Reading protein modifications with interaction domains

Bruce T. Seet, Ivan Dikic‡, Ming-Ming Zhou§ and Tony Pawson*||*

Abstract | Proteins are controlled by a vast and dynamic array of post-translational modifications, many of which create binding sites for specific protein-interaction domains. We propose that these domains, working together, read the state of the proteome and therefore couple post-translational modifications to cellular organization. We also identify common strategies through which modification-dependent interactions synergize to regulate cell behaviour.

Phosphoinositide
A phosphorylated derivative of the glycerolipid phosphatidyl-inositol.

Allosteric regulation
The regulation of a protein's activity through a conformational change that is induced by the binding of a ligand or the addition of a post-translational modification at a region other than the substrate-binding site.

**Samuel Lunenfeld Research Institute, Mount Sinai Hospital,600 University Avenue, Toronto, Ontario M5G 1X5, Canada.*
‡Institute of Biochemistry II, Goethe University School of Medicine, University Hospital, Building 75, Theodour-Stern-Kai 7, D-60590 Frankfurt, Germany.
§Department of Molecular Physiology and Biophysics, Mount Sinai School of Medicine, New York University, One Gustave L. Levy Place, New York, New York 10029, USA.
||Department of Medical Genetics and Microbiology, University of Toronto, Toronto, Ontario M5S 1A8, Canada.
Correspondence to T.P.
e-mail: pawson@mshri.on.ca
doi:10.1038/nrm1960

The proteins that underlie the behaviour of any cell type are produced by selective gene expression and expanded by alternative RNA splicing. However, a cell is not a static entity — it continuously responds to cues from its external and internal environments, such as growth factors, ligands on the surface of adjacent cells, the extracellular matrix, electrical excitation, shear stress, cell-cycle checkpoints, DNA damage, oxygen tension and nutrient status. The cellular response to changing conditions is frequently mediated by the reversible covalent modification of existing molecules, for which the prototype in eukaryotic cells is protein phosphorylation of the hydroxyamino acids Tyr, Ser and Thr. However, proteins can be altered by a diverse set of post-translational modifications (PTMs), including the methylation of Arg residues, the methylation, acetylation, ubiquitylation or sumoylation of Lys residues, or prolyl hydroxylation, among many others[1] (FIG. 1).

Phosphorylation of a single residue is unitary, because only one phosphate moiety can be added to an amino acid. However, an individual polypeptide chain can potentially be phosphorylated at several sites, thereby producing many phosphoisoforms, each of which can potentially have a distinct biological activity[1]. Other PTMs can be yet more diverse. For example, a Lys residue can be mono-, di- or trimethylated, or mono- or polyubiquitylated. Furthermore, the same polypeptide chain can be modified by different classes of PTM, which generates an even larger number of possible variants[1]. Such dynamic and combinatorial modifications are not restricted to proteins. For example, phosphoinositides are modified by the variable phosphorylation of their inositol headgroup, which produces several distinct phosphoinositide species in membranes that each have different biological properties[2].

A crucial issue for understanding cellular regulation is to define how this vast and constantly varying spectrum of protein and lipid modifications is interpreted to produce alterations in cellular phenotype. One mechanism by which PTMs can alter a protein's function is to directly induce a new conformational state, and this is exemplified by the phosphorylation of the activation segment of protein kinases, which reorganizes the active site of the kinases into a productive configuration. However, such conventional allosteric regulation usually involves a highly specialized set of molecular interactions, which are specifically tailored to a particular polypeptide, and therefore cannot readily be incorporated into one type of protein from another during the course of evolution. This limits this type of regulation from being a general solution to the problem of interpreting PTMs on a proteome-wide scale.

Instead, we argue that the cell uses modular protein-interaction domains as a broad device to decode the state of its protein and lipid modifications, with different domains being dedicated to the selective recognition of distinct PTMs or phospholipids. Consistent with the generality of this mechanism, the human proteome has many classes of interaction domain that collectively recognize numerous different protein or lipid modifications, with each domain being represented tens or hundreds of times in cellular proteins[3]. The cassette-like structure of interaction domains provides a simple way of coupling a given protein to a particular PTM[4] (FIG. 1), and is well suited to the coordinated recognition and deciphering of several distinct PTMs in combination. Furthermore, although these interactions seem simple in isolation, they can nonetheless produce complex biochemical responses in the context of multidomain proteins and multiprotein assemblies. This has prompted an analogy to the construction of languages from simpler units of

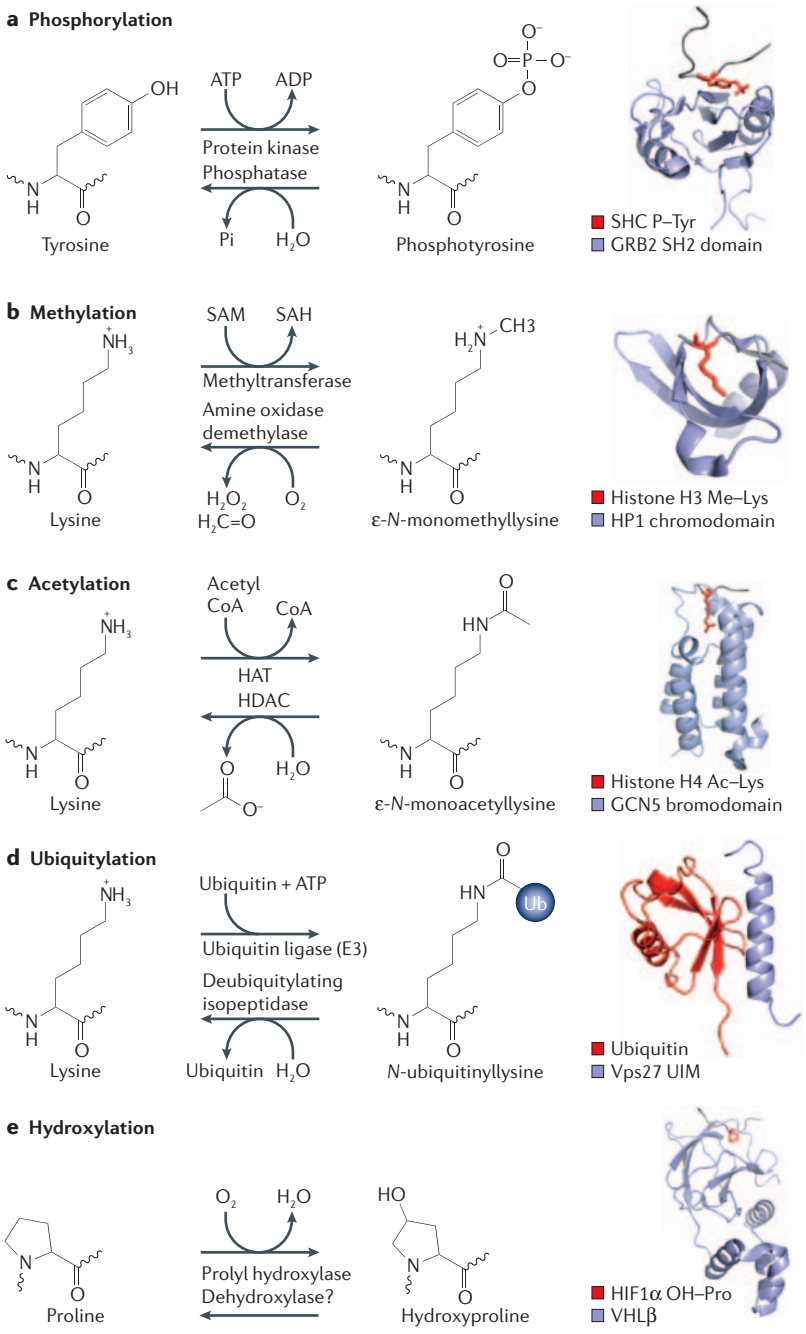

Figure 1 | **Example post-translational modification reactions and structures of protein-interaction-domain–ligand complexes.** Various amino-acid side chains can be modified by, for example: phosphorylation (**a**); methylation (**b**); acetylation (**c**); ubiquitylation (**d**); and hydroxylation (**e**). The enzymes that are involved in adding and removing these post-translational modifications are shown on the reaction arrows. The structures on the far right show examples of protein-interaction domains (pale purple) in complex with their respective ligands (red). The structures were obtained from the Protein Data Bank (accession codes 1JYR, 1Q3L, 1E6I, 1Q0W and 1LM8 for parts **a**–**e**, respectively) and were manipulated using Pymol (Delano Scientific). GCN5, general control of amino-acid-synthesis protein-5; GRB2, growth-factor-receptor-bound protein-2; HAT, histone acetyltransferase; HDAC, histone deacetylase; HIF1α, hypoxia-inducible factor-1α; HP1, heterochromatin protein-1; Me–Lys, methylated lysine; OH–Pro, hydroxylated proline; Pi, inorganic phosphate; P–Tyr, phosphotyrosine; SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine; SH2, Src-homology-2; SHC, SH2-domain-containing transforming protein; Ub, ubiquitin; UIM, ubiquitin-interacting motif; VHLβ, von Hippel–Lindau protein-β; Vps27, vacuolar protein sorting protein-27.

words and phrases, and has engendered the hypothesis that protein domains represent a basic syntactic unit of cellular organization[5]. In this article, we briefly discuss the common ways in which PTMs and interaction domains synergize to regulate cellular processes, and we provide specific examples that involve phosphorylation, methylation, acetylation, ubiquitylation and sumoylation (FIGS 1,2). This is not intended to be a comprehensive analysis, but rather aims to highlight the general strategies through which PTMs exert their effects.

PTM-dependent interactions: common strategies

In the following subsections we briefly discuss the common mechanisms by which PTM-dependent interactions regulate cellular processes.

PTM-induced interactions. Interaction domains often recognize short peptide motifs that are embedded in target proteins, but do not bind stably until the peptide has acquired an appropriate PTM (FIGS 1,2a). Such domains usually have a conserved binding pocket for the modified residue and a more variable surface that selectively engages the flanking amino acids, and thereby distinguishes between different peptide motifs with the same PTM[6–9]. Both the domains and the peptide motifs that they recognize are modular in design and can therefore, in principle, be incorporated into many different proteins.

Cooperative interactions and multi-site PTMs. PTM-dependent interactions can be cooperative, such that a signal is only generated after two or more sites on the same protein have been modified. This can be achieved in various ways. First, a doubly modified motif can be recognized in an obligatory fashion by two tandem interaction domains, as is the case for the recognition of two phosphotyrosine (pTyr)-containing motifs by the tandem Src-homology-2 (SH2) domains of the ZAP-70 (ζ-chain (T-cell receptor)-associated protein kinase 70 kDa) protein tyrosine kinase[10] (FIG. 2b). This can increase both the affinity and the specificity of the interaction. Second, a single domain can possess two binding pockets for the modified residues. For example, the single SH2 domain of the APS protein (adaptor molecule containing pleckstrin-homology (PH) and SH2 domains protein) binds to two pTyr residues in the activation loop of the insulin receptor kinase; furthermore, two APS SH2 domains form a non-covalent dimer, which potentially stabilizes the activated receptor[11]. Third, a domain with a single binding pocket can bind specifically to a protein that carries several modifications. For example, the WD40-repeat domain of the *Saccharomyces cerevisiae* protein Cdc4 (cell-division cycle-4) only binds to its target, Sic1 (substrate inhibitor of cyclin-dependent protein kinase-1), when the target has been phosphorylated during the G1 phase of the cell cycle on at least six Ser/Thr residues (FIG. 2c). As Cdc4 is the substrate-targeting subunit of an SCF (Skp1–Cul1–F-box) E3 ubiquitin ligase complex, phosphorylation of Sic1 leads to its polyubiquitylation and degradation by the proteasome. This, in turn, is necessary for the onset of DNA synthesis, because Sic1

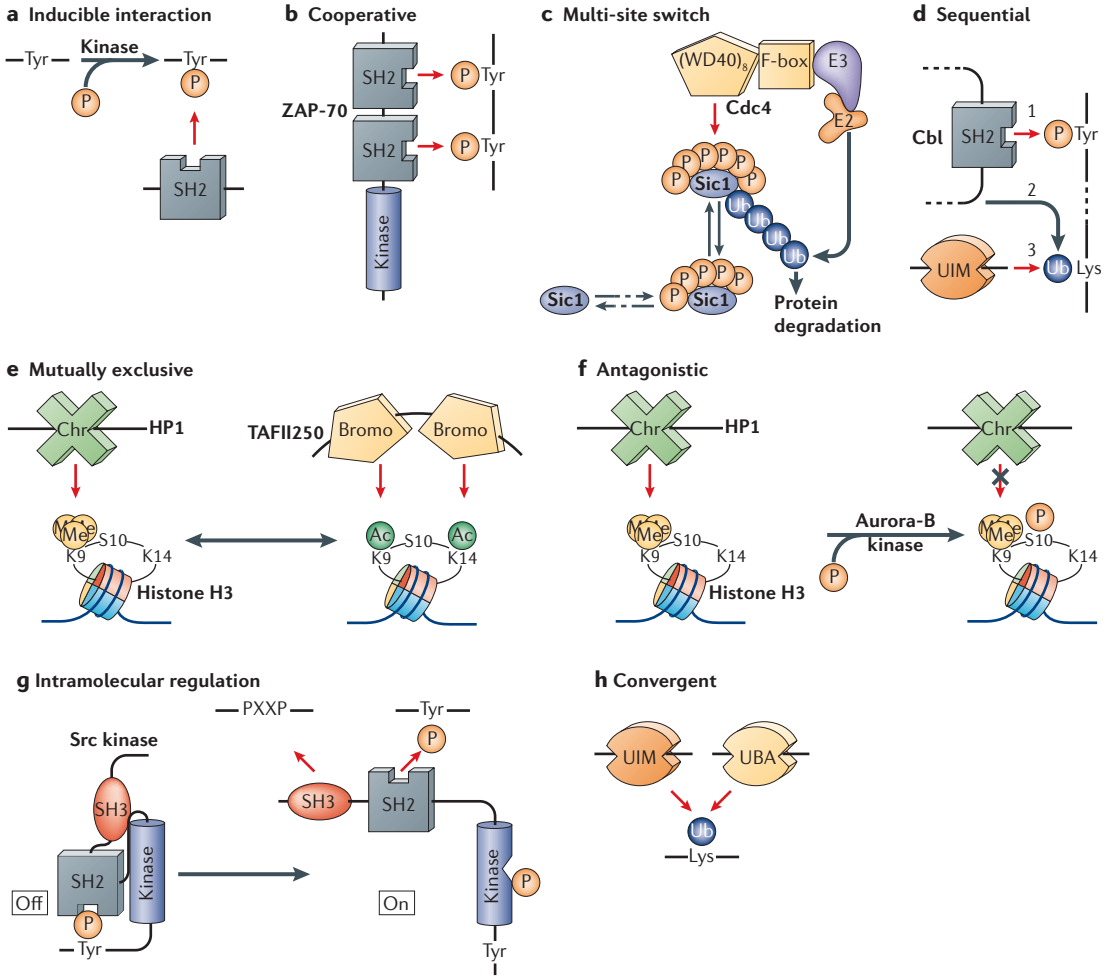

Figure 2 | **Examples of the regulatory properties of post-translational modifications and their interaction domains.** Examples of the regulatory modes that can be achieved using interaction-domain–post-translational-modification interactions. **a** | Inducible interactions. A peptide with a phosphorylated (P) Tyr residue is recognized by its target Src-homology-2 (SH2) domain. **b** | Cooperative interactions. Tandem SH2 domains in ZAP-70 (ζ-chain (T-cell receptor)-associated protein kinase 70 kDa) bind in an obligatory fashion to two phosphotyrosine (pTyr) sites in ITAMs (immunoreceptor Tyr-based activation motifs). **c** | Multi-site switch. Sic1 (substrate inhibitor of cyclin-dependent protein kinase-1) requires phosphorylation on at least six sites before it can bind to the SCF (Skp1–Cul1–F-box) E3 ubiquitin (Ub)-ligase subunit Cdc4 (cell-division cycle-4), which is a WD40-repeat-containing protein. This switch-like interaction leads to Sic1 ubiquitylation and controls cell-cycle progression in *Saccharomyces cerevisiae*. **d** | Sequential interactions. The recruitment of the SH2 domain of the E3 ubiquitin ligase Cbl (Casitas B-lineage lymphoma proto-oncogene) to a pTyr site (step 1) precedes, and is required for, the transfer of ubiquitin to a target protein (step 2). The target protein is subsequently recognized by ubiquitin-interacting motif (UIM)-containing proteins (step 3). **e** | Mutually exclusive interactions. When Lys9 of histone H3 is trimethylated (Me), it binds to the chromodomain (Chr) of heterochromatin protein-1 (HP1), whereas histone H3 that is acetylated (Ac) at Lys9 and Lys14 binds to the bromodomains (Bromo) of TAFII250 (TATA-binding protein-associated factor-II250). The acetylation and methylation of Lys9 cannot occur simultaneously. **f** | Antagonistic modifications. Trimethylated Lys9 of histone H3 binds to the HP1 chromodomain, and this interaction is antagonized by the phosphorylation of the neighbouring Ser10 residue by Aurora-B kinase. **g** | Intramolecular interactions. Src kinase is autoinhibited through an interaction between its SH2 domain and a pTyr in the C-terminal tail of the protein ('off' state). The release of this intramolecular interaction by dephosphorylation of the tail, or through the presence of a combination of activating ligands (such as Pro-rich or pTyr ligands), results in an active kinase ('on' state). **h** | Convergent interactions. A single post-translational-modification can be recognized by different types of protein-interaction domain. For example, ubiquitin binds to UIMs and ubiquitin-associated (UBA) domains, as well as to other ubiquitin-binding domains. E2, ubiquitin-conjugating enzyme; SH3, Src-homology-3.

Src-homology-2 (SH2) domain
A 100-residue domain that binds to particular phosphorylated Tyr sequences in proteins.

Pleckstrin-homology domain
(PH domain). The PH domain is ~120 amino acids. It typically interacts with various phosphoinositides, and is thereby involved in targeting proteins to membranes.

WD40-repeat domain
A repeat sequence of 40–60 amino acids that usually ends with Trp and Asp (WD). Consecutive repeats fold into a circular β-propeller structure.

E3 ubiquitin ligase
An enzyme that functions with a ubiquitin-conjugating enzyme (E2) to link one or more ubiquitin molecules to a target protein, which marks the protein for subsequent recognition by ubiquitin-binding domains. SCF-type ubiquitin ligases are one of the principal classes of E3 ligase. They are complexes that consist of SKP1, cullin and F-box proteins.

Proteasome
A large multiprotein complex that is responsible for degrading intracellular proteins that have been tagged for destruction by the addition of ubiquitin.

Breast-cancer-susceptibility protein-1 C-terminal domain
(BRCT domain). The BRCT domain is 90–100 amino acids and occurs either as a single element or as multiple repeats. It binds to phosphopeptides in several proteins that are involved in DNA-damage response and DNA repair.

inhibits S-phase cyclin-dependent kinase (CDK) activity. The requirement for multi-site phosphorylation of Sic1 sets a threshold for the level of G1 CDK activity that is necessary to promote Sic1 destruction, and it precipitates a switch-like entry into S phase[12,13].

In a complementary scenario, two tandem domains can fold into a common structure that recognizes a singly modified peptide, as has been seen for tandem BRCT (breast-cancer-susceptibility protein-1 C-terminal) domains that together bind to a phosphoserine (pSer)-containing motif[14].

Table 1 | **Phosphoserine/phosphothreonine-binding domains***

Domain/protein	Consensus-binding motif	Functions
14-3-3	R-S-X-pS-X-P	Survival, cell cycle, cytoskeleton, metabolism, signalling
TPR repeat	pS-Q	Nonsense-mediated decay
FHA	pT-X-X-D	Cell cycle, DNA repair, transport
MH2	pS-X-pS	TGFβ signalling
WW	pS-P	*Cis–trans* prolyl isomerization
WD40	L-P-pT-P	Cell cycle, ubiquitylation
BRCT	pS-X-X-Y	DNA-damage response, cell cycle
Polo box	S-pS-P	Cell cycle
FF	pS-Y/F/W-pS	Transcription, splicing
NT-Cgt1	$(Y\text{-}S\text{-}P\text{-}T\text{-}pS\text{-}P\text{-}S)_3$	RNA 5′-capping
SRI	$(Y\text{-}pS\text{-}P\text{-}T\text{-}pS\text{-}P\text{-}S)_2$	Histone methylation
β-arrestin	pS/pT…pS/pT…pS/pT	GPCR downregulation and signalling
Arm repeat	pS-pS-L-pS-A-L-pS	β-catenin signalling

*Domain families are indicated, together with examples of consensus phosphoserine/phosphothreonine (pSer/pThr) motifs that specific family members recognize and cellular processes that are regulated by these interactions. In the case of 14-3-3, the entire protein is composed of the pSer/pThr-binding fold. The indicated consensus motifs are not exhaustive. Some families are largely dedicated to pSer/pThr recognition (for example, 14-3-3 proteins), whereas others have only a few members that are known to be involved in phosphopeptide binding (for example, WW and WD40 domains). In the case of tetratricopeptide repeat (TPR)-, WD40-domain- and armadillo (Arm)-repeat-containing proteins, the phosphopeptide-binding domains are composed of several peptide repeats. β-arrestin is a scaffolding protein that binds preferentially to activated G-protein-coupled receptors (GPCRs) following multi-site phosphorylation in the GPCR C-terminal tail. FF domains can also bind to motifs that contain acidic residues (phosphorylation is not strictly required for ligand binding). BRCT, breast-cancer-susceptibility protein-1 C-terminal; FHA, forkhead-associated; MH2, MAD-homology-2; NT-Cgt1, nucleotidyltransferase domain of the mRNA-capping enzyme Cgt1 subunit; pS, phosphoserine; pT, phosphothreonine; SRI, Set2 Rpb1 interacting; TGFβ, transforming growth factor-β; X, any amino acid.

Sequential PTM-dependent interactions. One PTM-dependent protein–protein interaction can be required for a subsequent modification and interaction (FIG. 2d). Protein phosphorylation, for example, can create a binding site for the substrate-targeting domain of an E3 ubiquitin ligase[13], as noted above. The phosphorylated protein is then ubiquitylated and is consequently recognized by the ubiquitin-binding domain (UBD) of a protein that regulates events such as endocytosis or proteolysis[15] (see later). In the case of human CDC4 (also known as FBW7), its WD40-repeat domain binds to phosphothreonine (pThr)-containing sites on cyclin E, a regulatory subunit of the cell-cycle kinase CDK2. This leads to cyclin-E polyubiquitylation, its recruitment to a UBD on the proteasome and its degradation[16–18] (FIG. 3). Loss-of-function mutations that ablate the binding of CDC4 to phosphorylated cyclin E, including substitutions of the Arg residues in the WD40-repeat domain that normally coordinate the phosphorylated Thr residue, yield aberrantly high levels of cyclin E that are associated with some human cancers[17,19].

Along the lines of sequential interactions, a specific type of PTM-dependent interaction (that is, pTyr–SH2-domain or ubiquitin–UBD) can be used in a reiterated fashion to generate more extended signalling networks[20]. For example, signalling downstream of the T-cell receptor (TCR)–ZAP-70 complex involves a network of tyrosine kinases, scaffolds and effectors that is assembled through numerous interdependent pTyr–SH2-domain interactions[21].

Mutually exclusive PTMs and interactions. PTMs, and the creation of binding sites for protein-interaction domains, can be mutually exclusive (FIG. 2e). For example, one Lys residue can potentially be acetylated, methylated, ubiquitylated or sumoylated. However, these different PTMs cannot occur simultaneously, and the binding properties of the Lys site therefore vary depending on the type of modification that it receives.

Antagonistic PTMs and interactions. A PTM that is attached to one amino acid can antagonize the ability of an adjacent modified residue to recruit a binding partner (FIG. 2f). For example, the trimethylation of Lys9 in the N-terminal tail of histone H3 recruits the chromodomain of heterochromatin protein-1 (HP1), which modifies chromatin structure to repress gene expression. During mitosis, the Aurora-B kinase phosphorylates Ser10 of histone H3 (a known 14-3-3 protein-binding site), which displaces HP1 even though Lys9 remains in a trimethylated state[22,23] (FIG. 2f). Dislodging HP1 might allow other regulatory proteins that are involved in chromosome organization to interact with chromatin during mitosis.

Intramolecular PTM-dependent interactions. A protein that sustains a PTM can undergo an intramolecular interaction if it also contains a protein-interaction domain that binds to the modified site (FIG. 2g). This, in turn, can block the ability of the protein-interaction domain to engage an exogenous ligand, and it can also elicit a conformational change that can inhibit the activity of a linked catalytic domain[24].

Convergent recognition of a PTM. A single PTM is frequently recognized by several different classes of interaction domain, as is indicated in FIG. 2h for ubiquitin recognition by different UBDs. In another example, there are numerous domain classes with entirely different folds that selectively recognize pSer/pThr-containing motifs[25], although they can discriminate between the modified sites on the basis of the flanking sequence (TABLE 1). The diversity of such binding domains might expand the ability of Ser/Thr phosphorylation to control regulatory pathways.

Protein phosphorylation

Protein phosphorylation and phosphorylation-dependent interactions provide examples of the principles that were highlighted in the preceding section, which we summarize in the following subsections.

pTyr-dependent interactions. Protein tyrosine kinases can activate intracellular pathways through the inducible recruitment of proteins with SH2 domains. SH2 domains typically bind to pTyr-containing peptide motifs of 5–8

Ubiquitin-binding domains
The collective term that is given to modular interaction domains that bind to ubiquitin.

Scaffold
A protein that supports the assembly of a multiprotein complex through interactions with other proteins.

Chromodomain
A protein domain that often binds to methylated Lys residues in target proteins.

14-3-3 proteins
A family of proteins that bind to phosphorylated Ser/Thr residues in a context-specific manner.

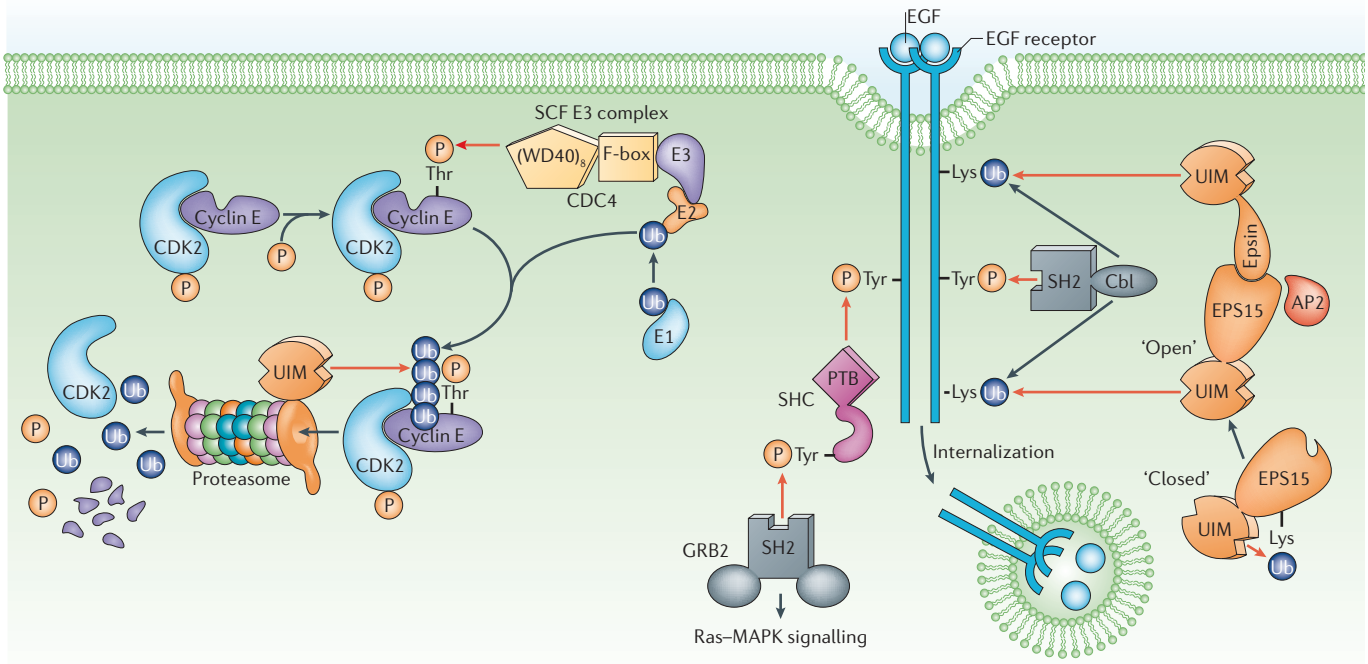

Figure 3 | **Networks of modification-dependent interactions regulate cellular processes.** Networks of phosphorylation (P) and ubiquitin (Ub)-dependent protein interactions regulate the cell cycle, as well as growth-factor-induced signalling and endocytic trafficking. Inducible post-translational modification (PTM)-dependent interactions are highlighted by red arrows. Sequential PTM-dependent interactions generate biological pathways, as can be seen in two examples. In the first example, mammalian cyclin E activates cyclin-dependent kinase-2 (CDK2) to promote the G1-to-S phase transition in the cell cycle. The phosphorylation of cyclin E on a Thr residue is required for its recognition by the WD40-repeat domain of the targeting subunit — CDC4 (cell-division cycle-4) — of an SCF (SKP1–CUL1–F-box) E3 ubiquitin-ligase complex. This interaction leads to the addition of a Lys48-linked polyubiquitin chain to cyclin E, which results in its subsequent recruitment to the proteasome by a ubiquitin-interacting motif (UIM) and its destruction. In the second example, epidermal growth factor (EGF) receptor autophosphorylation produces phosphotyrosine (pTyr) sites that can recruit the Src-homology-2 (SH2) domain of the E3 ubiquitin ligase Cbl (Casitas B-lineage lymphoma proto-oncogene). Cbl monoubiquitylates the receptor to provide docking sites for UIM-containing proteins such as epsin and EPS15 (EGF-receptor-pathway substrate-15), which are involved in endocytic membrane trafficking. The EPS15 UIM also undergoes an intramolecular interaction with a monoubiquitylated site, which regulates EPS15 activity (see the 'closed' and 'open' conformations). The binding of the phosphotyrosine-binding (PTB) domain of SHC (SH2-domain-containing transforming protein) to a pTyr site on the EGF receptor induces the phosphorylation of a Tyr residue in SHC. This pTyr, in turn, recruits the SH2 domain of GRB2 (growth-factor-receptor-bound protein-2) and activates downstream signals. The figure also shows the convergence of distinct interaction domains on the recognition of a single PTM (for example, SH2 and PTB domains for pTyr) and the modification of the same polypeptide by different types of PTM (for example, the multi-site phosphorylation and ubiquitylation of the EGF receptor). AP2, adaptor protein complex-2; E1, ubiquitin-activating enzyme; E2, ubiquitin-conjugating enzyme; MAPK, mitogen-activated protein kinase.

residues in a manner that depends on ligand phosphorylation and the identity of the flanking amino acids[26] (FIG. 2a). Activated receptor tyrosine kinases (RTKs), such as the β-platelet-derived growth factor receptor or the epidermal growth factor receptor (EGFR), become phosphorylated at multiple Tyr sites, and each of these sites selectively binds the SH2 domain of one or more cytoplasmic signalling proteins, which, in turn, activate specific intracellular signalling pathways[27–29] (FIGS 3,4a). Among the SH2-domain-containing proteins that are recruited to an autophosphorylated RTK such as EGFR is the Cbl (Casitas B-lineage lymphoma proto-oncogene) E3 ubiquitin ligase, which subsequently ubiquitylates the activated receptor on Lys residues and forms binding sites for proteins with ubiquitin-interacting motifs (UIMs) that are involved in receptor endocytosis[15] (FIGS 2d,3).

These data show that an activated RTK is modified at multiple sites by phosphorylation and ubiquitylation to yield motifs that are specifically recognized by the interaction domains of cytoplasmic effectors and regulators (FIG. 4a). As is discussed in the following sections, this strategy of multi-site protein modification followed by the selective recruitment of effectors through PTM-dependent interaction domains is more widely used by key regulatory proteins. For example, this mechanism is used by histones to modify chromatin organization (FIG. 4b) and by RNA polymerase II to recruit transcriptional regulators (FIG. 4c), although in these cases the principal PTMs that are recognized by interaction domains are acetylated/methylated Lys residues or pSer residues, respectively, as opposed to pTyr residues for RTKs (FIG. 4a).

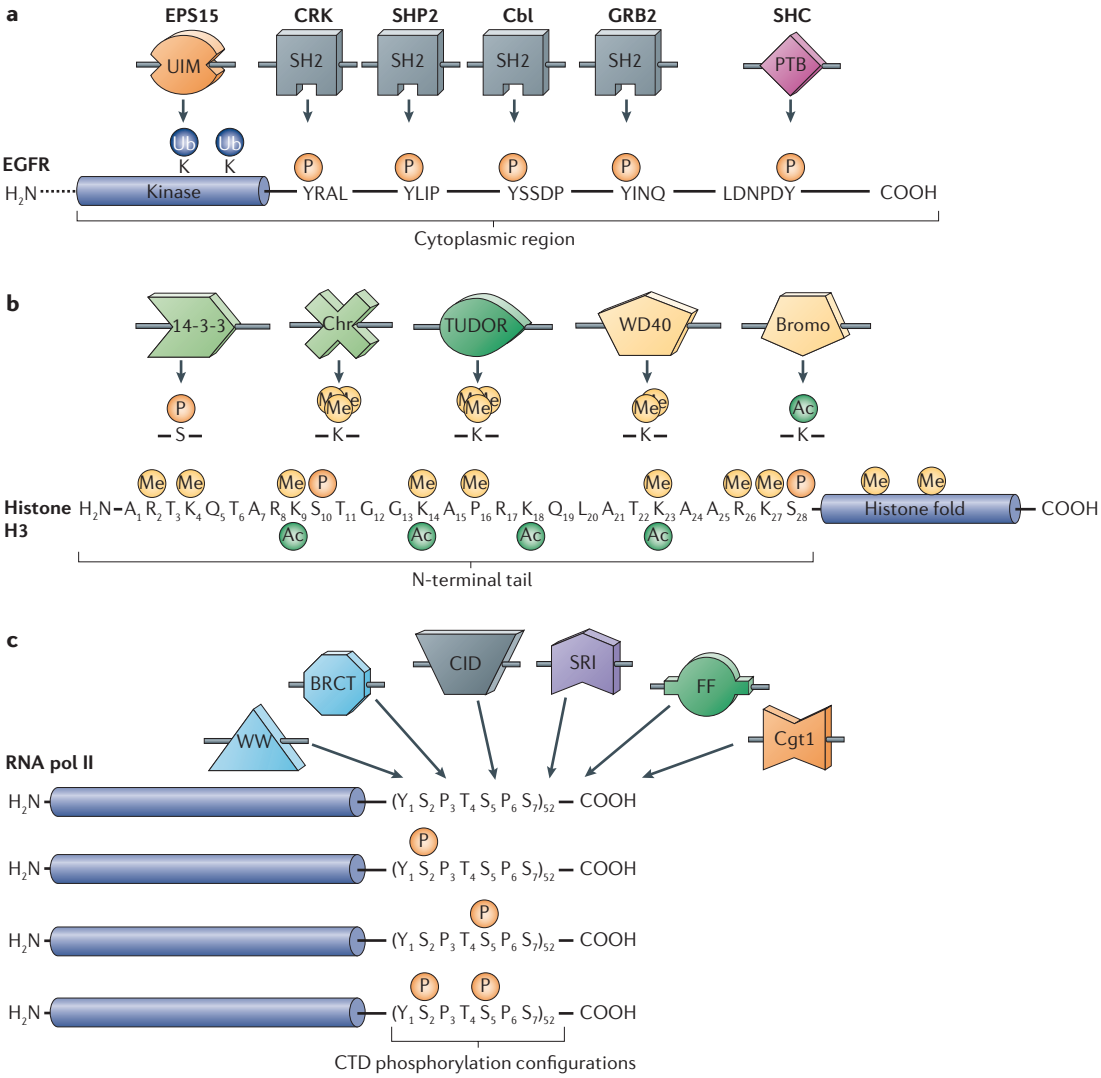

Figure 4 | **The combinatorial use of post-translational modifications. a** | The activated epidermal growth factor receptor (EGFR) is a tyrosine kinase that autophosphorylates at multiple sites, each of which selectively binds to one or more effectors that contain Src-homology-2 (SH2) or phosphotyrosine-binding (PTB) domains. One such protein, the E3 ubiquitin ligase Cbl (Casitas B-lineage lymphoma proto-oncogene), induces the multi-site ubiquitylation of EGFR, which creates binding sites for proteins with ubiquitin-interacting motifs (UIMs) that are involved in endocytosis (for example, EPS15 (EGF-receptor-pathway substrate-15)). Representative post-translational modifications and interacting proteins are indicated. The N-terminal extracellular and transmembrane regions of the receptor are not shown for clarity. **b** | Histone H3 has a flexible N-terminal tail that contains several sites that can be phosphorylated (P), methylated (Me) or acetylated (Ac), which creates binding sites for 14-3-3 proteins, chromodomains (Chr), TUDOR domains, a WD40-repeat domain and bromodomains (Bromo). These dynamic interactions determine events such as chromatin remodelling, gene expression or gene silencing. **c** | The C-terminal domain (CTD) of human RNA polymerase II (RNA pol II) contains 52 repeats of a heptad sequence (YSPTSPS). This sequence has several potential phosphorylation sites, of which Ser2 and Ser5 are best understood. Various phosphorylation patterns allow the CTD to recruit factors that contain WW, BRCT (breast-cancer-susceptibility protein-1 C-terminal), CID (CTD-interaction domain), SRI (Set2 Rpb1 interacting) or FF domains (the FF domains are tandemly arrayed, which might allow contacts with several phosphorylated CTD repeats). The guanyltransferase domain of the *Candida albicans* mRNA-capping enzyme Cgt1 subunit uses its nucleotidyltransferase domain to engage the Ser-phosphorylated CTD. GRB2, growth-factor-receptor-bound protein-2; SHC, SH2-domain-containing transforming protein; SHP2, Src-homology phosphatase-2; Ub, ubiquitin.

Src-family kinases
Kinases that belong to the Src family of tyrosine kinases, which is the largest of the non-receptor-tyrosine-kinase families. Members include Src, Yes, Fyn, Lck, Lyn, Blk, Hck, Fgr and Yrk.

As mentioned above for ZAP-70, SH2 domains can bind cooperatively to bisphosphorylated sites (FIG. 2b). The tandem SH2 domains of ZAP-70 selectively engage doubly phosphorylated ITAMs (immunoreceptor Tyr-based activation motifs) in the signalling subunits of the TCR[10]. Tyr phosphorylation can also promote the intramolecular regulation of SH2-domain-containing proteins (FIG. 2g). The prototypic example involves Src-family kinases, which are inactivated by the phosphorylation of a Tyr residue in their C-terminal tail.

This creates a binding site for the SH2 domain of Src that leads to an autoinhibited conformation, in which the kinase domain is locked in an inactive state and the SH2 and Src-homology-3 (SH3) domains are masked by their binding to internal ligands[24].

In addition to SH2 domains, subsets of PTB (phosphotyrosine-binding) and C2 (conserved region-2 of protein kinase C) domains have converged on pTyr recognition, even though they are structurally quite distinct from SH2 domains[30,31]. PTB and SH2 domains typically function sequentially. Docking proteins, such as IRS1 (insulin-receptor substrate-1) or SHC (SH2-domain-containing transforming protein), bind through their PTB domains to pTyr sites on activated RTKs and are then themselves phosphorylated at sites that recruit SH2-domain-containing proteins, which leads to the activation of signalling pathways[32,33] (FIG. 3).

pSer/pThr-dependent interactions. Ser/Thr phosphorylation is more prevalent than Tyr phosphorylation, and there is a correspondingly larger array of domains that selectively bind to pSer/pThr sites[25] (TABLE 1). However, many of the themes that are noted above can be discerned for pSer/pThr-binding proteins, such as 14-3-3 proteins. These proteins form non-covalent dimers, which can consequently bind to two pSer/pThr-containing peptides that have an appropriate consensus sequence[34]. As a result, 14-3-3 proteins can regulate the conformation and catalytic activity of multiphosphorylated enzymes, and can control the interactions and localizations of phosphorylated ligands[35].

Many proteins are phosphorylated at several Ser/Thr sites, which potentially produces combinatorial or cooperative effects. In the context of signal transduction, the MAD-homology-2 (MH2) domain of SMAD transcription factors preferentially binds to motifs that contain at least two pSer sites, as can be found in the juxtamembrane region of type-I transforming growth factor-β receptors or in the tail of receptor-regulated SMADs (R-SMADs)[36].

Strikingly, the C-terminal domain (CTD) of human RNA polymerase II has 52 repeats of a heptad sequence (YSPTSPS), each of which can be phosphorylated at Ser2 and/or Ser5. This yields four possible phosphoisoforms per repeat (that is, unphosphorylated, pSer2, pSer5 and pSer2–pSer5) (FIG. 4c). The phosphorylated CTD recruits factors that couple transcription to RNA modifications, such as the 5′-capping enzyme and 3′-processing and splicing factors. These proteins can bind to the CTD through pSer-recognition modules such as WW, BRCT, FF and SRI (Set2 Rpb1 interacting) domains or through CTD-interaction domains (CIDs), which preferentially bind to phosphorylated repeats but do not directly contact the Ser phosphates[37]. The RNA-polymerase-II CTD highlights the dynamic nature of multi-site PTMs and attendant protein–protein interactions, because the CTD repeats are initially phosphorylated at Ser5 during initiation, but shift to being phosphorylated at both Ser2 and Ser5 (and subsequently Ser2 alone) during elongation. The CTD can therefore recruit different binding partners at distinct stages of transcription owing to the fluctuating phosphorylation state of the heptad repeats. For example, during initiation, the guanyltransferase domain of the mRNA-capping enzyme Cgt1 subunit binds to three extended CTD repeats and selectively engages the pSer5 residues of the flanking repeats, with the central repeat being looped out[38]. As transcription proceeds, the SRI domain of the Set2 histone methyltransferase, which couples transcriptional elongation to the methylation of Lys36 of histone H3, preferentially binds to two consecutive heptad repeats in which both Ser2 and Ser5 are phosphorylated (REF. 39). Methylated Lys36 of histone H3, in turn, might regulate transcription by recruiting the chromodomain of a histone deacetylase complex[40], and this exemplifies a PTM-dependent interaction of the type that is discussed in the next section.

Protein methylation and acetylation

Peptide motifs that contain Lys residues can be methylated or acetylated, which can lead to their recognition by chromodomains or bromodomains, respectively. Such domains are found in proteins that regulate chromatin structure and gene expression. These Lys modifications are a particularly prominent feature of the flexible N- and C-terminal tails of histones, and they are important for coupling histones to changes in chromatin organization and the epigenetic control of gene expression (FIG. 4b). As with phosphorylation-dependent recognition, these interactions are rather simple, in the sense that a single chromodomain or bromodomain typically recognizes a suitably modified Lys residue in the context of a short peptide motif of defined sequence[8,41,42]. However, there are interesting and physiologically important subtleties that can be understood in the context of the general strategies for PTM-dependent interactions that were defined above. For example, similar to the tandem SH2 domains of ZAP-70, which bind cooperatively to bisphosporylated ITAMs, TAFII250 (TATA-binding protein-associated factor-II250) has two adjacent bromodomains that bind preferentially to doubly acetylated motifs on histone H3 (REF. 43) (FIG. 2e). In addition, similar to double BRCT domains, which fold together to bind to a single pSer-containing motif, two tandemly arranged chromodomains in the human protein CHD1 (chromodomain–helicase-domain–DNA-binding-domain protein-1) cooperatively interact with a single methylated Lys residue in histone H3 by forming a single peptide-binding groove at the interface between the domains[44].

Recent evidence indicates that some WD40, TUDOR and MBT (malignant brain tumour) domains, together with chromodomains, have converged on a common ability to recognize methylated Lys motifs[45]. For example, the WD40-repeat domain of WDR5 (WD-repeat protein-5) — a common subunit of histone H3 Lys4 methyltransferase — preferentially binds to dimethylated Lys4 in histone H3 (REF. 46). The propeller-like WD40-repeat domain of WDR5 selectively engages both the dimethylated Lys and the preceding N-terminal residues, notably an Arg residue at the –2 position (that is, the second residue N-terminal to the modified Lys) that fills the central channel of the domain[47]. These data illustrate

Src-homology-3 (SH3) domain
A protein sequence of ~50 amino acids that binds to Pro-rich regions of proteins. Some SH3 domains have been identified that bind to atypical non-Pro-based motifs.

Phosphotyrosine-binding domain
(PTB domain). A domain of 100–150 amino acids. Some PTB domains bind to specific phosphotyrosine sites, which usually have the consensus sequence Asn-Pro-X-pTyr (NPXpY).

Conserved region-2 of protein kinase C domain
(C2 domain). A domain that was originally found to bind to lipids in a Ca^{2+}-dependent manner. However, an exception has been identified in the C2 domain of protein kinase Cδ, which binds to specific phosphotyrosine-containing peptides.

WW domain
A protein domain of ~35 amino acids that binds to Pro-rich peptide motifs or, in some cases, to pSer/pThr-Pro motifs.

FF domain
A protein domain of 50–60 amino acids. FF domains are always arranged in tandem repeats and bind to acidic or phosphorylated peptide motifs.

SRI domain
(Set2 Rpb1 interacting domain). An ~100-amino-acid domain that is conserved among a number of putative Set2 homologues. The SRI domain of Set2 binds to the phosphorylated C-terminal domain of RNA polymerase II.

C-terminal-interaction domains
(CIDs). CIDs are domains that bind to phosphorylated heptad repeats in the C-terminal domain of RNA polymerase II.

two more general points. First, different types of interaction domain can engage a PTM through distinct structural mechanisms, which indicates that they have evolved separately to recognize the same modification. For example, chromodomains and TUDOR domains use conserved aromatic residues to envelop the ζ-methyl groups of a methylated Lys residue in a hydrophobic cage[8,42,48], whereas WD40 domains form non-conventional hydrogen bonds with the two ζ-methyl groups of dimethylated Lys4 of histone H3 (REF. 47). Second, some domain folds are sufficiently versatile that distinct family members recognize different types of PTM. For example, WD40-repeat domains can use the same surface to engage either a pThr-containing motif (as is the case for the WD40-repeat domain of Cdc4) or a methylated Lys site (as described for WDR5). Together, these findings underscore the widespread use of PTMs in generating sites for protein–protein interactions.

Histones are a good example of proteins that can sustain several modifications of different types, including acetylation, methylation, phosphorylation and ubiquitylation. This allows the production of many possible combinations of PTM, which can, in turn, be read by families of protein-interaction domains (FIG. 4b). These interactions can be cooperative (a diacetylated Lys motif in histone H3 binds to TAFII250), independent (the acetylation of Lys9 and Lys14 of histone H3 does not interfere with the binding of a 14-3-3 protein to phosphorylated Ser10)[49], mutually exclusive (Lys9 of histone H3 can be either acetylated or methylated), sequential (the phosphorylation of the RNA polymerase II CTD promotes the methylation of Lys36 of histone H3), or antagonistic (as noted above for the phosphorylation of Ser10 of histone H3, which blocks the binding of the HP1 chromodomain to trimethylated Lys9 during mitosis) (FIGS 2e,f,4b). Histone modifications and the resulting protein interactions provide a powerful way to rearrange chromatin function. For example, the phosphorylation of Ser139 of histone γH2AX by the ATM (ataxia telangiectasia mutated) protein kinase following DNA damage provides an optimal binding site for the tandem BRCT domains of MDC1 (mediator of DNA-damage checkpoint protein-1), which functions as a scaffold to recruit factors that are required for the repair of DNA double-strand breaks[50].

The varied modifications that have been observed for histones are also found on other chromosomal proteins and transcriptional regulators such as human p53 and CBP (cAMP-response-element-binding-protein-binding protein), and on virally-encoded polypeptides (for example, human immunodeficiency virus (HIV) Tat)[51]. These PTMs, and the molecular interactions they induce, probably provide a dynamic interplay between chromatin structural changes (or remodelling) and gene transcription or silencing.

Protein ubiquitylation

Many proteins are modified by ubiquitin or ubiquitin-like proteins such as SUMO (small ubiquitin-like modifier) and NEDD8 (neuronal-precursor-cell-expressed developmentally downregulated protein-8). Ubiquitylation is inducible (for example, by growth factors, cytokines, DNA damage and cell-cycle transitions), reversible, and occurs through the regulated function of ubiquitin-activating (E1), ubiquitin-conjugating (E2) and ubiquitin-ligase (E3) enzymes[52,53].

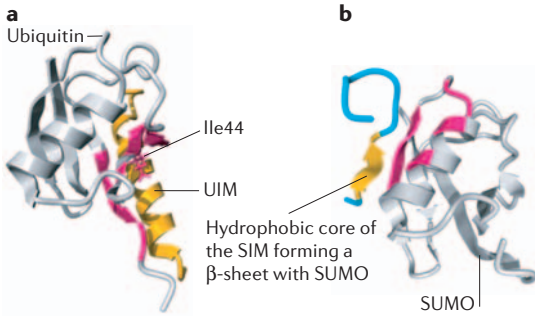

Figure 5 | **Comparison of the related folds of ubiquitin and SUMO. a** | The ubiquitin-interacting motif (UIM) of vacuolar protein sorting protein-27 (Vps27) bound to ubiquitin (Protein Data Bank (PDB) accession code 1Q0W). The UIM peptide (yellow) is shown bound to ubiquitin (grey), with the binding site for the UIM on ubiquitin (the Ile44 patch) shown in pink. The side chain of Ile44 is shown in a stick representation. **b** | The PIASx (protein inhibitor of activated STAT x) SUMO-interacting motif (SIM) bound to SUMO (small ubiquitin-like modifier)[64,74] (PDB accession code 2ASQ). The SIM peptide is shown in blue, with its hydrophobic core in yellow. SUMO is in grey, and the amino acids of its SIM-binding site are shown in pink. There are numerous ubiquitin-binding domains, including UIMs, UBMs (ubiquitin-binding motifs), CUE (coupling of ubiquitin conjugation to endoplasmic reticulum degradation) domains, ubiquitin-associated (UBA) domains, GAT (Gga and TOM1) domains, NZF (Npl4 zinc finger) motifs, UEV (ubiquitin-conjugating enzyme variant) motifs, PAZ (polyubiquitin-associated zinc finger) motifs, VHS (Vps27, HRS, STAM) domains and GLUE (GRAM-like ubiquitin-binding in Eap45) domains.

The diversity of ubiquitin interactions. Ubiquitin is a 76-residue protein that is most often linked through its C-terminal glycine to the ε-NH_2 group of Lys side chains. Variations of this simple mechanism are used to modify target proteins in diverse ways. A single ubiquitin can be attached to a single site in an acceptor protein (monoubiquitylation) (FIG. 1d) or to several Lys residues in the same protein (multiple monoubiquitylation, also known as multiubiquitylation[54]), which is akin to multi-site phosphorylation. In addition, ubiquitin contains seven Lys residues that can themselves be modified to form polyubiquitin chains, which can be built of homotypic (Lys48- or Lys63-) or heterotypic (mixed Lys29-, Lys48- and Lys63-) linkages[55].

Ubiquitin is essentially a transferable interaction domain, which, following its attachment to targets, is recognized by proteins that contain specialized interaction modules (UBDs; FIGS 1d,5a). So far, at least 11 families of UBD have been identified. These are structurally different, but generally bind to the same hydrophobic patch on ubiquitin that is centred on Ile44 (FIG. 5a), albeit

Histone deacetylase
An enzyme that removes acetyl groups from Lys residues of a histone protein. Histone acetyltransferases function in the opposite manner to add acetyl groups to Lys residues of a histone protein.

Bromodomain
An evolutionarily conserved protein domain that often binds to acetylated Lys residues in target proteins.

Epigenetic
Heritable information that is encoded by modifications of the genome and chromatin components, which affect gene expression without changing the nucleotide sequence.

TUDOR domain
A conserved chromodomain-like protein domain. Some TUDOR domains bind to methylated Lys or Arg residues.

with relatively low affinity (in the micromolar range)[56]. However, some UBDs — such as ubiquitin-binding motifs (UBMs) of translesion DNA polymerases and the Rabex-5 ubiquitin-binding zinc finger (RUZ) domain of Rabex-5, a Rab5 guanine nucleotide-exchange factor — bind to sites on ubiquitin that are distinct from the Ile44 patch[57–59]. A single ubiquitin chain might therefore recruit several targets through distinct binding surfaces for UBDs. Furthermore, many ubiquitin-binding proteins have several UBDs, and can therefore be coupled to distinct ubiquitin signals simultaneously. These features seem ideally suited to the rapid assembly and disassembly of ubiquitin-based protein-interaction networks.

These different types of ubiquitin modification and ubiquitin-dependent interaction have numerous biological effects[15]. For example, proteins tagged with Lys48-linked polyubiquitin chains are degraded following their recruitment to the proteasome by ubiquitin-binding receptors. Lys48-linked chains form a closed structure, which can be preferentially recognized by some ubiquitin-associated (UBA) domains[60]. On the other hand, monoubiquitin that is attached to an endosomal cargo is recognized by ubiquitin-binding proteins of the sorting machinery, which direct the trafficking route of the monoubiquitylated polypeptide (FIG. 3). In addition, the activation of kinases in the nuclear factor-κB signalling pathway[61], membrane fusion during organelle reassembly after mitosis and post-replicative DNA repair[62] are also governed by the specific recognition of ubiquitin-modified sites by different interacting partners. In the case of DNA repair, DNA damage promotes the monoubiquitylation of Lys164 of PCNA (proliferating cell nuclear antigen), a polymerase processivity factor that forms a sliding clamp around DNA. This monoubiquitin modification, in turn, stabilizes an interaction with Y-family translesion synthesis (TLS) polymerases, which contain UBDs (a UBM and a ubiquitin-binding zinc finger (UBZ)), and this leads to the efficient bypassing of DNA lesions, which would otherwise block replication[59].

This paradigm apparently also holds for ubiquitin-like proteins. For example, SUMO-binding proteins contain a SUMO-interacting motif (SIM), which forms a β-strand that binds in a parallel or anti-parallel orientation to the β2 strand of SUMO[63,64] (FIG. 5b). A sumoylated Lys residue therefore recruits a different set of effectors than if the same residue was ubiquitylated. For example, the sumoylation — as opposed to the ubiquitylation — of PCNA at Lys164 promotes high fidelity replication by recruiting a helicase (Srs2) that contains a SIM to replication forks, thereby preventing inappropriate recombinational repair[65,66].

In some cases, this type of switch is controlled by the phosphorylation of a nearby residue. For example, the sumoylation of Lys403 in the transcription factor MEF2A (myocyte-specific enhancer factor-2A) represses its transcriptional activity, and this is blocked by the acetylation of the same residue, which consequently stimulates MEF2A-mediated transcription[67]. Phosphorylation of MEF2A at Ser408 promotes Lys403 sumoylation, and therefore sustains the MEF2A repressed state, which, in cerebellar granule neurons, leads to increased postsynaptic differentiation. However, an activity-dependent rise in the intracellular Ca^{2+} level in these cells leads to the dephosphorylation of Ser408 by calcineurin, and a consequent switch of Lys403 from a sumoylated to an acetylated state, which results in transcriptionally active MEF2A that attenuates synapse formation. This system uses a combination of the strategies that were outlined above — the sumoylation and acetylation of Lys403 are mutually exclusive, with acetylation blocking the repressive influence of sumoylation, and the phosphorylation/dephosphorylation of Ser408 functions as a switch to allow Lys403 to toggle between its sumoylated and acetylated states. Such data emphasize the dynamic interplay between different types of PTM.

Intramolecular ubiquitin interactions. UBD-containing proteins are themselves often monoubiquitylated in a manner that requires their UBD[68], and this can result in an intramolecular interaction between attached ubiquitin and the UBD[69] (see, for example, EPS15 (epidermal-growth-factor-receptor-pathway substrate-15) in FIG. 3). The primary consequence of this intramolecular ubiquitin–UBD interaction is the imposition of a closed conformation, which inhibits the association of the UBD with ubiquitylated sites on other proteins. Monoubiquitylation therefore has a dual role. It functions as a positive signal to recruit ubiquitin-binding proteins (for example, as a sorting tag on trafficking cargoes), but it also provides a 'switch-off' signal for UBD-containing proteins (for example, endocytic adaptor proteins). This ubiquitin-mediated autoinhibition might prevent UBDs from binding to free ubiquitin or to ubiquitin that is attached to non-physiological targets, and thereby allows the accumulation of an autoinhibited pool of UBD-containing proteins that can be rapidly mobilized for signalling purposes by being targeted to specific ubiquitylated binding partners. This mode of autoregulation is similar to the intramolecular SH2-domain–pTyr interaction that occurs for Src tyrosine kinases (FIG. 2g).

Interaction affinities and specificities

PTMs have varied effects on the stability of protein–protein interactions and have correspondingly diverse biological properties. At one extreme, the association of an SH2 domain with a pTyr-containing peptide usually obtains about half of its binding energy from recognizing the phosphate group, and this association is dominated by an Arg residue that is located at the base of a specialized pTyr-binding pocket[4,70]. As a result, the binding of an SH2 domain to an unphosphorylated peptide is very weak, and the interaction is effectively phosphorylation dependent. Optimal SH2-domain–phosphopeptide interactions typically have high on and off rates, which results in dissociation constants of ~0.1–1 μM. These high rates probably contribute to the rapid docking of SH2-domain-containing proteins onto activated RTKs and the subsequent activation of signalling pathways, as well as to the prompt downregulation of these pathways[26]. Various other factors can influence the affinity and specificity of such phosphorylation-dependent complexes, such as the presence of further interaction domains

(FIG. 2b). The nature of the complexes that actually form in cells depends on a combination of the intrinsic affinities and kinetics of potential interacting partners and their relative local concentrations. This point has been emphasized by a systematic chip-based analysis of SH2 domains, which were probed with pTyr-containing peptides that were derived from the ERBB family of RTKs. This experiment indicated that the overexpression of the ERBB2 RTK — as has been observed in some human cancers — could lead to the recruitment of SH2-domain-containing proteins that would not normally bind to the activated receptor at physiological expression levels, and this recruitment potentially contributes to the oncogenic activity of the amplified receptor[71].

Bromodomains and chromodomains typically bind to acetylated and methylated motifs, respectively, with dissociation constants that are equal to or greater than 10 μM, which is a significantly weaker affinity than for SH2-domain–phosphopeptide complexes[51]. This reflects the fact that bromodomains and chromodomains are usually components of large, multidomain polypeptides that form a series of molecular interactions in the nucleus. Their binding to acetylated or methylated sites might selectively rearrange protein–protein associations in multiprotein complexes, rather than inducing *de novo* recruitment, as is the case for SH2-domain-containing proteins binding to pTyr sites. This use of multiple interaction surfaces can also be seen in the association of TLS polymerases with PCNA. This association is mediated in part by a PCNA-interacting peptide that binds to PCNA independently of any PCNA modifications and is enhanced by PCNA ubiquitylation, which is recognized by the TLS polymerase UBDs[59].

A related strategy is used in signalling from G-protein-coupled receptors (GPCRs). Following activation, GPCRs become phosphorylated on multiple Ser/Thr sites in the C-terminal tail, which directly contact β-arrestin scaffold proteins that are involved in endocytosis and signalling. However, β-arrestin also interacts with other, conformation-dependent sites on activated receptors, and might thereby preferentially sense a specific state of the receptor that has undergone both an agonist-induced conformational change and C-terminal multi-site phosphorylation[72].

At the opposite extreme to SH2 domains, the armadillo-repeat domain of β-catenin forms an extended interface with 100 C-terminal residues of the E-cadherin cytoplasmic region, and this interaction does not require phosphorylation. However, phosphorylation of multiple Ser residues in an unstructured 15-residue segment of E-cadherin results in novel contacts with β-catenin (TABLE 1) and a consequent ordering of this region of the E-cadherin tail[73]. In this case, the phosphorylation-dependent interaction is not necessary for binding *per se*, but it exerts a more subtle effect on the structure of the complex. Taken together, these observations indicate that the effects of PTM-regulated protein–protein interactions are highly context-dependent. They can be switch-like, stabilize weak interactions, function as coincidence detectors or reorganize existing complexes.

Conclusion

PTMs provide a dynamic means to modify intracellular proteins and to elicit changes in cellular function. PTMs are recognized by a wide array of modular interaction domains, which read the state of the proteome. In combination, the large set of PTMs and corresponding interaction domains — only a few of which could be mentioned here — provide a versatile mechanism to orchestrate cellular behaviour. Here, we identified some basic attributes of combinatorial PTM-dependent interactions that govern the assembly of regulatory pathways in cells.

1. Yang, X. J. Multisite protein modification and intramolecular signaling. *Oncogene* **24**, 1653–1662 (2005).
An excellent review that describes various ways in which multi-site post-translational modifications can be used to coordinate protein function in a cell.
2. Cullen, P. J., Cozier, G. E., Banting, G. & Mellor, H. Modular phosphoinositide-binding domains — their role in signalling and membrane trafficking. *Curr. Biol.* **11**, R882–R893 (2001).
3. Venter, J. C. The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
4. Kuriyan, J. & Cowburn, D. Modular peptide recognition domains in eukaryotic signaling. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 259–288 (1997).
5. Gimona, M. Protein linguistics — a grammar for modular protein assembly? *Nature Rev. Mol. Cell Biol.* **7**, 68–73 (2006).
6. Waksman, G., Shoelson, S., Pant, N., Cowburn, D. & Kuriyan, J. Binding of a high affinity phosphotyrosyl peptide in the Src SH2 domain: crystal structures of the complexed and peptide-free forms. *Cell* **72**, 779–790 (1993).
A seminal structural study that describes the modular nature of the Src SH2 domain and its interaction with a pTyr-containing peptide.
7. Owen, D. J. *et al.* The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase Gcn5p. *EMBO J.* **19**, 6141–6149 (2000).
8. Nielsen, P. R. *et al.* Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* **416**, 103–107 (2002).
9. Durocher, D. *et al.* The molecular basis of FHA domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms. *Mol. Cell* **6**, 1169–1182 (2000).
10. Hatada, M. H. *et al.* Molecular basis for interaction of the protein tyrosine kinase ZAP-70 with the T-cell receptor. *Nature* **377**, 32–38 (1995).
11. Hu, J., Liu, J., Ghirlando, R., Saltiel, A. R. & Hubbard, S. R. Structural basis for recruitment of the adaptor protein APS to the activated insulin receptor. *Mol. Cell* **12**, 1379–1389 (2003).
Describes a crystal structure that provides an example of how multiple molecular interactions, including the homodimerization of the interaction domain, can cooperatively enhance recruitment to a substrate.
12. Nash, P. *et al.* Multisite phosphorylation of a CDK inhibitor sets a threshold for the onset of DNA replication. *Nature* **414**, 516–523 (2001).
13. Orlicky, S., Tang, X., Willems, A., Tyers, M. & Sicheri, F. Structural basis for phosphodependent substrate selection and orientation by the SCF^{Cdc4} ubiquitin ligase. *Cell* **112**, 243–256 (2003).
References 12 and 13 provide insights into how a threshold number of multiple weak phosphorylation sites on Sic1 cooperate to result in a high-affinity interaction with Cdc4.
14. Clapperton, J. A. *et al.* Structure and mechanism of BRCA1 BRCT domain recognition of phosphorylated BACH1 with implications for cancer. *Nature Struct. Mol. Biol.* **11**, 512–518 (2004).
15. Haglund, K. & Dikic, I. Ubiquitylation and cell signaling. *EMBO J.* **24**, 3353–3359 (2005).
16. Ye, X. *et al.* Recognition of phosphodegron motifs in human cyclin E by the SCF^{Fbw7} ubiquitin ligase. *J. Biol. Chem.* **279**, 50110–50119 (2004).
17. Strohmaier, H. *et al.* Human F-box protein hCdc4 targets cyclin E for proteolysis and is mutated in a breast cancer cell line. *Nature* **413**, 268–279 (2001).
18. Welcker, M. *et al.* Multisite phosphorylaton by Cdk2 and GSK3 controls cyclin E degradation. *Mol. Cell* **12**, 381–392 (2003).
19. Rajagopalan, H. *et al.* Inactivation of hCDC4 can cause chromosomal instability. *Nature* **428**, 77–81 (2004).
20. Pawson, T. & Nash, P. Assembly of cell regulatory systems through protein interaction domains. *Science* **300**, 445–452 (2003).
21. Jordan, M. S., Singer, A. L. & Koretzky, G. A. Adaptors as central mediators of signal transduction in immune cells. *Nature Immunol.* **4**, 110–116 (2003).
22. Fischle, W. *et al.* Regulation of HP1–chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438**, 1090–1091 (2005).
23. Hirota, T., Lipp, J. J., Toh, B. H. & Peters, J. M. Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438**, 1176–1180 (2005).
References 22 and 23 provide an example of how PTMs can be used to antagonize interactions with regulatory proteins in order to control gene expression.
24. Sicheri, F., Moarefi, I. & Kuriyan, J. Crystal structure of the Src family tyrosine kinase Hck. *Nature* **385**, 602–609 (1997).

This paper reveals how the SH2 and SH3 domains of the Src family tyrosine kinase Hck intramolecularly interact with modified peptide motifs in Hck to regulate the kinase activity of the protein.
25. Yaffe, M. B. & Elia, A. E. Phosphoserine/threonine-binding domains. *Curr. Opin. Cell Biol.* **13**, 131–138 (2001).
26. Bradshaw, J. M. & Waksman, G. Molecular recognition by SH2 domains. *Adv. Protein Chem.* **61**, 161–210 (2002).
27. Heldin, C.-H., Ostman, A. & Ronnstrand, L. Signal transduction via platelet-derived growth factor receptors. *Biochem. Biophys. Acta* **1378**, F79–F113 (1998).
28. Hunter, T. Signaling — 2000 and beyond. *Cell* **100**, 113–127 (2000).
29. Schulze, W. X., Deng, L. & Mann, M. Phosphotyrosine interactome of the ErbB-receptor kinase family. *Mol. Syst. Biol.* **1038**, E1–E13 (2005).
30. Kavanaugh, W. M., Turck, C. W. & Williams, L. T. PTB domain binding to signaling proteins through a sequence motif containing phosphotyrosine. *Science* **268**, 1177–1179 (1995).
31. Benes, C. H. *et al.* The C2 domain of PKCδ is a phosphotyrosine binding domain. *Cell* **121**, 158–160 (2005).
32. Sun, X. J. *et al.* Structure of the insulin receptor substrate IRS-1 defines a unique signal transduction protein. *Nature* **352**, 73–77 (1991).
33. Uhlik, M. T. *et al.* Structural and evolutionary division of phosphotyrosine binding (PTB) domains. *J. Mol. Biol.* **345**, 1–20 (2005).
34. Yaffe, M. B. *et al.* The structural basis for 14-3-3: phosphopeptide binding specificity. *Cell* **91**, 961–971 (1997).
35. MacKintosh, C. Dynamic interactions between 14-3-3 proteins and phosphoproteins regulate diverse cellular processes. *Biochem. J.* **381**, 329–342 (2004).
36. Wu, J. W. *et al.* Crystal structure of a phosphorylated Smad2. Recognition of phosphoserine by the MH2 domain and insights on Smad function in TGF-signaling. *Mol. Cell* **8**, 1277–1289 (2001).
37. Meinhart, A., Kamenski, T., Hoeppner, S., Baumli, S. & Cramer, P. A structural perspective of CTD function. *Genes Dev.* **19**, 1401–1415 (2005).
38. Fabrega, C., Shen, V., Shuman, S. & Lima, C. D. Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol. Cell* **11**, 1549–1561 (2003).
39. Li, M. *et al.* Solution structure of the Set2–Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proc. Natl Acad. Sci. USA* **102**, 17636–17641 (2005).
40. Keogh, M. C. *et al.* Cotranscriptional Set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123**, 593–605 (2005).
41. Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **203**, 1074–1080 (2001).
42. Jacobs, S. A. & Khorasanizadeh, S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **295**, 2080–2083 (2002).
43. Jacobson, R. H., Ladurner, A. G., King, D. S. & Tjian, R. Structure and function of a human TAFII250 double bromodomain module. *Science* **288**, 1422–1425 (2000).
44. Flanagan, J. F. *et al.* Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* **438**, 1090–1091 (2005).

References 43 and 44 illustrate the use of tandem interaction domains to recognize two adjacent modified peptide motifs.
45. Kim, J. *et al.* Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep.* **7**, 397–403 (2006).
46. Wysocka, J. *et al.* WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121**, 859–872 (2005).
47. Han, Z. *et al.* Structural basis for the specific recognition of methylated histone H3 lysine 4 by the WD-40 protein WDR5. *Mol. Cell* **22**, 137–144 (2006).
48. Huang, Y., Fang, J., Bedford, M. T., Zhang, Y. & Xu, R. M. Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science* **312**, 748–751 (2006).
49. Macdonald, N. *et al.* Molecular basis for the recognition of phosphorylated and phosphoacetylated histone H3 by 14-3-3. *Mol. Cell* **20**, 199–211 (2005).
50. Stucki, M. *et al.* MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. *Cell* **123**, 1213–1226 (2005).
51. Yan, K. S. & Zhou, M.-M. in *Modular Protein Domains* Ch. 11 (eds Cesarini, G., Gimona, M., Sudol, M. & Yaffe, M.) 227–236 (Wiley-VCH, Weinheim, 2005).
52. Pickart, C. M. Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* **70**, 503–533 (2001).
53. Pickart, C. M. & Eddins, M. J. Ubiquitin: structures, functions, mechanisms. *Biochim. Biophys. Acta* **1695**, 55–72 (2004).
54. Haglund, K., Di Fiore, P. P. & Dikic, I. Distinct monoubiquitin signals in receptor endocytosis. *Trends Biochem.* **28**, 598–603 (2003).
55. Pickart, C. M. & Fushman, D. Polyubiquitin chains: polymeric protein signals. *Curr. Opin. Chem. Biol.* **8**, 610–616 (2004).
56. Hicke, L., Schubert, H. L. & Hill, C. P. Ubiquitin-binding domains. *Nature Rev. Mol. Cell Biol.* **6**, 610–621 (2005).
57. Penengo, L. *et al.* Crystal structure of the ubiquitin binding domains of Rabex-5 reveals two modes of interaction with ubiquitin. *Cell* **124,** 1183–1195 (2006).
58. Lee, S. *et al.* Structural basis for ubiquitin recognition and autoubiquitination by Rabex-5. *Nature Struct. Mol. Biol.* **13**, 186–188 (2006).
59. Bienko, M. *et al.* Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. *Science* **310**, 1821–1824 (2005).
This paper reveals novel ubiquitin interactions and their role in the repair of DNA lesions.
60. Varadan, R., Assfalg, M., Raasi, S., Pickart, C. & Fushman, D. Structural determinants for selective recognition of a Lys48-linked polyubiquitin chain by a UBA domain. *Mol. Cell* **18**, 687–698 (2005).
61. Krappmann, D. & Scheidereit, C. A pervasive role of ubiquitin conjugation in activation and termination of IκB kinase pathways. *EMBO Rep.* **6**, 321–326 (2005).
62. Stelter, P. & Ulrich, H. D. Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. *Nature* **425**, 188–191 (2003).
63. Song, J., Durrin, L. K., Wilkinson, T. A., Krontiris, T. G. & Chen, Y. Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. *Proc. Natl Acad. Sci. USA* **101**, 14373–14378 (2004).
64. Hecker, C. M., Rabiller, M., Haglund, K., Bayer, P. & Dikic, I. Specification of SUMO1- and SUMO2-interacting motifs. *J. Biol. Chem.* **281**, 16117–16127 (2006).
65. Pfander, B., Moldovan, G. L., Sacher, M., Hoege, C. & Jentsch, S. SUMO-modified PCNA recruits Srs2 to prevent recombination during S phase. *Nature* **436**, 428–433 (2005).
66. Papouli, E. *et al.* Crosstalk between SUMO and ubiquitin on PCNA is mediated by recruitment of the helicase Srs2p. *Mol. Cell* **19**, 123–133 (2005).
67. Shalizi, A. *et al.* A calcium-regulated MEF2 sumoylation switch controls postsynaptic differentiation. *Science* **311**, 1012–1017 (2006).
68. Polo, S. *et al.* A single motif responsible for ubiquitin recognition and monoubiquitination in endocytic proteins. *Nature* **416**, 451–455 (2002).
69. Hoeller, D. *et al.* Regulation of ubiquitin-binding proteins by monoubiquitination. *Nature Cell Biol.* **8**, 163–169 (2006).
70. Piccione, E. *et al.* Phosphatidylinositol 3-kinase p85 SH2 domain specificity defined by direct phosphopeptide/SH2 domain binding. *Biochem.* **32**, 3197–3202 (1993).
71. Jones, R. B., Gordus, A., Krall, J. A. & Macbeath, G. A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature* **439**, 168–174 (2006).
72. Hanson, S. M. *et al.* Differential interaction of spin-labeled arrestin with inactive and active phosphorhodopsin. *Proc. Natl Acad. Sci. USA* **103**, 4900–4905 (2006).
73. Huber, A. H. & Weis, W. L. The structure of the β-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by β-catenin. *Cell* **105**, 391–402 (2001).
74. Song, J., Zhang, Z., Hu, W. & Chen, Y. Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: a reversal of the bound orientation. *J. Biol. Chem.* **280**, 40122–40129 (2005).

Acknowledgements
We are indebted to C. Lim and M. Rabiller for assistance with the figures, and to R. Linding and M. Tyers for insightful comments. B.T.S. is funded by a fellowship from the Cancer Research Institute (New York, USA). Work in the authors' laboratories is funded by the Canadian Institutes for Health Research (CIHR) and the National Cancer Institute of Canada and Genome Canada (T.P.), the Deutsche Forschungsgemeinschaft and the German–Israeli Foundation (I.D.), and the United States National Institutes of Health (M.-M.Z). T.P. is a distinguished investigator of the CIHR.

Competing interests statement
The authors declare no competing financial interests.

DATABASES
The following terms in this article are linked online to:
Protein Data Bank: http://www.rcsb.org/pdb
1E6I | 1JYR | 1LM8 | 1Q0W | 1Q3L | 2ASQ
UniProtKB: http://ca.expasy.org/sprot
Aurora-B | Cdc4 | FBW7 | MDC1 | NEDD8 | Rabex-5 | Sic1 | TAFII250 | ZAP-70

FURTHER INFORMATION
The Pawson Lab: http://pawsonlab.mshri.on.ca

POSTER
Reading protein and phospholipid modifications with interaction domains:
http://www.nature.com/nrm/poster/modifications
Access to this links box is available online.