binding motifs can then be fused to output domains that activate or repress transcription or stimulate cleavage or recombination events, generating designer transcription factors, endonucleases, and recombinases with desired sequence specificity (133–139a). The principles of modular design are critical to the construction of these sophisticated proteins. The challenge of generating a novel, site-specific DNA modifying enzyme can be reduced by combining independently acting modules—multiple recognition domains as well as catalytic elements—into a larger protein with more complex behavior. The ability to generate targeted DNA modifying enzymes on demand has striking implications for functional genomics and gene therapy.

event between these two proteins (118). In addition, the modular organization of signaling proteins can contribute to oncogenesis. Mutational loss or recombination of these regulatory interactions can cause improper activation of important signaling molecules. In the classic case of the v-Src oncogene, a key tyrosine residue involved in an intramolecular interaction with an SH2 domain that autoinhibits the c-Src tyrosine kinase is lost, contributing to oncogenic transformation (119–121). Modularity can thus result in a trade-off between evolvability and fragility.

These studies are consistent with the hypothesis that the modular organization of signaling proteins allows for the facile reconnection of signaling components to yield new pathways and biological responses. In addition, these synthetic approaches present a potentially useful way to systematically perturb and alter complex signaling circuits in a way that may facilitate elucidation of basic systems properties controlling complex biological responses. Moreover, these approaches may allow the rational engineering of cells that could carry out new specific therapeutic functions.

CONCLUSIONS: MODULARITY AND EVOLVABILITY OF BIOLOGICAL REGULATORY SYSTEMS

Highly modular architectures are not only found in eukaryotic signaling systems but also in many other systems, including transcription, proteolysis, and cellular trafficking (1, 122, 123). These systems are characterized by the use of increasingly general, portable elements that can be genetically interchanged to mediate new regulatory connections. The exact domains and motifs that implement these connections vary to some extent, but there may be some pressure to maintain a degree of evolvability in such systems. The reuse of similar modular domains in different contexts represents standardization of the means of communication between protein nodes. Standardization is a central feature of highly complex and evolvable systems (124).

Why might there be selective pressure to maintain modularity and evolvability, given that cellular systems cannot actually foresee the need to change their response behaviors? Presumably, in a constantly changing and competitive environment, the lack of an ability to rapidly evolve novel responses might prove to be a disadvantage. In many cases, highly integrated, nonmodular systems perform in a more efficient, optimal manner, but such performance would only be optimal for a specific and unchanging environment. Hence, modular systems would be more robust to accommodating and buffering against change. During the course of evolution, as environmental pressures shift, there

is likely a constant push and pull between the efficiency of integration on one hand and the flexibility and adaptability provided by modularity on the other. A study modeling network development by standard evolutionary algorithms found that modular network structures and motifs evolved spontaneously in response to shifting evolutionary goals (124a). Even in engineered systems such as electronic circuits, where modular components provide an advantage in circuit development, there is often pressure to minimize and integrate circuits once they are well developed. This optimization and integration can lead to a loss of the modularity that was critical during development (e.g., components in integrated circuits do not have transferable functions). Similarly, one might expect that modularity could easily be lost in fundamental housekeeping biological processes, which do not change significantly over evolution. In support of this model, recent bioinformatics studies indicate tissue-specific proteins, especially those associated with evolutionarily newer functions, tend to have a more modular composition than those proteins that are globally expressed and have a housekeeping function (125). Hopefully as more families of closely related genomes are sequenced, we will gain insight into the actual paths by which new signaling pathways have arisen over the course of evolution.

SUMMARY POINTS

1. Eukaryotic signaling proteins use modular strategies to achieve specific circuit connectivities. These strategies are often characterized by a physical and functional separation between elements of the protein that carry out catalytic functions and elements that determine upstream and downstream partner linkages.
2. Docking interactions between recognition grooves and peptide motifs allow for some degree of separation between recognition events and catalytic function, allowing for more adaptable interactions. In addition, the peptide motifs are readily transferable, creating new potential pathway linkages, whereas the docking grooves are less modular.
3. Protein interactions mediated by specialized modular domains allow for standardized, transferable interactions between catalytic elements and their targets or effectors. The interacting regions in this case are bidirectionally transferable, as either the domain or ligand can readily be exchanged onto a new protein, conferring new functional linkages.
4. Proteins specialized for protein interaction, termed scaffold or adapter proteins, further separate catalysis from molecular recognition. This separation allows the same catalytic molecule to be used in multiple distinct pathways with minimal cross-signaling.
5. The relative ease of transferability inherent to recognition events mediated by standardized modular domains may facilitate the evolution of new connections in signaling pathways, hence the development of complex signaling behaviors.
6. Common catalytic domains and protein interaction domains are recombined together in many different combinations in metazoans to yield complex targeting and regulation of catalytic activity. These multidomain proteins are enriched in cell signaling and other complex processes and are more likely to be expressed in a tissue-specific manner.

7. The principle that modular architecture contributes to the development of complexity can be exploited in the design of synthetic signaling systems, allowing the construction of proteins with novel regulation and behavior from a toolkit of common components.

FUTURE ISSUES TO BE RESOLVED

1. Comparative genomics studies will help illuminate the evolutionary origins of complex signaling proteins and networks.
2. Detailed analysis of gene expression in specific tissues will test the prediction that evolutionarily newer tissues and processes, such as the brain and immune system, will be enriched for the expression of highly modular proteins.
3. Given the separation between catalysis and molecular recognition in metazoan signaling systems, pharmacological disruption of protein interactions will continue to hold promise for more "surgical" interference with specific signaling events compared with the more general tactic of inhibiting catalytic activity.
4. Docking interactions, modular domains, and scaffold proteins will be combined in synthetic signaling pathways with increasingly sophisticated behaviors in vivo, with potential research and therapeutic applications.

ACKNOWLEDGMENTS

We thank C. Tang, C. Voigt, A. Arkin, J. Dueber, N. Helman, C. Bashor, and other members of the Lim lab for their helpful comments on the manuscript.

LITERATURE CITED

1. Carroll SB. 2005. *PLoS Biol.* 3:e245
2. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. 2001. *Nature* 409:860–921
3. Rubin GM. 2001. *Nature* 409:820–21
4. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, et al. 2001. *Science* 291:1304–51
5. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. 1999. *Nature* 402:C47–52
6. Reményi A, Scholer HR, Wilmanns M. 2004. *Nat. Struct. Mol. Biol.* 11:812–15

6a. Ptashne M, Gann A. 2002. *Genes and Signals*. Cold Spring Harbor, NY: Cold Spring Harbor Lab.

7. Ihmels J, Bergmann S, Gerami-Nejad M, Yanai I, McClellan M, et al. 2005. *Science* 309:938–40
8. Kirschner M, Gerhart J. 1998. *Proc. Natl. Acad. Sci. USA* 95:8420–27
9. Monod J, Changeux JP, Jacob F. 1963. *J. Mol. Biol.* 6:306–29
10. Perutz MF. 1978. *Sci. Am.* 239:92–125
11. Pawson T, Nash P. 2003. *Science* 300:445–52
12. Johnson SA, Hunter T. 2005. *Nat. Methods* 2:17–25
13. Fischer E. 1894. *Ber. Dtsch. Chem. Ges.* 27:2985
14. Krishnaswamy S. 2005. *J. Thromb. Haemost.* 3:54–67

15. Overall CM. 2002. *Mol. Biotechnol.* 22:51–86
16. Biondi RM, Nebreda AR. 2003. *Biochem. J.* 372:1–13
17. Holland PM, Cooper JA. 1999. *Curr. Biol.* 9:R329–31
18. Kemp BE, Bylund DB, Huang TS, Krebs EG. 1975. *Proc. Natl. Acad. Sci. USA* 72:3448–52
19. Hutti JE, Jarrell ET, Chang JD, Abbott DW, Storz P, et al. 2004. *Nat. Methods* 1:27–29
19a. Songyang Z, Blechner S, Hoagland N, Hoekstra MF, Piwnica-Worms H, Cantley LC. 1994. *Curr. Biol.* 4:973–82
20. Miller WT. 2003. *Acc. Chem. Res.* 36:393–400
20a. Verma R, Annan RS, Huddleston MJ, Carr SA, Reynard G, Deshaies RJ. 1997. *Science* 278:455–60
21. Sharrocks AD, Yang SH, Galanis A. 2000. *Trends Biochem. Sci.* 25:448–53
22. Tanoue T, Adachi M, Moriguchi T, Nishida E. 2000. *Nat. Cell Biol.* 2:110–16
23. Jacobs D, Glossip D, Xing H, Muslin AJ, Kornfeld K. 1999. *Genes Dev.* 13:163–7524
24. Chang CI, Xu BE, Akella R, Cobb MH, Goldsmith EJ. 2002. *Mol. Cell* 9:1241–49
25. Heo YS, Kim SK, Seo CI, Kim YK, Sung BJ, et al. 2004. *EMBO J.* 23:2185–95
26. Reményi A, Good MC, Bhattacharyya RP, Lim WA. 2005. *Mol. Cell*. 20:951–62
27. Grewal S, Molina DM, Bardwell L. 2006. *Cell Signal.* 18:123–34
28. Kusari AB, Molina DM, Sabbagh W Jr, Lau CS, Bardwell L. 2004. *J. Cell. Biol.* 164:267–77
29. Ho DT, Bardwell AJ, Abdollahi M, Bardwell L. 2003. *J. Biol. Chem.* 278:32662–72
30. Nguyen AN, Ikner AD, Shiozaki M, Warren SM, Shiozaki K. 2002. *Mol. Biol. Cell* 13:2651–63
31. Tanoue T, Maeda R, Adachi M, Nishida E. 2001. *EMBO J.* 20:466–79
31a. Bardwell L, Thorner J. 1996. *Trends Biochem. Sci*. 21:373–74
32. Barsyte-Lovejoy D, Galanis A, Sharrocks AD. 2002. *J. Biol. Chem.* 277:9896–903
33. Lee T, Hoofnagle AN, Kabuyama Y, Stroud J, Min XS, et al. 2004. *Mol. Cell* 14:43–55
34. Fantz DA, Jacobs D, Glossip D, Kornfeld K. 2001. *J. Biol. Chem.* 276:27256–65
34a. Bardwell AJ, Flatauer LJ, Matsukuma K, Thorner J, Bardwell L. 2001. *J. Biol. Chem*. 276:10374–86
35. Takekawa M, Tatebayashi K, Saito H. 2005. *Mol. Cell* 18:295–306
36. Tatebayashi K, Takekawa M, Saito H. 2003. *EMBO J.* 22:3624–34
37. Biondi RM, Cheung PC, Casamayor A, Deak M, Currie RA, Alessi DR. 2000. *EMBO J.* 19:979–88
38. Frodin M, Jensen CJ, Merienne K, Gammeltoft S. 2000. *EMBO J.* 19:2924–34
39. Biondi RM, Komander D, Thomas CC, Lizcano JM, Deak M, et al. 2002. *EMBO J.* 21:4219–28
40. Fiol CJ, Haseman JH, Wang YH, Roach PJ, Roeske RW, et al. 1988. *Arch. Biochem. Biophys.* 267:797–802
41. Dajani R, Fraser E, Roe SM, Young N, Good V, et al. 2001. *Cell* 105:721–32
42. Frame S, Cohen P, Biondi RM. 2001. *Mol. Cell* 7:1321–27
42a. Bardwell L, Cook JG, Chang EC, Cairns BR, Thorner J. 1996. *Mol. Cell. Biol*. 16:3637–50
43. Breitkreutz A, Boucher L, Tyers M. 2001. *Curr. Biol.* 11:1266–71
44. Bridges D, Moorhead GB. 2004. *Science STKE* 2004:re10
45. Durocher D, Jackson SP. 2002. *FEBS Lett.* 513:58–66
46. Glover JN, Williams RS, Lee MS. 2004. *Trends Biochem. Sci.* 29:579–85
47. Harris BZ, Lim WA. 2001. *J. Cell Sci.* 114:3219–31

48. Lemmon MA, Ferguson KM. 2000. *Biochem. J.* 350(Pt. 1):1–18
49. Pawson T, Gish GD, Nash P. 2001. *Trends Cell Biol.* 11:504–11
50. Schlessinger J, Lemmon MA. 2003. *Science STKE* 2003:re12
51. Sondermann H, Kuriyan J. 2005. *Cell* 121:158–60
52. Zarrinpar A, Bhattacharyya RP, Lim WA. 2003. *Science STKE* 2003:re8
53. Marengere LE, Pawson T. 1994. *J. Cell Sci. Suppl.* 18:97–104
54. Dueber JE, Yeh BJ, Bhattacharyya RP, Lim WA. 2004. *Curr. Opin. Struct. Biol.* 14:690–99
55. Lim WA. 2002. *Curr. Opin. Struct. Biol.* 12:61–68
56. Prehoda KE, Scott JA, Mullins RD, Lim WA. 2000. *Science* 290:801–6
57. Rohatgi R, Ho HY, Kirschner MW. 2000. *J. Cell Biol.* 150:1299–310
58. Sicheri F, Moarefi I, Kuriyan J. 1997. *Nature* 385:602–9
59. Williams JC, Weijland A, Gonfloni S, Thompson A, Courtneidge SA, et al. 1997. *J. Mol. Biol.* 274:757–75
60. Xu W, Harrison SC, Eck MJ. 1997. *Nature* 385:595–602
61. Letunic I, Copley RR, Schmidt S, Ciccarelli FD, Doerks T, et al. 2004. *Nucleic Acids Res.* 32:D142–44
62. Schultz J, Milpetz F, Bork P, Ponting CP. 1998. *Proc. Natl. Acad. Sci. USA* 95:5857–64
63. Berry DM, Nash P, Liu SK, Pawson T, McGlade CJ. 2002. *Curr. Biol.* 12:1336–41
64. Harkiolaki M, Lewitzky M, Gilbert RJ, Jones EY, Bourette RP, et al. 2003. *EMBO J.* 22:2571–82
65. Liu Q, Berry D, Nash P, Pawson T, McGlade CJ, Li SC. 2003. *Mol. Cell* 11:471–81
66. Lewitzky M, Harkiolaki M, Domart MC, Jones EY, Feller SM. 2004. *J. Biol. Chem.* 279:28724–32
67. Groemping Y, Lapouge K, Smerdon SJ, Rittinger K. 2003. *Cell* 113:343–55
68. Li SC, Gish G, Yang D, Coffey AJ, Forman-Kay JD, et al. 1999. *Curr. Biol.* 9:1355–62
69. Benes CH, Wu N, Elia AE, Dharia T, Cantley LC, Soltoff SP. 2005. *Cell* 121:271–80
70. Bu JY, Shaw AS, Chan AC. 1995. *Proc. Natl. Acad. Sci. USA* 92:5106–10
71. Iwashima M, Irving BA, van Oers NS, Chan AC, Weiss A. 1994. *Science* 263:1136–39
72. Pluskey S, Wandless TJ, Walsh CT, Shoelson SE. 1995. *J. Biol. Chem.* 270:2897–900
73. Zarrinpar A, Park SH, Lim WA. 2003. *Nature* 426:676–80
74. Landgraf C, Panni S, Montecchi-Palazzi L, Castagnoli L, Schneider-Mergener J, et al. 2004. *PLoS Biol.* 2:e14
75. King N, Hittinger CT, Carroll SB. 2003. *Science* 301:361–63
76. Pawson T, Scott JD. 1997. *Science* 278:2075–80
77. Lowe ED, Tews I, Cheng KY, Brown NR, Gul S, et al. 2002. *Biochemistry* 41:15625–34
78. Loog M, Morgan DO. 2005. *Nature* 434:104–8
79. Burack WR, Shaw AS. 2000. *Curr. Opin. Cell. Biol.* 12:211–16
80. Elion EA. 2001. *J. Cell Sci.* 114:3967–78
81. Garrington TP, Johnson GL. 1999. *Curr. Opin. Cell. Biol.* 11:211–18
82. Morrison DK, Davis RJ. 2003. *Annu. Rev. Cell. Dev. Biol.* 19:91–118
83. Whitmarsh AJ, Davis RJ. 1998. *Trends Biochem. Sci.* 23:481–85
84. Wong W, Scott JD. 2004. *Nat. Rev. Mol. Cell Biol.* 5:959–70
85. Zarrinpar A, Bhattacharyya RP, Nittler MP, Lim WA. 2004. *Mol. Cell* 14:825–32
86. Choi KY, Satterberg B, Lyons DM, Elion EA. 1994. *Cell* 78:499–512
87. Marcus S, Polverino A, Barr M, Wigler M. 1994. *Proc. Natl. Acad. Sci. USA* 91:7762–66
88. Posas F, Saito H. 1997. *Science* 276:1702–5
89. Printen JA, Sprague GF Jr. 1994. *Genetics* 138:609–19
89a. Truckses DM, Bloomekatz JE, Thorner J. 2006. *Mol. Cell. Biol.* 26:912–28

89b. Harris K, Lamson RE, Nelson B, Hughes TR, Marton MJ, et al. 2001. *Curr. Biol.* 11:1815–24
90. Li W, Fan J, Woodley DT. 2001. *Oncogene* 20:6403–17
91. Zhang M, Wang W. 2003. *Acc. Chem. Res.* 36:530–38
92. Feliciello A, Gottesman ME, Avvedimento EV. 2001. *J. Mol. Biol.* 308:99–114
93. Muller J, Cacace AM, Lyons WE, McGill CB, Morrison DK. 2000. *Mol. Cell Biol.* 20:5529–39
94. Sierralta J, Mendoza C. 2004. *Brain Res. Brain Res. Rev.* 47:105–15
95. Raab M, da Silva AJ, Findell PR, Rudd CE. 1997. *Immunity* 6:155–64
96. Zhang W, Trible RP, Zhu M, Liu SK, McGlade CJ, Samelson LE. 2000. *J. Biol. Chem.* 275:23355–61
97. Muller J, Ory S, Copeland T, Piwnica-Worms H, Morrison DK. 2001. *Mol. Cell* 8:983–93
98. Nguyen A, Burack WR, Stock JL, Kortum R, Chaika OV, et al. 2002. *Mol. Cell Biol.* 22:3035–45
99. Ory S, Zhou M, Conrads TP, Veenstra TD, Morrison DK. 2003. *Curr. Biol.* 13:1356–64
100. Matheny SA, Chen C, Kortum RL, Razidlo GL, Lewis RE, White MA. 2004. *Nature* 427:256–60
101. Flotho A, Simpson DM, Qi M, Elion EA. 2004. *J. Biol. Chem.* 279:47391–401
101a. Bhattacharyya RP, Reményi A, Good MC, Bashor CJ, Falick AM, Lim WA. 2006. *Science* 311:822–26
102. Whitmarsh AJ, Cavanagh J, Tournier C, Yasuda J, Davis RJ. 1998. *Science* 281:1671–74
103. Willoughby EA, Perkins GR, Collins MK, Whitmarsh AJ. 2003. *J. Biol. Chem.* 278:10731–36
104. Willoughby EA, Collins MK. 2005. *J. Biol. Chem.* 280:25651–58
105. Coghlan VM, Perrino BA, Howard M, Langeberg LK, Hicks JB, et al. 1995. *Science* 267:108–11
106. Klauck TM, Faux MC, Labudda K, Langeberg LK, Jaken S, Scott JD. 1996. *Science* 271:1589–92
107. Schillace RV, Scott JD. 1999. *Curr. Biol.* 9:321–24
108. Dodge KL, Khouangsathiene S, Kapiloff MS, Mouton R, Hill EV, et al. 2001. *EMBO J.* 20:1921–30
109. McCahill A, McSorley T, Huston E, Hill EV, Lynch MJ, et al. 2005. *Cell Signal* 17:1158–73
110. Scott K, Zuker CS. 1998. *Nature* 395:805–8
111. Tsunoda S, Sierralta J, Sun Y, Bodner R, Suzuki E, et al. 1997. *Nature* 388:243–49
112. Tsunoda S, Zuker CS. 1999. *Cell Calcium* 26:165–71
113. Benner SA, Sismour AM. 2005. *Nat. Rev. Genet.* 6:533–43
114. Pawson T, Linding R. 2005. *FEBS Lett.* 579:1808–14
115. Park SH, Zarrinpar A, Lim WA. 2003. *Science* 299:1061–64
116. Howard PL, Chia MC, Del Rizzo S, Liu FF, Pawson T. 2003. *Proc. Natl. Acad. Sci. USA* 100:11267–72
117. Dueber JE, Yeh BJ, Chak K, Lim WA. 2003. *Science* 301:1904–8
118. McDonald C, Vacratsis PO, Bliska JB, Dixon JE. 2003. *J. Biol. Chem.* 278:18514–23
119. Cooper JA, Gould KL, Cartwright CA, Hunter T. 1986. *Science* 231:1431–34
120. Kmiecik TE, Shalloway D. 1987. *Cell* 49:65–73
121. Piwnica-Worms H, Saunders KB, Roberts TM, Smith AE, Cheng SH. 1987. *Cell* 49:75–82

122. Jackson PK, Eldridge AG, Freed E, Furstenthal L, Hsu JY, et al. 2000. *Trends Cell Biol.* 10:429–39
123. Vale RD, Milligan RA. 2000. *Science* 288:88–95
124. Csete M, Doyle J. 2004. *Trends Biotechnol.* 22:446–50
124a. Kashtan N, Alon U. 2005. *Proc. Natl. Acad. Sci. USA* 102:13773–78
125. Cohen-Gihon I, Lancet D, Yanai I. 2005. *Trends Genet.* 21:210–13
126. Bae KH, Kwon YD, Shin HC, Hwang MS, Ryu EH, et al. 2003. *Nat. Biotechnol.* 21:275–80
127. Dreier B, Beerli RR, Segal DJ, Flippin JD, Barbas CF 3rd. 2001. *J. Biol. Chem.* 276:29466–78
128. Dreier B, Fuller RP, Segal DJ, Lund C, Blancafort P, et al. 2005. *J. Biol. Chem.* 280:35588–97
129. Dreier B, Segal DJ, Barbas CF 3rd. 2000. *J. Mol. Biol.* 303:489–502
130. Segal DJ, Dreier B, Beerli RR, Barbas CF 3rd. 1999. *Proc. Natl. Acad. Sci. USA* 96:2758–63
130a. Rebar EJ, Pabo CO. 1994. *Science* 263:671–73
131. Klug A. 1999. *J. Mol. Biol.* 293:215–18
132. Segal DJ, Beerli RR, Blancafort P, Dreier B, Effertz K, et al. 2003. *Biochemistry* 42:2137–48
133. Beerli RR, Segal DJ, Dreier B, Barbas CF 3rd. 1998. *Proc. Natl. Acad. Sci. USA* 95:14628–33
134. Alwin S, Gere MB, Guhl E, Effertz K, Barbas CF 3rd, et al. 2005. *Mol. Ther.* 12:610–17
135. Beerli RR, Dreier B, Barbas CF 3rd. 2000. *Proc. Natl. Acad. Sci. USA* 97:1495–500
136. Blancafort P, Chen EI, Gonzalez B, Bergquist S, Zijlstra A, et al. 2005. *Proc. Natl. Acad. Sci. USA* 102:11716–21
137. Klug A. 2005. *FEBS Lett.* 579:892–94
138. Segal DJ, Goncalves J, Eberhardy S, Swan CH, Torbett BE, et al. 2004. *J. Biol. Chem.* 279:14509–19
139. Tan W, Zhu K, Segal DJ, Barbas CF 3rd, Chow SA. 2004. *J. Virol.* 78:1301–13
139a. Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, et al. 2005. *Nature* 435:646–51
140. Bax B, Carter PS, Lewis C, Guy AR, Bridges A, et al. 2001. *Structure* 9:1143–52
141. Dajani R, Fraser E, Roe SM, Yeo M, Good VM, et al. 2003. *EMBO J.* 22:494–501
142. Stradal TE, Rottner K, Disanza A, Confalonieri S, Innocenti M, Scita G. 2004. *Trends Cell Biol.* 14:303–11
143. Lei M, Lu W, Meng W, Parrini MC, Eck MJ, et al. 2000. *Cell* 102:387–97
144. Rohrschneider LR, Fuller JF, Wolf I, Liu Y, Lucas DM. 2000. *Genes Dev.* 14:505–20
145. Murata Y, Iwasaki H, Sasaki M, Inaba K, Okamura Y. 2005. *Nature* 435:1239–43
146. Sands Z, Grottesi A, Sansom MS. 2005. *Curr. Biol.* 15:R44–47

Annual Review
of Biochemistry

Volume 75, 2006

Contents