

## **The switches.ELM resource: a compendium of conditional regulatory interaction interfaces.**

**Authors:** Kim Van Roey<sup>1</sup>, Holger Dinkel<sup>1,†</sup>, Robert J. Weatheritt<sup>1,†</sup>, Toby J. Gibson<sup>1</sup> and Norman E. Davey<sup>1,\*</sup>

<sup>1</sup>Structural and Computational Biology Unit, European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69117 Heidelberg, Germany.

\*To whom correspondence should be addressed. E-mail: [davey@embl.de](mailto:davey@embl.de).

†These authors contributed equally to this work.

**One Sentence Summary:** A resource centered on short linear motifs provides a repository and exploratory tool for conditional protein interactions.

**Abstract:** Short linear motifs (SLiMs) are protein interaction sites that play an important role in cell regulation by controlling protein activity, localization, and local abundance. The functionality of a SLiM can be modulated in a context-dependent manner to induce a gain, loss, or exchange of binding partners, which will affect the function of the SLiM-containing protein. As such, these conditional interactions underlie molecular decision-making in cell signaling. We identified multiple types of pre- and post-translational switch mechanisms that can regulate the function of a SLiM and thereby control its interactions. The collected examples of experimentally characterized SLiM-based switch mechanisms were curated in the freely accessible switches.ELM resource (<http://switches.elm.eu.org>). On the basis of these examples, we defined and integrated rules to analyze SLiMs for putative regulatory switch mechanisms. We applied these rules to known validated SLiMs, providing evidence that over half of these are likely to be pre- or post-translationally regulated. In addition, we showed that post-translationally modified sites are enriched around SLiMs, which enables cooperative and integrative regulation of protein interaction interfaces. We foresee switches.ELM complementing available resources to extend our knowledge of the molecular mechanisms underlying cell signaling.

## **Introduction**

Tight regulation of protein function is vital for reliable and robust control of cell physiology and is enforced at every stage of a protein's life cycle (1). This includes control of local protein abundance by modulating gene transcription, mRNA decay, translation efficiency, protein degradation, subcellular localization, and scaffolding (1, 2). Furthermore, pre-translational mechanisms, such as alternative promoter usage, alternative splicing, and RNA editing, can produce functionally distinct protein species derived from the same gene (3-5). Most commonly however, protein function is controlled at the post-translational level through interactions mainly with other proteins but also with other cellular components including membranes, small molecules, and nucleic acids (6-9). Regulatory interactions are often context-dependent and the function of a protein can be altered by rewiring its interaction network in response to a specific signal. As such, these conditional interactions control whether and how subsequent signaling proceeds to elicit appropriate responses to internal and external cues, and thereby mediate molecular decision-making in cell regulation (10-16).

Regulatory protein interactions are often mediated by motifs [also referred to as short linear motifs (SLiMs)], a subset of compact and degenerate binding modules typically located in intrinsically disordered regions (IDRs) of the proteome that lack a stable tertiary structure under native conditions (13, 17-20). SLiMs are a spatially efficient and convergently evolvable solution for encoding interaction interfaces (21, 22). They are ubiquitous in eukaryotic proteomes and mediate many regulatory functions, including directing ligand binding, providing docking sites for modifying enzymes, controlling protein stability, and acting as signals to target proteins to specific subcellular locations. Typically, only 3 to 4 residues of a SLiM determine the majority of the binding specificity and affinity, hence motif-mediated interactions occur with low affinity, are transient and reversible, and can be easily modulated (15, 20, 22). These intrinsic properties make SLiMs ideal regulatory modules and enable them to conditionally switch between their 'on' and 'off' states or between multiple, functionally distinct 'on' states. Switching of a motif between different functional states can be mediated by multiple mechanisms that often depend on post-translational modifications (PTMs) and the cooperative or competitive use of multiple overlapping or adjacent SLiMs (7, 15, 22, 23). Conditional switching of motif functionality can stimulate a gain, loss, or exchange of binding partners, and thereby regulate the function of SLiM-containing proteins in a context-dependent manner.

Pre-translational mechanisms can irreversibly control the presence and binding properties of a SLiM, respectively, by producing protein isoforms that lack a specific motif or have a motif with altered flanking regions (5, 24-28). Post-translationally, several mechanisms, most commonly protein modification, can reversibly regulate motif function by switching a single motif on or off, or by gradually altering its binding strength (29-34). When multiple SLiMs are present, they can be used cooperatively to mediate multivalent binding, having a more than additive effect on protein interaction strength (7, 35-38). Alternatively, multiple overlapping or adjacent SLiMs can bind competitively to mutually exclusive interaction partners. A signal can shift the balance of the competition in favor of a specific binding partner (39, 40), or mediate a sequential exchange of binding partners by altering the specificity of the SLiM-containing region in a stepwise manner (41, 42). These different modes of motif regulation are widely used in the cell and mediate conditional, combinatorial, and fine-tuned control of protein function (15). This includes controlling the subcellular localization of proteins or protein isoforms, temporally and spatially regulating their stability, and directing the assembly of functionally distinct signaling complexes depending on prevailing conditions (1, 15, 37, 38).

Despite being prevalent in and imperative for dynamic and robust cell regulation, the mechanisms that control SLiM-mediated interactions through context-dependent switching of motif function, and thereby control the function of SLiM-containing proteins and the processes they regulate, are not fully captured in the currently available resources. Here, we present the switches.ELM resource (<http://switches.elm.eu.org>), which consists of a categorized repository of 710 manually curated, experimentally validated instances of SLiM-based switch mechanisms that we collected from the literature. On the basis of these instances, we defined a set of rules to suggest switch mechanisms for SLiMs. These rules were integrated in the resource to enable users to explore a motif of interest for putative

regulatory mechanisms and to direct them to the relevant literature and data sources. We then applied these rules to analyze the 2,070 experimentally validated SLiMs that are manually curated in the Eukaryotic Linear Motif (ELM) database (43). This survey provides suggestive evidence that over half of the characterized SLiMs are likely to be pre- or post-translationally regulated. In addition, we elaborated on a previous study (23) and found that modification sites are enriched in regions flanking the validated SLiMs from ELM, and that a large portion of modification sites regulating annotated switch instances

are modified by multiple enzymes. Finally, we found that SLiMs involved in the curated switches are more likely to be located in non-constitutive exons compared to SLiMs in general, and are thus more likely to be removed from specific protein isoforms.

## **Results**

### *Collection and categorization of switches.ELM data*

In the context of switches.ELM, a ‘switch mechanism’ is a regulatory process that controls SLiM-mediated interactions by modulating the function of a SLiM, and hence affects the function of the SLiM-containing protein due to the concomitant gain, loss, or exchange of binding partners. The modulation of motif function can be mediated by a pre-translational mechanism, the addition of a PTM, or the binding of a molecule. The experimentally validated instances of SLiM-based switch mechanisms curated in the switches.ELM resource were collected by a literature survey (see Materials and Methods). On the basis of the mechanism that mediates switching of motif functionality, these instances were classified into different switch types and subtypes, listed with examples in Table 1 and explained in

detail here. It should be taken into consideration that the curated switch instances were regarded in isolation; however, in a biological context, these definitions partially overlap and the different mechanisms combine to mediate higher-order regulatory functions that allow integration of multiple signals. Also, some instances have a unique regulatory mechanism and are currently labeled as uncategorized switches. As more instances accumulate, these switches may be worthy of a novel switch type or subtype.

Simple ‘binary’ switches shift between two distinct functional states of a motif. This can be either the ‘on’ and ‘off’ states, or two ‘on’ states with different affinities for a single binding partner. On the basis of the mechanism that regulates switching between the two states, different subtypes were defined. A first subtype comprises the ‘altered physicochemical compatibility’ switches that are regulated by addition of a PTM to a residue in or adjacent to a SLiM. Due to their specific structural and physicochemical properties, PTMs can alter the intrinsic specificity or affinity of a SLiM for its binding partner (29-31). For example, the addition of a PTM to a specific residue within a SLiM is often a prerequisite for a SLiM-mediated interaction (44), whereas a PTM adjacent to a SLiM can enhance the strength of an interaction (45). Alternatively, a PTM can partially or completely inhibit an interaction by rendering a SLiM physicochemically incompatible with its binding partner (46). The ‘allostery’ switches are a second subtype comprising SLiMs or SLiM-binding globular protein domains whose binding properties are modulated indirectly by allosteric effects induced by the addition of a PTM or by the binding of an effector molecule to a site that is distinct from the actual interaction interface (35, 47-49). ‘Pre-translational’ switches are the third and final subtype of binary switches. These are mediated by pre-translational mechanisms that irreversibly alter motif presence or function, either by isoform-specific removal of a SLiM-containing exon or by generating isoforms whose SLiMs show different binding specificity or affinity due to variations in

their flanking regions (5, 25-27, 50-53).

Other, more complex mechanisms exist to regulate the function of a single motif. The ‘cumulative’ switches modulate motif functionality by multisite modification, meaning the addition of PTMs to multiple residues in or adjacent to the motif (32). A subtype of cumulative switches is the ‘rheostatic’ switches, which gradually alter the affinity of a motif for a single binding partner by the addition of multiple PTMs that additively strengthen or weaken the interaction. This enables fine-tuned control of motif activity depending on the strength of a signal or integration of multiple signals (33, 34, 54). Another more complex switch type acting on single motifs consists of ‘pre-assembly’ switches, which require prior formation of a complex as a prerequisite for a SLiM-mediated interaction to occur. A subtype of the pre-assembly switch is the prerequisite formation of a ‘composite binding site’ that spans more than one component of the complex. Because neither component on its own contains a functional motif-binding domain, binding of the motif only occurs in the context of the active, fully assembled complex (55).

The small footprint of SLiMs facilitates the occurrence of regions with high functional density containing multiple motifs that can bind cooperatively or competitively. ‘Avidity-sensing’ switches involve multiple motifs that mediate high-avidity binding, effected by low-affinity single interactions that have a more than additive effect on binding strength (7, 35, 37). These high-avidity binding events can be mediated by multiple motifs within a single protein, even if the motifs are separated by large distances. The motifs can also be distributed over more than one protein, thereby promoting multi-protein complex formation. High-avidity interactions are indispensable for cell signaling, as they mediate the assembly of large meta- stable signaling complexes at a biologically relevant time-scale and contribute to the switch-like activation of the emergent properties of these complexes (36, 38). Alternatively, adjacent or overlapping motifs can bind in a mutually exclusive

manner and thus promote competitive binding (22). The outcome of the competition depends on the intrinsic specificity and affinity of the SLiMs for, as well as the local abundance of, the competitors; modulating either of these can shift the balance in favor of a specific interactor. Mutually exclusive interaction interfaces are characterized by distinct ‘on’ states, and ‘specificity’ switches can mediate a regulated exchange of the distinct binding partners through different mechanisms, characterized by their subtypes. ‘Competition’ switches between mutually exclusive SLiMs depend on local target protein abundance, which can be regulated by changing the expression level or subcellular localization of the competitors, as well as by scaffolding (1, 15, 40, 56, 57). ‘Altered binding specificity’ switches tip the balance of the competition for mutually exclusive SLiMs in favor of one of the interactors by PTM-dependent modulation of the intrinsic specificity or affinity of the binding region (39). In addition, multiple, successive PTMs allow sequential switching of different binding partners in an ordered manner by stepwise alteration of binding specificity (41). A common and biologically important subset of competition switches involves mutually exclusive interactions where there is strong preferential binding of one of the competing interactors, either due to a large difference in the intrinsic specificity or affinity of the binding region for a specific competitor, or due to a large difference in the local abundance of the competitors. The interaction that is favored thus inhibits binding of a mutually exclusive interactor. ‘Motif hiding’ switches can preclude a motif from binding by sterically blocking its accessibility through binding of a protein to an overlapping or adjacent motif (58, 59). Similarly, ‘domain hiding’ switches can sterically block binding of a SLiM to a domain by binding of another molecule to an overlapping site on the domain (60, 61). With respect to the functional outcome, these switches act in a binary fashion, and activation of the masked binding region depends on active removal of the preferred interactor.

*Curation and visualization of switches.ELM data*

The first release (version 1.0, March 2013) of the switches.ELM resource contained 710 manually curated switch instances, involving 817 distinct interactions mediated by 664 different SLiMs in 409 proteins. For each of these instances, different types of information were annotated to accurately describe the switch mechanism (see Materials and Methods). Each switch instance contains one or more SLiM-mediated interactions, with each interaction comprising two interactors. A specific switch mechanism modulates the function of a SLiM in one of the interactors and thereby regulates its interactions. The switch mechanism depends on a pre-translational mechanism, the addition of a PTM, or binding of a molecule. It can be reversible or irreversible and can either positively or negatively affect a specific interaction. A positive effect can be either induction or enhancement of an interaction, whereas a negative effect can be partial inhibition or complete abrogation of an interaction. Also, interdependencies between different interactions were annotated. An interaction can be mutually exclusive with one or more other interactions or, in contrast, require one or more previous interactions to occur. Interactions that occur in a specific, defined order were numbered according to their position in the sequence of binding events. The context-dependency of the switches is captured by adding temporal, spatial and contextual conditionality to the interaction data. These data were either manually curated (table S1) or computationally inferred (table S2) (see Materials and Methods). This includes, when applicable, manual curation of the cell cycle phases during which a switch is active. Also, subcellular localization data were manually annotated, retrieved from manually curated localization data in ELM (43), or inferred from the Gene Ontology (GO) resource (62). Furthermore, pathways in which a switch is involved were manually annotated or inferred from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (63). For interactions regulated by PTMs, modifying enzymes were manually curated or inferred from the phosphoSitePlus (64) and phospho.ELM (65) resources. All curated and inferred data can be downloaded from the website in tab-separated values (tsv)

format. Data from version 1.0 are supplied in the supplementary tables S1 and S2.

The curated switches can be browsed by different categories, which are color-coded, or searched on the basis of all annotated information by user-defined search terms. Each switch instance can be viewed in a visual output that gives a clear and detailed overview of the regulatory mechanisms (Fig. S1), including a concise text description, the type of switch, the participating molecules, the binding interfaces that mediate the interactions, and information on how interactions are regulated. The curated and inferred cell cycle phase, subcellular localization, pathway and modifying enzyme data are provided as contextual information. For a sequence-centric view, an alignment of orthologous proteins corresponding to the protein region containing the SLiM of interest is shown (Fig. S1).

Biologically relevant data, including genomic data (66), interaction interfaces (43, 67), modification sites (64, 65), mutation data (66), and structural information (68, 69), are mapped to the region and displayed. A diagram of the complete protein architecture is also provided to situate a switch in the context of the whole protein. Data displayed in this overview includes other switch instances curated for the protein, SLiMs, protein domains, PTM sites and secondary structure. Detailed descriptions for all data used in the switches output view are shown upon hovering over a feature, whereas clicking on the feature navigates to the corresponding external data source (43, 62-68, 70-73).

#### *The switches.ELM exploratory analysis tool*

To enable users to submit and analyze a motif of interest for putative switch mechanisms regulating its function, a rules-based analysis tool was developed and integrated in switches.ELM (Fig. S2). The tool is designed to be exploratory and is thus deliberately over-predictive, as it puts forward all hypotheses concerning regulation of the entered motif

that can be inferred given the available data. Further reading of the relevant literature, which is presented to the user, is necessary to gain an understanding of the true regulatory mechanisms. The switch mechanisms for a motif suggested by the tool are inferred from rules that were defined for the different switch types and subtypes on the basis of SLiM-specific modes of regulation and on the curated switch instances (see Materials and Methods).

Some SLiM classes annotated in ELM are not intrinsically active and require switching to be functional. For motifs that can be mapped to the motif patterns defined for the different motif classes in ELM (43), rules specific to these motif classes are applied. For example, some motif classes require a PTM for binding, whereas others are known to be allosterically regulated or to depend on formation of a composite binding site. Also, some SLiM classes are known to mediate multivalent binding when at least two motifs of the same class are present in a protein. Finally, some motifs have been observed to mediate intra-molecular interactions, which enable regulation by motif and domain hiding if a protein contains both the motif and the corresponding motif-binding domain.

Additional general rules were defined on the basis of the collected switch instances. The presence of one or more experimentally determined PTM sites, retrieved from phosphoSitePlus (64) and phospho.ELM (65), within 20 residues up- or downstream of a motif is suggestive of an altered physicochemical compatibility or rheostatic switch mechanism, respectively. For motifs contained within a non-constitutive exon that is specifically removed from one or more protein isoforms [on the basis of isoform data curated in UniProtKB (66)], pre-translational switching is suggested. The presence of overlapping or adjacent (within 20 residues) motifs might confer regulation by competition or motif hiding. Altered binding specificity is proposed as a switch mechanism if the conditions for both competition and altered physicochemical compatibility are met.

### *Analysis of ELM instances for putative switch mechanisms*

To suggest novel regulatory mechanisms for previously characterized SLiMs, we applied the set of rules describing the switch mechanisms to the 2,070 experimentally validated SLiMs curated in the ELM database as of November 2012 (43) (see Materials and Methods). A total of 1,789 switch mechanisms were proposed for the 2,070 ELM instances, and hence, on average a motif annotated in ELM was suggested to be regulated by 0.86 switch mechanisms (Fig. 1A). Confidence in the suggested switch mechanisms varies. Several of these are inherently correct, such as pre-translational switching of functional motifs in non-constitutive exons, whereas others, such as rheostatic switches, would require extensive biochemical characterization to be validated. In all, 1,058 (51.1%) of the 2,070 ELM instances had proteomic or genomic evidence suggestive of a pre- or post-translational regulatory switch mechanism (Fig. 1B). 375 of these (18.1%) had experimental evidence for a switch mechanism available and were curated in switches.ELM. 681 (32.9%) ELM instances were modified by PTMs on residues within or adjacent (within 5 residues of a defined residue) to the motif (Fig. 1A), whereas 214 (10.3%) were modified on defined residues (excluding modification sites used to induce binding to protein domains that recognize a specific PTM on a specific motif residue). These defined residues are typically in direct contact with the binding partner and contribute most of the binding specificity and affinity. Hence, the addition of a PTM to such a residue will likely have a strong effect on the binding properties of the motif. In addition, 218 instances (10.5%) were modified by PTMs on multiple adjacent positions (within 5 residues of a defined residue), potentially enabling integration of upstream signals acting rheostatically, through priming, or through independent modification sites (Fig. 1A). 115 (5.6%) of the SLiMs annotated in ELM occurred within 20 residues of each other, suggesting that they may mediate mutually exclusive binding events (Fig. 1A). 304 (14.6%)

of the ELM instances occurred in non-constitutive exons, corroborating a recent study showing preferential splicing of SLiM-containing exons (24) (Fig. 1A). All switch mechanisms suggested by the switches.ELM tool to regulate motifs curated in ELM are listed in the supplementary table S3 and are available as an interactive results page on the switches.ELM website.

Analysis of ELM instances suggested many switch mechanisms that might regulate known functional motifs. Although direct experimental confirmation could be found in the literature for some of these, evidence was either circumstantial or absent for many others. Some representative examples of the latter are discussed here to demonstrate how hypotheses developed from suggestions made by switches.ELM can direct further experimental elucidation of these systems and how the tool can be useful to guide and focus experimental research aimed at interpreting high-throughput modification data.

Example 1. The transcription intermediary factor 1- $\beta$  (TIF1- $\beta$ ) functions as a co-repressor of heterochromatin protein 1 (HP1) homologs to mediate transcriptional repression of genes, including cell cycle enhancers like cyclin-A2 (*CCNA2*) (74). The switches.ELM tool suggested several switch mechanisms for the HP1-binding motif of TIF1- $\beta$ , including an altered physicochemical compatibility switch mediated by phosphorylation of the Ser<sup>489</sup> residue. This suggestion resulted from integration of data from a large-scale study of cell cycle-regulated protein phosphorylation, which showed increased phosphorylation of TIF1- $\beta$  in cells arrested in M phase compared to cells arrested in G1 phase of the cell cycle (75). Two sites with increased phosphorylation were Ser<sup>473</sup> and Ser<sup>489</sup>, which are adjacent to and part of, respectively, the HP1-binding motif. Phosphorylation of Ser<sup>473</sup> has previously been shown to inhibit binding of TIF1- $\beta$  to HP1 and to be associated with induction of *CCNA2* expression at S phase (74). Thus, direct phosphorylation of the motif on Ser<sup>489</sup> might serve the same purpose, because it occurs under similar conditions and probably has the

same, if not stronger, effect on binding to HP1. This example shows how the switches.ELM tool offers testable hypotheses by integrating motif and PTM data from low- and high-throughput experiments.

Example 2. The erythropoietin (EPO) receptor (EPOR-F) stimulates proliferation and terminal differentiation of committed erythroid progenitor cells in response to its ligand. EPO-induced dimerization of the receptor initiates Janus kinase (JAK) and signal transducer and activator of transcription (STAT) signaling. Recruitment of a JAK dimer to the EPOR-F results in tyrosine-specific phosphorylation of the receptor, which creates docking sites for STAT proteins. The latter become phosphorylated by JAK, translocate to the nucleus and initiate transcription of target genes, including anti-apoptotic factors such as B-cell lymphoma 2 (*BCL2*) (76, 77). Early-stage progenitor cells predominantly express the truncated EPOR-T isoform, which acts as a dominant-negative receptor for EPO-mediated signaling. In addition, cells transfected with EPOR-T are more susceptible to apoptosis (78). The switches.ELM tool suggested that these observations might be explained by a pre-translational switch that removes two SH2-binding motifs, which are involved in phospho-dependent recruitment of STAT proteins, from the EPOR-T isoform.

#### *Analysis of ELM instances for enrichment of motif-proximal modification sites*

A large subset of the altered physicochemical compatibility switches curated in switches.ELM comprised of SLiMs that are regulated by the addition of a specific PTM to a specific residue in the motif, which induces binding to PTM-binding domains such as the SH2, PTB, and Bromo domains (6, 29, 31). However, many motifs annotated in switches.ELM are regulated by the addition of PTMs to residues that are distinct from these position-specific PTM sites that induce binding to PTM-binding domains. Of the 710 curated switches, 80 are regulated by this mechanism, and the vast majority of their

switching modifications occur within 5 residues of the motif [81 (76.3%) of the 114 modification sites involved in this subset of switches]. To assess if this is a general feature of SLiMs, the available modification data (64, 65) were mapped to the ELM instances (43) to determine whether there was an enrichment of modification sites around validated functional SLiMs. The residues adjacent (within 5 residues of a defined residue) to validated motif instances had an enrichment of modification sites [690 (3.6%) of the 19,112 residues are modified] when compared to a background dataset, in which 2,151 (2.9%) of the 73,005 residues are modified (see Materials and Methods). The proportion of PTMs present in each set was calculated and a statistically significant enrichment of modification sites adjacent to known motifs was found (Fig. 1C). Similar results were obtained when modification data were restricted to only high-throughput experimentation, suggesting that the observation is not an acquisition bias due to annotation of the modifications in the motif literature. Also, when ELM motif instances belonging to a MOD class, which are recognized and modified by enzymes, were omitted from the analysis, the enrichment of modification sites adjacent to motif instances was still significant. These findings are in agreement with results from a previous study analyzing coupling between motifs and phosphorylation events (23).

#### *Analysis of switches.ELM instances*

Previous studies indicate that motifs are enriched in IDRs (22). We found that 479 (72.1%) of the 664 SLiMs that participate in curated switches occurred in predicted IDRs. In comparison, 1,428 (68.9%) of the 2,070 SLiM instances annotated in ELM occur in predicted IDRs (Fig. 2A). This slight difference is not statistically significant and can be explained by the large number of glycosylation sites, which often occur in globular regions, in the ELM dataset. Hence, motifs in the switches.ELM repository concur with previous observations that motifs are enriched in IDRs (22), however, as expected, they are

no more likely to occur in IDRs than annotated functional motifs in general.

Recent analyses have shown motifs to be enriched in non-constitutive exons, which enables isoform-specific rewiring of protein interaction networks (5, 24, 25). In addition, motifs involved in switches curated in switches.ELM were more likely to be located in non-constitutive exons compared to ELM instances. We find that 304 (14.7%) of the 2,070 SLiMs in ELM occurred in non-constitutive exons, whereas this was true for 126 (23.7%) of 532 SLiMs in switches.ELM (excluding the 132 motifs involved in pre-translational switches) (Fig. 2B). This difference is statistically significant, and did not result from the bias towards human SLiMs in switches.ELM or the superior coverage for human SLiMs in the available protein isoform data [112 (28.4%) of 395 human SLiMs from switches.ELM compared with 232 (19.9%) of 1,168 human SLiMs from ELM] (Fig. 2B). This indicates a tight regulation of the switch mechanisms at both pre- and post-translational levels.

Modifying enzymes could be inferred for 182 (42.6%) of the 427 modification sites involved in regulating PTM-dependent switches curated in switches.ELM (see Materials and Methods) (table S2). Of the 182 sites for which modifying enzymes could be inferred, 83 (46.6%) have experimental evidence for modification by two or more distinct enzymes, emphasizing the integrative nature of these switch mechanisms (Fig. 2C). Over 50 families of modifying enzymes were inferred to regulate the curated switches, with the CDC2 (cell division control protein 2, also known as cyclin dependent kinase 1), MAPK (mitogen activated protein kinase), and the tyrosine-protein kinase SRC families being particularly well represented.

## **Discussion**

Intrinsically disordered interfaces, and SLiMs in particular, enable rapid and robust regulation of protein function and activity in response to prevailing conditions and therefore combinatorial control of cell physiology (11, 15, 22, 23). Consequently, dynamic, tunable interactions mediated by intrinsically disordered regions of higher eukaryotic proteomes underlie a substantial portion of the ability of cellular networks to cooperatively make decisions (11, 15, 79-82). The compactness of SLiM interaction interfaces facilitates low-affinity interactions that can be easily modulated and permits high functional density. This enables SLiMs to act as switchable modules that can respond to multiple input signals to direct cell signaling in a context-dependent manner. An additional consequence of the short length and degeneracy of motifs is their propensity for de novo evolution, because a functional motif can be gained or lost by a few or even a single mutation (21, 22). The convergent evolution of new motif instances enabling recruitment of existing motif-binding domains can add functionality to proteins by rewiring their interaction networks (21, 83). Furthermore, the addition of SLiMs and modification sites to proteins enables highly complex interface regulation to evolve in a modular and additive way.

The instances of SLiM-based switch mechanisms collected in the switches.ELM resource show that many examples of motif-mediated decision-making interfaces have now been experimentally characterized. Furthermore, using the available genomic and proteomic data (43, 64-66), we showed evidence that many of the known motifs are regulated either pre- or post-translationally and revealed that modified residues were enriched around known motifs, suggesting that motif-based switches are common in higher Eukaryotes. This synergistic relationship between motifs and modification sites, and the important regulatory functionality afforded by these cooperatively regulated interfaces, offers an appealing functional explanation for the extensive intrinsic disorder content of higher eukaryotic proteomes (84) and for the many PTMs mapped to these regions (17, 19). These observations indicate that modification sites should be considered as a mechanism for

switching the intrinsic binding specificity of the underlying sequence, through a mechanism that is distinct from the canonical PTM-dependent sites recognized by PTM-binding domains such as the SH2, PTB, and Bromo domains (6).

By separating SLiM-based regulatory systems into their most basic parts, such as individual binding interfaces and PTM sites, we can gain insight into the emergent properties enabled by cooperative use of a multiplicity of these simple modules. A publicly available database of experimentally validated motif-based switches will assist research on cell regulatory mechanisms and help extend our knowledge on how these systems conditionally direct cell signaling. As we gain a better understanding of recurring regulatory mechanisms, experimental design will advance to reflect this improved information, providing us with a more complete view of these complex interfaces. By integrating the available information, the rules-based switches.ELM exploratory analysis tool can guide this design by identifying possible context-dependent modes of regulation.

The switches.ELM resource may also prove useful for systems biologists. Currently, interaction data are captured in a binary fashion in bioinformatics resources, with limited information on specific requirements depending on cellular context (85). The switches.ELM resource adds temporal, spatial, and contextual conditionality to one-dimensional interaction data to capture their context-dependency and interdependency. This conditional information can be incorporated into rule-based systems models describing the complex regulatory interaction networks in the cell, thereby increasing the quality and utility of such approaches. To facilitate data exchange and analysis, collaborative efforts are underway to integrate switches.ELM with current systems biology resources. In conclusion, switches.ELM adds to and is complementary with a growing list of resources (43, 86-89) that allow a better understanding of SLiMs and can help to direct their in-depth experimental characterization.

## Materials and Methods

### *Collection, curation and visualization of switches.ELM data*

Experimentally validated motif-based switches were collected by a literature survey of publications retrieved from directed searches in PubMed and references annotated in the ELM database (43). Relevant publications were further mined for related citations. In addition, we used the switches.ELM tool to analyze SLiM instances curated in ELM and retrieve literature that experimentally validated the suggested switch mechanisms. Users are encouraged to use the website's submission form to submit publications describing experimentally validated switches not yet annotated in the repository, which will be reviewed and added to the database.

Each switch instance is described by one or more interactions, each of which comprises two interactors that are identified by a name and identifier from an external resource [UniProtKB for proteins (66), the Chemical Entities of Biological Interest (ChEBI) database for small molecules (70), and the European Nucleotide Archive (ENA) for nucleotide sequences (71)]. For each protein interactor the binding region that mediates the interaction is annotated using an identifier referencing an external resource [ELM for SLiMs (43), and the Protein family (Pfam) (67) or InterPro databases (72) for protein domains] and the start and stop position of the binding region in the sequence. If the position of the binding region given by the authors in a publication is different from that found in the available resources, the positions that map this region to the canonical sequence in UniProtKB were chosen.

Each switch instance is categorized by a switch type and subtype that describes how the

function of a motif in one of the interactors is switched to regulate its interactions. This either involves a pre-translational mechanism, the addition of a PTM or the binding of a molecule to a specific interactor. PTMs are annotated by the position of the modified residue in the sequence and the type of modification, using the PSI-MOD controlled vocabulary for protein modifications (90), whereas effector molecules are identified by reference to an external resource (66, 70, 71). The switch mechanism can be reversible or irreversible and can be positive (either induce or enhance an interaction) or negative (either inhibit or abrogate an interaction). Interactions can be intra-molecular, mutually exclusive, require preceding interactions, or have a specific order, in which case they are numbered according to their position in the sequence of binding events.

Each switch instance is provided with a concise text description and a list of publications from which it was curated. If available, structural information [Protein Data Bank (PDB) identifier (68)] and affinity values (dissociation constant in micromolar) for the interactions are collected. Contextual information was manually curated from the literature, retrieved from manually curated data in ELM (43), or computationally inferred from external resources [subcellular localization data from GO (62), pathway data from KEGG (63), and enzymatic modification data from phosphoSitePlus (64) and phospho.ELM (65)]. Subcellular localization for 353 switch instances could be inferred from co-localization in a cellular compartment of at least two proteins involved in the switch. Similarly, 210 switch instances were mapped to KEGG pathways, such as the ErbB pathway (fig. S3), by co-occurrence in a pathway of at least two proteins involved in the switch. Enzymatic modification data could be mapped to 182 of the modification sites annotated in switches.ELM.

Biologically relevant data for the sequence-centric view of the protein region surrounding a SLiM whose function is regulated by a switch mechanism were integrated from

various resources, and all data on a switch instance output page are linked to the corresponding resources for more detailed information [ELM (43) for validated SLiMs; Pfam (67) and InterPro (72) for protein domains; phosphoSitePlus (64) and phospho.ELM (65) for modification sites; UniProtKB (66) for proteins, protein isoform data, single nucleotide polymorphisms (SNPs) and experimental mutations; ChEBI (70) for small molecules; ENA (71) for nucleotide sequences; KEGG (63) and Reactome (73) for pathways; PDB (68) for structural information; GO (62) for GO terms; PubMed for Publications; disorder scores calculated by IUPred (69); and sequence matches to ELM regular expressions (43)].

#### *Rules to analyze motifs for putative switch mechanisms*

The motif-class specific rules were developed from motif classes annotated in ELM that are known to be regulated by a specific switch mechanism and are applied to motifs that can be mapped to the regular expressions of these ELM motif classes using the CompariMotif tool (91):

- (i) Altered physicochemical compatibility switches are suggested for ELM classes whose binding is known to be induced by PTM (LIG\_14-3-3 classes, LIG\_AGCK\_PIF\_1, LIG\_BRCT\_BRCA1 classes, LIG\_BRCT\_MDC1\_1, LIG\_FHA classes, LIG\_ODPH\_VHL\_1, LIG\_PTB\_Phospho\_1, LIG\_SCF\_FBW7 classes, LIG\_SCF-TrCP1\_1, LIG\_SH2 classes, LIG\_WW\_Pin1\_4, MOD\_TYR\_ITAM, MOD\_TYR\_ITIM, MOD\_TYR\_ITSM) and substrate motifs for modifying enzymes whose recognition depends on a priming PTM on the motif (MOD\_CK1\_1, MOD\_GSK3\_1).
- (ii) Allosteric switches are suggested for motif classes known to be regulated by allosteric mechanisms (LIG\_CORNBOX, LIG\_NRBOX, TRG\_ENDOCYTIC\_2, TRG\_LysEnd\_ApsAcLL\_1).
- (iii) Pre-assembly switches are suggested for motifs that require pre-formation of a

composite binding site (LIG\_Integrin\_isoDGR\_1, LIG\_RGD, LIG\_SCF\_Skp2-Cks1\_1).

(iv) Avidity-sensing is suggested for ELM classes known to mediate multivalent binding (MOD\_TYR\_ITAM, LIG\_14-3-3 classes, LIG\_EH\_1, LIG\_AP2alpha\_2, LIG\_SH2 classes, LIG\_SH3 classes). Whereas a single MOD\_TYR\_ITAM motif mediates multivalent binding, an avidity-sensing switch mechanism is only suggested for the other motif classes when there are at least two instances of the same motif class present within a protein.

(v) Motif and domain hiding, through an intra-molecular interaction, is suggested for LIG\_SH2 classes, LIG\_SH3 classes and LIG\_PDZ classes if the protein containing the motif also contains the corresponding motif-binding domain.

General rules were developed on the basis of the curated switch instances and depended on integration of experimentally derived modification sites (64, 65), sequence information for protein isoforms (66), and experimentally validated SLiMs and sequence matches to ELM regular expressions (43):

(i) Altered physicochemical compatibility switching is suggested if there is an experimentally validated modified residue in or adjacent (within 20 residues) to the motif.

(ii) Pre-translational switches are suggested for motifs contained within a non-constitutive exon that is specifically removed from one or more isoforms of a protein.

(iii) Rheostatic switching is suggested if multiple experimentally validated modified residues are present in or adjacent (within 20 residues) to the motif.

(iv) Competition and motif hiding switches are suggested for motifs overlapping or adjacent (within 20 residues) to experimentally validated SLiMs or sequences matching the ELM regular expressions.

(v) Altered binding specificity switches are suggested for motifs overlapping or adjacent (within 20 residues) to experimentally validated SLiMs or sequences matching the ELM regular expressions if there is an experimentally validated modified residue present in or adjacent (within 20 residues) to at least one motif.

### *Analysis of ELM instances for putative switch mechanisms*

The ELM dataset (as of November 2012) consists of 2,070 experimentally validated functional SLiMs (ELM instances) in 1,325 different proteins. The ELM instances were manually curated and categorized into 178 motif classes (43). The rules applied for the analysis of ELM instances differed from those integrated in the switches.ELM analysis tool, which were less stringent to reflect the exploratory function of the tool. Altered physicochemical compatibility and rheostatic switching was suggested for ELM instances with one or more experimentally derived modification sites, respectively, in a motif or within 5 residues of the motif, instead of the 20 residues specified for the switches.ELM tool. In addition, only experimentally validated motifs annotated in ELM were considered here to determine the presence of overlapping or adjacent motifs. Sequences matching the ELM regular expressions were used by the switches.ELM tool, but were not taken into account in the analysis of ELM instances.

### *Analysis of ELM instances for enrichment of motif-proximal modification sites*

A dataset of experimentally derived modification sites (acetylation, methylation, phosphorylation, sumoylation, ubiquitylation) was retrieved from phosphoSitePlus (64) and phospho.ELM (65) through November 2012. A high-throughput modification site data set was also created from definitions made by these source databases. These modification sites were mapped to the 1,325 SLiM-containing proteins from the ELM dataset (containing 2,070 SLiMs) or to the ELM dataset excluding the MOD motif classes (1,622 SLiMs in 1,084 proteins) (43). Two sets of modification sites were created: ‘adjacent sites’ are modification sites within 5 residues of but not including a defined motif residue (690 modifications on 19,112 residues); ‘background sites’ are those greater than 5 and less

than or equal to 25 residues from a defined motif residue (2,151 modifications on 73,005 residues). The number of modification sites at each position relative to the motif was calculated and corrected for the non-uniform distribution of sample sizes at different distances from the defined motif. The proportion of modified residues at each position in the 'adjacent' and 'background' sets of positions were compared using a two-tailed Mann-Whitney U test. This is a non-parametric test for assessing whether observations in one sample tend to be larger than observations in another sample, in this case, whether amino acids within 5 residues of a motif tend to have more PTM sites than residues between 5 and 25 residues from a motif.

#### *Analysis of switches.ELM instances*

The switches.ELM dataset (as of March 2013) contained 710 experimentally validated switches comprising 817 distinct interactions that are mediated by 664 different SLiMs in 409 proteins. For all SLiMs in the switches.ELM and ELM datasets, the average disorder score was calculated as the mean IUPred score (69) for residues between the start and end position of the SLiMs. A cut-off of 0.4 was used (22), with values equal to and above this cut-off indicating disordered regions and values below 0.4 indicating globular regions. The proportion of disordered motifs in both datasets was compared using the Fisher's Exact test. Similarly, the number of SLiMs contained within a non-constitutive exon [assigned from protein isoform data in UniProtKB (66)] was determined for both datasets and the counts were compared using the Fisher's Exact test. For this analysis, the 132 SLiMs that were involved in pre-translational switches were not included in the switches.ELM dataset. Counts for the number of modifying enzymes per modification site were derived from mapping modification data (64, 65) to the modification sites involved in PTM-dependent switches. The modifying enzymes were grouped according to protein families defined in UniProtKB (66).

## Supplementary Materials

Fig. S1. Visual output of switch instances in switches.ELM.

Fig. S2. Visual output of the switches.ELM exploratory analysis tool.

Fig. S3. Switch instances in switches.ELM mapped to the ErbB signaling pathway annotated in KEGG.

Table S1. Curated data in the first version of switches.ELM.

Table S2. Inferred data in the first version of switches.ELM.

Table S3. Putative switch mechanisms for experimentally validated functional motifs curated in the ELM database.

## References and Notes

1. M. M. Babu, R. van der Lee, N. S. de Groot, J. Gsponer, Intrinsically disordered proteins: regulation and disease. *Curr. Opin. Struct. Biol.* **21**, 432-440 (2011).
2. B. Schwanhausser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, M. Selbach, Global quantification of mammalian gene expression control. *Nature* **473**, 337-342 (2011).
3. R. V. Davuluri, Y. Suzuki, S. Sugano, C. Plass, T. H. Huang, The functional consequences of alternative promoter use in mammalian genomes. *Trends Genet.* **24**, 167-177 (2008).
4. S. Stamm, S. Ben-Ari, I. Rafalska, Y. Tang, Z. Zhang, D. Toiber, T. A. Thanaraj, H. Soreq, Function of alternative splicing. *Gene* **344**, 1-20 (2005).
5. R. J. Weatheritt, T. J. Gibson, Linear motifs: lost in (pre)translation. *Trends Biochem. Sci.* **37**, 333-341 (2012).
6. T. Hunter, The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol. Cell* **28**, 730-738 (2007).
7. T. J. Gibson, Cell regulation: determined to signal discrete cooperation. *Trends*

- Biochem. Sci.* **34**, 471-482 (2009).
8. B. Novak, O. Kapuy, M. R. Domingo-Sananes, J. J. Tyson, Regulated protein kinases and phosphatases in cell cycle decisions. *Curr. Opin. Cell Biol.* **22**, 801-808 (2010).
  9. J. Pines, I. Hagan, The Renaissance or the cuckoo clock. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **366**, 3625-3634 (2011).
  10. P. E. Wright, H. J. Dyson, Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293**, 321-331 (1999).
  11. C. A. Galea, Y. Wang, S. G. Sivakolundu, R. W. Kriwacki, Regulation of cell division by intrinsically unstructured proteins: intrinsic flexibility, modularity, and signaling conduits. *Biochemistry* **47**, 7598-7609 (2008).
  12. A. K. Dunker, C. J. Oldfield, J. Meng, P. Romero, J. Y. Yang, J. W. Chen, V. Vacic, Z. Obradovic, V. N. Uversky, The unfoldomics decade: an update on intrinsically disordered proteins. *BMC Genomics* **9 Suppl 2**, S1 (2008).
  13. P. E. Wright, H. J. Dyson, Linking folding and binding. *Curr. Opin. Struct. Biol.* **19**, 31-38 (2009).
  14. C. T. Pawson, J. D. Scott, Signal integration through blending, bolstering and bifurcating of intracellular information. *Nat. Struct. Mol. Biol.* **17**, 653-658 (2010).
  15. K. Van Roey, T. J. Gibson, N. E. Davey, Motif switches: decision-making in cell regulation. *Curr. Opin. Struct. Biol.* **22**, 378-385 (2012).
  16. M. Lewitzky, P. C. Simister, S. M. Feller, Beyond 'furballs' and 'dumpling soups' - towards a molecular architecture of signaling complexes and networks. *FEBS Letters* **586**, 2740-2750 (2012).
  17. L. M. Iakoucheva, P. Radivojac, C. J. Brown, T. R. O'Connor, J. G. Sikes, Z. Obradovic, A. K. Dunker, The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32**, 1037-1049 (2004).
  18. H. Xie, S. Vucetic, L. M. Iakoucheva, C. J. Oldfield, A. K. Dunker, Z. Obradovic, V.N. Uversky, Functional anthology of intrinsic disorder. 3. Ligands, post-

translational modifications, and diseases associated with intrinsically disordered proteins. *J. Proteome. Res.* **6**, 1917-1932 (2007).

19. M. O. Collins, L. Yu, I. Campuzano, S. G. Grant, J. S. Choudhary, Phosphoproteomic analysis of the mouse brain cytosol reveals a predominance of protein phosphorylation in regions of intrinsic sequence disorder. *Mol. Cell Proteomics* **7**, 1331-1348 (2008).
20. F. Diella, N. Haslam, C. Chica, A. Budd, S. Michael, N. P. Brown, G. Trave, T. J. Gibson, Understanding eukaryotic linear motifs and their role in cell signaling and regulation. *Front. Biosci.* **13**, 6580-6603 (2008).
21. V. Neduva, R. B. Russell, Linear motifs: evolutionary interaction switches. *FEBS Lett.* **579**, 3342-3345 (2005).
22. N. E. Davey, K. Van Roey, R. J. Weatheritt, G. Toedt, B. Uyar, B. Altenberg, A. Budd, F. Diella, H. Dinkel, T. J. Gibson, Attributes of short linear motifs. *Mol. Biosyst.* **8**, 268-281 (2012).
23. E. Akiva, G. Friedlander, Z. Itzhaki, H. Margalit, A dynamic view of domain-motif interactions. *PLoS Comput. Biol.* **8**, e1002341 (2012).
24. R. J. Weatheritt, N. E. Davey, T. J. Gibson, Linear motifs confer functional diversity onto splice variants. *Nucleic Acids Res.* **40**, 7123-7131 (2012).
25. M. Buljan, G. Chalancon, S. Eustermann, G. P. Wagner, M. Fuxreiter, A. Bateman, M. M. Babu, Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. *Mol. Cell* **46**, 871-883 (2012).
26. M. J. Davis, C. J. Shin, N. Jing, M. A. Ragan, Rewiring the dynamic interactome. *Mol. Biosyst.* **8**, 2054-2066, 2013 (2012).
27. J. D. Ellis, M. Barrios-Rodiles, R. Colak, M. Irimia, T. Kim, J. A. Calarco, X. Wang, Q. Pan, D. O'Hanlon, P. M. Kim, J. L. Wrana, B. J. Blencowe, Tissue-specific alternative splicing remodels protein-protein interaction networks. *Mol. Cell* **46**, 884-892 (2012).

28. C. Shin, J. L. Manley, Cell signalling and the control of pre-mRNA splicing. *Nat. Rev. Mol. Cell Biol.* **5**, 727-738 (2004).
29. T. Pawson, M. Raina, P. Nash, Interaction domains: from simple binding events to complex cellular behavior. *FEBS Lett.* **513**, 2-10 (2002).
30. B. T. Seet, I. Dikic, M. M. Zhou, T. Pawson, Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Biol.* **7**, 473-483 (2006).
31. Y. L. Deribe, T. Pawson, I. Dikic, Post-translational modifications in signal integration. *Nat. Struct. Mol. Biol.* **17**, 666-672 (2010).
32. C. Salazar, A. Brummer, L. Alberghina, T. Hofer, Timing control in regulatory networks by multisite protein modifications. *Trends Cell. Biol.* **20**, 634-641 (2010).
33. P. Di Lello, L. M. Jenkins, T. N. Jones, B. D. Nguyen, T. Hara, H. Yamaguchi, J. D. Dikeakos, E. Appella, P. Legault, J. G. Omichinski, Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIF and the activation domain of p53. *Mol. Cell* **22**, 731-740 (2006).
34. C. W. Lee, J. C. Ferreon, A. C. Ferreon, M. Arai, P. E. Wright, Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19290-19295 (2010).
35. A. Whitty, Cooperativity and biological complexity. *Nat. Chem. Biol.* **4**, 435-439 (2008).
36. B. Kostecky, A. T. Saurin, A. Purkiss, P. J. Parker, N. Q. McDonald, Recognition of an intra-chain tandem 14-3-3 binding site within PKCepsilon. *EMBO Rep.* **10**, 983-989 (2009).
37. L. Balagopalan, N. P. Coussens, E. Sherman, L. E. Samelson, C. L. Sommers, The LAT story: a tale of cooperativity, coordination, and choreography. *Cold Spring Harb. Perspect. Biol.* **2**, a005512 (2010).
38. P. Li, S. Banjade, H. C. Cheng, S. Kim, B. Chen, L. Guo, M. Llaguno, J. V. Hollingsworth, D. S. King, S. F. Banani, P. S. Russo, Q. X. Jiang, B. T. Nixon, M. K.

- Rosen, Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336-340 (2012).
39. C. L. Oxley, N. J. Anthis, E. D. Lowe, I. Vakonakis, I. D. Campbell, K. L. Wegener, An integrin phosphorylation switch: the effect of beta3 integrin tail phosphorylation on Dok1 and talin binding. *J. Biol. Chem.* **283**, 5420-5426 (2008).
  40. C. J. Moore, S. J. Winder, Dystroglycan versatility in cell adhesion: a tale of multiple motifs. *Cell Commun. Signal.* **8**, 3 (2010).
  41. E. Aragon, N. Goerner, A. I. Zaromytidou, Q. Xi, A. Escobedo, J. Massague, M. J. Macias, A Smad action turnover switch operated by WW domain readers of a phosphoserine code. *Genes Dev.* **25**, 1275-1288 (2011).
  42. M. Koivomagi, E. Valk, R. Venta, A. Iofik, M. Lepiku, E. R. Balog, S. M. Rubin, D. O. Morgan, M. Loog, Cascades of multisite phosphorylation control Sic1 destruction at the onset of S phase. *Nature* **480**, 128-131 (2011).
  43. H. Dinkel, S. Michael, R. J. Weatheritt, N. E. Davey, K. Van Roey, B. Altenberg, G. Toedt, B. Uyar, M. Seiler, A. Budd, L. Jodicke, M. A. Dammert, C. Schroeter, M. Hammer, T. Schmidt, P. Jehl, C. McGuigan, M. Dymecka, C. Chica, K. Luck, A. Via, A. Chatr-Aryamontri, N. Haslam, G. Grebnev, R. J. Edwards, M. O. Steinmetz, H. Meiselbach, F. Diella, T. J. Gibson, ELM--the database of eukaryotic linear motifs. *Nucleic Acids Res.* **40**, D242-251 (2012).
  44. Z. Gao, H. Y. Poon, L. Li, X. Li, E. Palmesino, D. Glubrecht, K. Colwill, I. Dutta, A. Kania, T. Pawson, R. Godbout, Splice-mediated motif switching regulates Disabled-1 phosphorylation and SH2 domain interactions. *Mol. Cell Biol.* **32**, 2794-2808 (2012).
  45. P. Stehmeier, S. Muller, Phospho-regulated SUMO interaction modules connect the SUMO system to CK2 signaling. *Mol. Cell* **33**, 400-409 (2009).
  46. A. Kaffman, N. M. Rank, E. K. O'Shea, Phosphorylation regulates association of the transcription factor Pho4 with its import receptor Pse1/Kap121. *Genes Dev.* **12**, 2673-2683 (1998).

47. A. W. Fenton, Allosteric: an illustrated definition for the 'second secret of life'. *Trends Biochem. Sci.* **33**, 420-425 (2008).
48. S. B. Padrick, M. K. Rosen, Physical mechanisms of signal integration by WASP family proteins. *Annu. Rev. Biochem.* **79**, 707-735 (2010).
49. R. Nussinov, C. J. Tsai, F. Xin, P. Radivojac, Allosteric post-translational modification codes. *Trends Biochem. Sci.* **37**, 447-455 (2012).
50. N. J. Anthis, K. L. Wegener, D. R. Critchley, I. D. Campbell, Structural diversity in integrin/talin interactions. *Structure* **18**, 1654-1666 (2010).
51. D. Bodmer, M. Ascano, R. Kuruvilla, Isoform-specific dephosphorylation of dynamin1 by calcineurin couples neurotrophin receptor endocytosis to axonal growth. *Neuron* **70**, 1085-1099 (2011).
52. B. Zhao, L. Li, Q. Lu, L. H. Wang, C. Y. Liu, Q. Lei, K. L. Guan, Angiotensin is a novel Hippo pathway component that inhibits YAP oncoprotein. *Genes Dev.* **25**, 51-63 (2011).
53. M. Ernkvist, O. Birot, I. Sinha, N. Veitonmaki, S. Nystrom, K. Aase, L. Holmgren, Differential roles of p80- and p130-angiotensin in the switch between migration and stabilization of endothelial cells. *Biochim. Biophys. Acta* **1783**, 429-437 (2008).
54. Q. Sun, R. A. Jackson, C. Ng, G. R. Guy, J. Sivaraman, Additional serine/threonine phosphorylation reduces binding affinity but preserves interface topography of substrate proteins to the c-Cbl TKB domain. *PLoS One* **5**, e12819 (2010).
55. B. Hao, N. Zheng, B. A. Schulman, G. Wu, J. J. Miller, M. Pagano, N. P. Pavletich, Structural basis of the Cks1-dependent recognition of p27(Kip1) by the SCF(Skp2) ubiquitin ligase. *Mol. Cell* **20**, 9-19 (2005).
56. K. Russo, E. Di Stasio, G. Macchia, G. Rosa, A. Brancaccio, T. C. Petrucci, Characterization of the beta-dystroglycan-growth factor receptor 2 (Grb2) interaction. *Biochem. Biophys. Res. Commun.* **274**, 93-98 (2000).
57. X. Huang, F. Poy, R. Zhang, A. Joachimiak, M. Sudol, M. J. Eck, Structure of a WW

- domain containing fragment of dystrophin in complex with beta-dystroglycan. *Nat. Struct. Biol.* **7**, 634-638 (2000).
58. N. J. Godde, G. M. D'Abaco, L. Paradiso, U. Novak, Efficient ADAM22 surface expression is mediated by phosphorylation-dependent interaction with 14-3-3 protein family members. *J. Cell Sci.* **119**, 3296-3305 (2006).
  59. T. Sugi, T. Oyama, T. Muto, S. Nakanishi, K. Morikawa, H. Jingami, Crystal structures of autoinhibitory PDZ domain of Tamalin: implications for metabotropic glutamate receptor trafficking regulation. *EMBO J.* **26**, 2192-2205 (2007).
  60. A. Pineda-Lucena, C. S. Ho, D. Y. Mao, Y. Sheng, R. C. Laister, R. Muhandiram, Y. Lu, B. T. Seet, S. Katz, T. Szyperski, L. Z. Penn, C. H. Arrowsmith, A structure-based model of the c-Myc/Bin1 protein interaction shows alternative splicing of Bin1 and c-Myc phosphorylation are key binding determinants. *J. Mol. Biol.* **351**, 182-194 (2005).
  61. K. Nagaki, H. Yamamura, S. Shimada, T. Saito, S. Hisanaga, M. Taoka, T. Isobe, T. Ichimura, 14-3-3 Mediates phosphorylation-dependent inhibition of the interaction between the ubiquitin E3 ligase Nedd4-2 and epithelial Na<sup>+</sup> channels. *Biochemistry* **45**, 6733-6740 (2006).
  62. M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, G. Sherlock, Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25-29 (2000).
  63. M. Kanehisa, S. Goto, Y. Sato, M. Furumichi, M. Tanabe, KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**, D109-114 (2012).
  64. P. V. Hornbeck, J. M. Kornhauser, S. Tkachev, B. Zhang, E. Skrzypek, B. Murray, V. Latham, M. Sullivan, PhosphoSitePlus: a comprehensive resource for investigating the

- structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* **40**, D261-270 (2012).
65. H. Dinkel, C. Chica, A. Via, C. M. Gould, L. J. Jensen, T. J. Gibson, F. Diella, Phospho.ELM: a database of phosphorylation sites--update 2011. *Nucleic Acids Res.* **39**, D261-267 (2011).
  66. T. U. Consortium, Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* **40**, D71-75 (2012).
  67. M. Punta, P. C. Coghill, R. Y. Eberhardt, J. Mistry, J. Tate, C. Boursnell, N. Pang, K. Forslund, G. Ceric, J. Clements, A. Heger, L. Holm, E. L. Sonnhammer, S. R. Eddy, A. Bateman, R. D. Finn, The Pfam protein families database. *Nucleic Acids Res.* **40**, D290-301 (2012).
  68. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, The Protein Data Bank. *Nucleic Acids Res.* **28**, 235-242 (2000).
  69. Z. Dosztanyi, V. Csizmok, P. Tompa, I. Simon, IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21**, 3433-3434 (2005).
  70. P. de Matos, R. Alcantara, A. Dekker, M. Ennis, J. Hastings, K. Haug, I. Spiteri, S. Turner, C. Steinbeck, Chemical Entities of Biological Interest: an update. *Nucleic Acids Res.* **38**, D249-254 (2010).
  71. R. Leinonen, R. Akhtar, E. Birney, L. Bower, A. Cerdeno-Tarraga, Y. Cheng, I. Cleland, N. Faruque, N. Goodgame, R. Gibson, G. Hoad, M. Jang, N. Pakseresht, S. Plaister, R. Radhakrishnan, K. Reddy, S. Sobhany, P. Ten Hoopen, R. Vaughan, V. Zalunin, G. Cochrane, The European Nucleotide Archive. *Nucleic Acids Res.* **39**, D28-31 (2011).
  72. S. Hunter, P. Jones, A. Mitchell, R. Apweiler, T. K. Attwood, A. Bateman, T. Bernard, D. Binns, P. Bork, S. Burge, E. de Castro, P. Coghill, M. Corbett, U. Das, L.

- Daugherty, L. Duquenne, R. D. Finn, M. Fraser, J. Gough, D. Haft, N. Hulo, D. Kahn, E. Kelly, I. Letunic, D. Lonsdale, R. Lopez, M. Madera, J. Maslen, C. McAnulla, J. McDowall, C. McMenamin, H. Mi, P. Mutowo-Muellenet, N. Mulder, D. Natale, C. Orengo, S. Pesseat, M. Punta, A. F. Quinn, C. Rivoire, A. Sangrador-Vegas, J. D. Selengut, C. J. Sigrist, M. Scheremetjew, J. Tate, M. Thimmajananathan, P. D. Thomas, C. H. Wu, C. Yeats, S. Y. Yong, InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res.* **40**, D306-312 (2012).
73. D. Croft, G. O'Kelly, G. Wu, R. Haw, M. Gillespie, L. Matthews, M. Caudy, P. Garapati, G. Gopinath, B. Jassal, S. Jupe, I. Kalatskaya, S. Mahajan, B. May, N. Ndegwa, E. Schmidt, V. Shamovsky, C. Yung, E. Birney, H. Hermjakob, P. D'Eustachio, L. Stein, Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* **39**, D691-697 (2011).
74. C. W. Chang, H. Y. Chou, Y. S. Lin, K. H. Huang, C. J. Chang, T. C. Hsu, S. C. Lee, Phosphorylation at Ser473 regulates heterochromatin protein 1 binding and corepressor function of TIF1beta/KAP1. *BMC Mol. Biol.* **9**, 61 (2008).
75. N. Dephoure, C. Zhou, J. Villen, S. A. Beausoleil, C. E. Bakalarski, S. J. Elledge, S. P. Gygi, A quantitative atlas of mitotic phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 10762-10767 (2008).
76. L. Wei, B. H. Han, Y. Li, C. L. Keogh, D. M. Holtzman, S. P. Yu, Cell death mechanism and protective effect of erythropoietin after focal ischemia in the whisker-barrel cortex of neonatal rats. *J. Pharmacol. Exp. Ther.* **317**, 109-116 (2006).
77. F. Zhang, S. Wang, G. Cao, Y. Gao, J. Chen, Signal transducers and activators of transcription 5 contributes to erythropoietin-mediated neuroprotection against hippocampal neuronal death after transient global cerebral ischemia. *Neurobiol. Dis.* **25**, 45-53 (2007).
78. Y. Nakamura, N. Komatsu, H. Nakauchi, A truncated erythropoietin receptor that fails to prevent programmed cell death of erythroid cells. *Science* **257**, 1138-1141

(1992).

79. H. J. Dyson, P. E. Wright, Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197-208 (2005).
80. A. K. Dunker, I. Silman, V. N. Uversky, J. L. Sussman, Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* **18**, 756-764 (2008).
81. J. R. Perkins, I. Diboun, B. H. Dessailly, J. G. Lees, C. Orengo, Transient protein-protein interactions: structural, functional, and network properties. *Structure* **18**, 1233-1243 (2010).
82. P. Tompa, Unstructural biology coming of age. *Curr. Opin. Struct. Biol.* **21**, 419-425 (2011).
83. R. P. Bhattacharyya, A. Remenyi, B. J. Yeh, W. A. Lim, Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu. Rev. Biochem.* **75**, 655-680 (2006).
84. J. J. Ward, J. S. Sodhi, L. J. McGuffin, B. F. Buxton, D. T. Jones, Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **337**, 635-645 (2004).
85. P. Aloy, R. B. Russell, Structural systems biology: modelling protein interactions. *Nat. Rev. Mol. Cell Biol.* **7**, 188-197 (2006).
86. V. Neduva, R. B. Russell, DILIMOT: discovery of linear motifs in proteins. *Nucleic Acids Res.* **34**, W350-355 (2006).
87. N. E. Davey, N. J. Haslam, D. C. Shields, R. J. Edwards, SLiMSearch 2.0: biological context for short linear motifs in proteins. *Nucleic Acids Res.* **39**, W56-60 (2011).
88. R. J. Weatheritt, P. Jehl, H. Dinkel, T. J. Gibson, iELM--a web server to explore short linear motif-mediated interactions. *Nucleic Acids Res.* **40**, W364-369 (2012).
89. A. N. Nguyen Ba, B. J. Yeh, D. van Dyk, A. R. Davidson, B. J. Andrews, E. L. Weiss, A. M. Moses, Proteome-wide discovery of evolutionary conserved sequences in disordered regions. *Sci. Signal.* **5**, rs1 (2012).

90. L. Montecchi-Palazzi, R. Beavis, P. A. Binz, R. J. Chalkley, J. Cottrell, D. Creasy, J. Shofstahl, S. L. Seymour, J. S. Garavelli, The PSI-MOD community standard for representation of protein modification data. *Nat. Biotechnol.* **26**, 864-866 (2008).
91. R. J. Edwards, N. E. Davey, D. C. Shields, CompariMotif: quick and easy comparisons of sequence motifs. *Bioinformatics* **24**, 1307-1309 (2008).
92. A. S. Kim, L. T. Kakalis, N. Abdul-Manan, G. A. Liu, M. K. Rosen, Autoinhibition and activation mechanisms of the Wiskott-Aldrich syndrome protein. *Nature* **404**, 151-158 (2000).
93. J. Y. Bu, A. S. Shaw, A. C. Chan, Analysis of the interaction of ZAP-70 and syk protein-tyrosine kinases with the T-cell antigen receptor by plasmon resonance. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5106-5110 (1995).
94. R. Sunder-Plassmann, F. Lialios, M. Madsen, S. Koyasu, E. L. Reinherz, Functional analysis of immunoreceptor tyrosine-based activation motif (ITAM)-mediated signal transduction: the two YxxL segments within a single CD3zeta-ITAM are functionally distinct. *Eur. J. Immunol.* **27**, 2001-2009 (1997).

**Acknowledgements:** We would like to thank Aino Järvelin, Markus Seiler, Bora Uyar and Niall Haslam for critical evaluation of the manuscript and switches.ELM resource. We would also like to thank our colleagues in the motif biology field for useful discussions and suggestions. **Funding:** This work was supported by an FP7 Health grant (No. 242129; SyBoSS) from the European Commission (K.V.R.), and an EMBL International Ph.D. Programme fellowship (R.J.W.), an EMBL Interdisciplinary Postdoctoral fellowship (EIPOD) (N.E.D.). **Author contributions:** K.V.R. and N.E.D. designed the study; K.V.R. and N.E.D. wrote the manuscript; K.V.R. and R.J.W. curated the data; N.E.D. and H.D. developed the switches.ELM website; N.E.D. performed the analyses; T.J.G. provided advice and edited the manuscript; and all authors discussed the results and commented on the manuscript. **Competing interests:** The authors declare that they have no competing

interests.

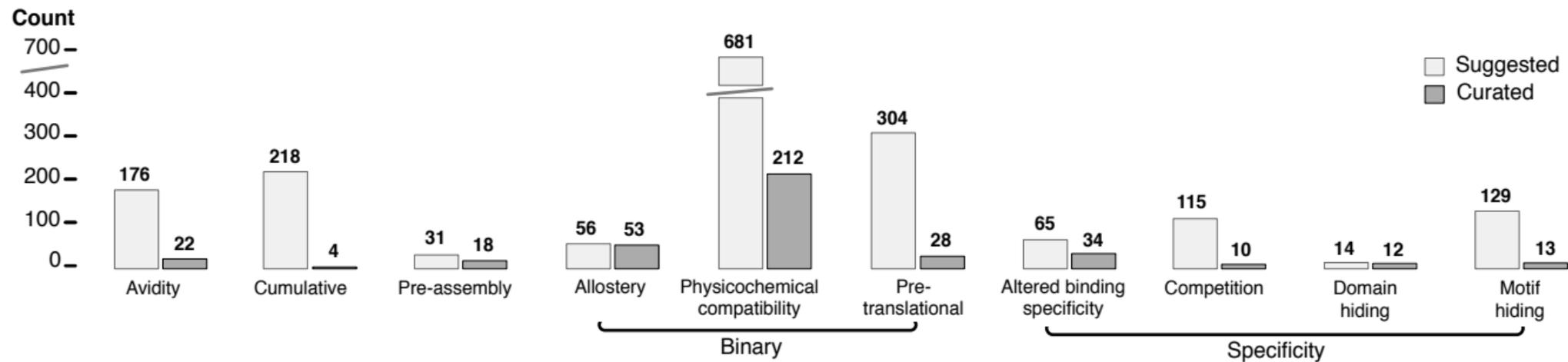
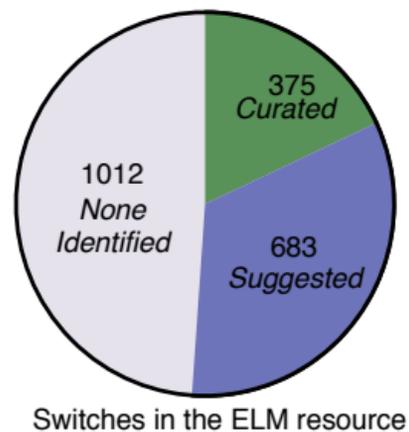
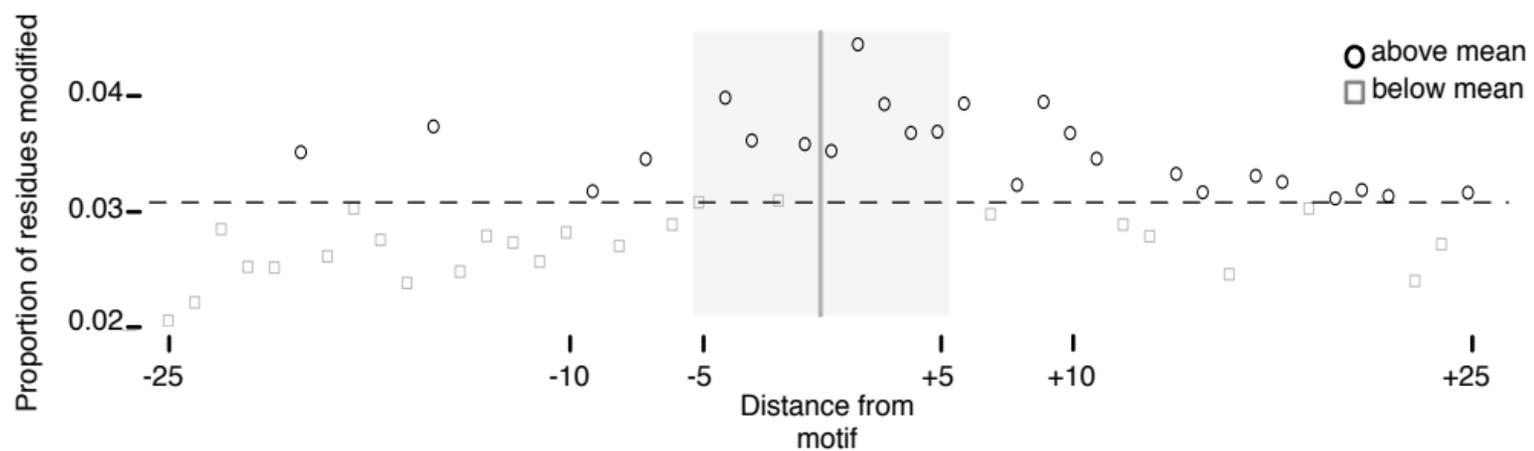
## FIGURE LEGENDS

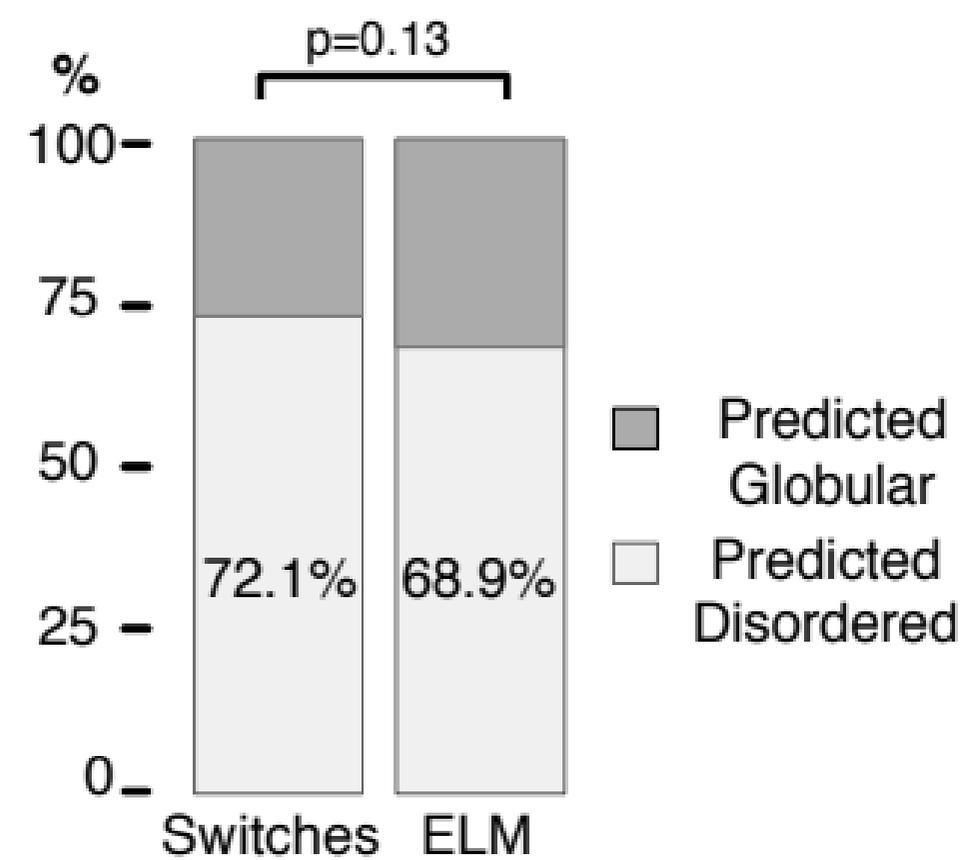
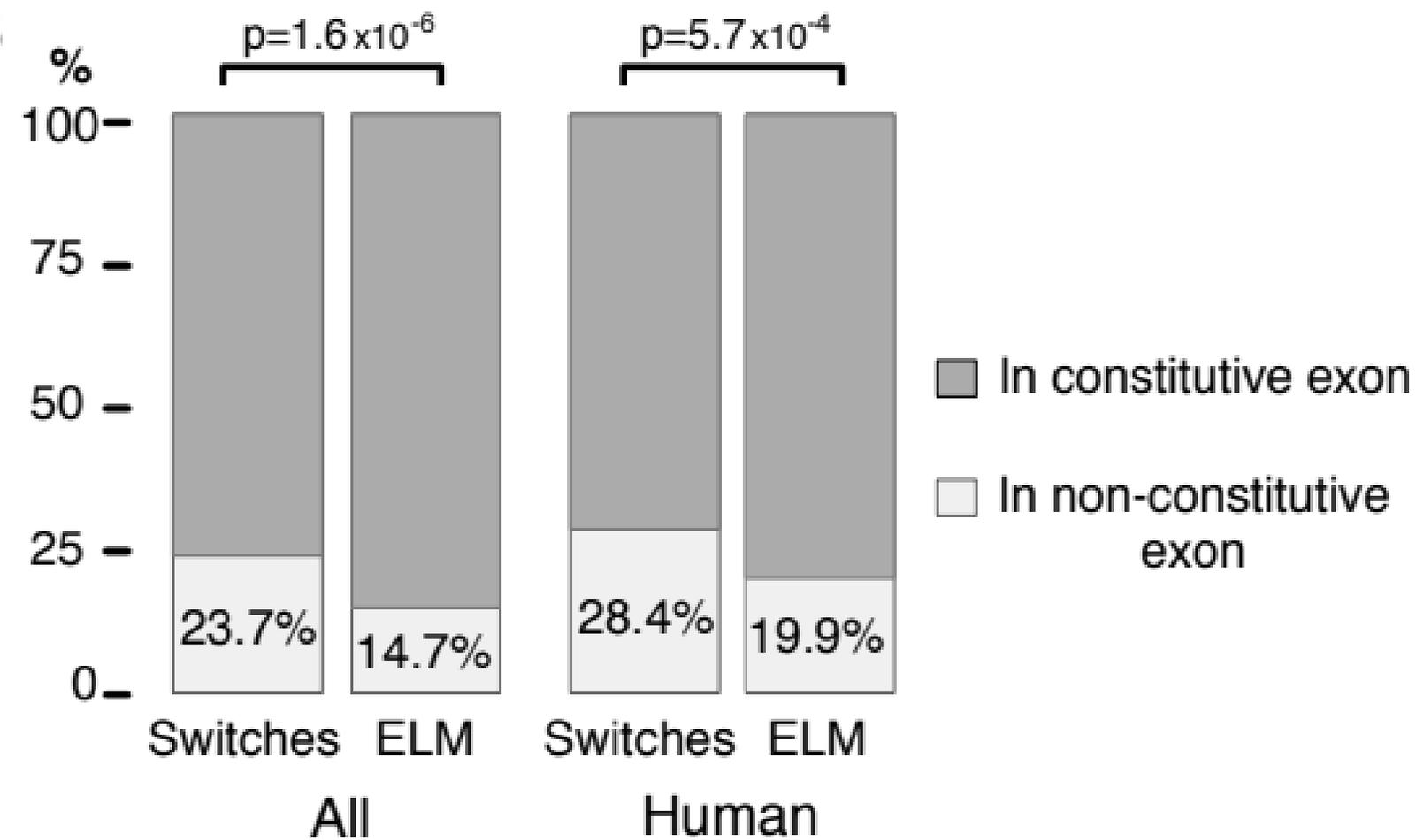
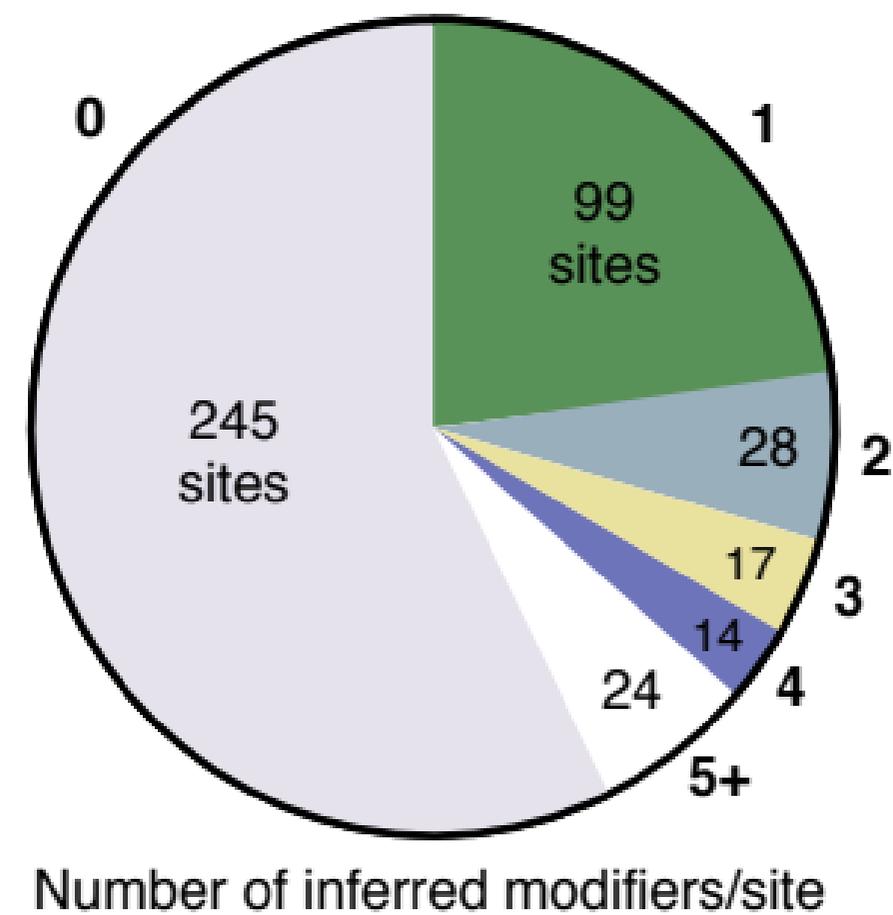
**Fig. 1.** Analysis of ELM instances. (A) Bar plot of the number of ELM instances suggested to be regulated by a specific switch mechanism (light grey) and the number of these already curated in switches.ELM (dark grey). (B) Pie chart of the number of ELM instances known to be regulated by a switch mechanism (Curated), suggested to be regulated by a switch mechanism using the switches.ELM rules (Suggested), and with no information that indicates regulation by a switch mechanism (None identified). (C) Plot of the proportion of known PTM sites flanking ELM instances. The central vertical line denotes the position of the motif. 25 residues before the first and after the last residue of the motif are plotted. The grey-shaded region indicates the adjacent positions (within 5 residues of the motif). The horizontal line denotes the mean proportion of residues modified for the 50 residues shown. Black circles denote positions that have more than the mean proportion of PTM sites, whereas grey squares denote positions that have less. Mann-Whitney tests show a statistically significant enrichment of PTM sites adjacent to known motifs ( $p=1.48 \times 10^{-4}$ ,  $U=378.5$ ,  $NA=11$ ,  $NB=40$ ;  $p=1.12 \times 10^{-3}$ ,  $U=353.5$ ,  $NA=11$ ,  $NB=40$  when only high-throughput PTM data were used; and  $p=3.56 \times 10^{-3}$ ,  $U=337.5$ ,  $NA=11$ ,  $NB=40$  when ELM MOD classes were omitted).

**Fig. 2.** Analysis of switches.ELM instances. (A) Proportion of motifs in the switches.ELM resource (479 out of 664 annotated SLiMs) and motifs in the ELM database (1,428 out of 2,070 annotated SLiMs) predicted to be intrinsically disordered. A Fisher's Exact test indicates the difference is not statistically significant ( $p=0.13$ ). (B) Proportion of motifs in the switches.ELM resource (All: 126 out of 532 annotated SLiMs; Human: 112 out of 395 annotated SLiMs) and motifs in the ELM database (All: 304 out of 2,070 annotated SLiMs; Human: 232 out of 1,168 annotated SLiMs) located in non-constitutive exons. A Fisher's Exact test indicates a statistically significant enrichment of SLiMs located in non-constitutive exons in switches.ELM compared to ELM, both when all motifs in the two datasets were considered, and when only human motifs were considered (All:  $p=1.57 \times 10^{-6}$ ; Human:  $p=5.66 \times 10^{-4}$ ). (C) Pie chart of the number of modification sites involved in PTM-dependent switches curated in switches.ELM grouped by the number of modifying enzymes inferred per modification site.

**Table 1:** Examples for the different SLiM-based switch mechanism types and

subtypes. Defined residues of a motif are in bold. Residues requiring phosphorylation in order for the motif to be functional are preceded by a lowercase p, whereas a residue that will be phosphorylated by a kinase that interacts with and modifies a motif is written between brackets.

**A****B****C**

**A****B****C**

**Table 1: Examples for the different SLiM-based switch mechanism types and subtypes.** Defined residues of a motif are in bold. Residues that are required to be phosphorylated in order for the motif to be functional are preceded by a lowercase p, while a residue that will be phosphorylated by a kinase that interacts with and modifies a motif is written between brackets.

Switch type – subtype (number of instances in switches.ELM)	Motif- containing protein <sup>1</sup>	Motif interface	Domain- containing protein <sup>1</sup>	Domain interface	Description	References
<b>Binary switch – Altered physicochemical compatibility (360)</b>	Dab1	185- <b>pYQTI</b> -188	Src	SH2 domain	Binding to Src is induced by tyrosine phosphorylation of the SH2-binding motif in Dab1.	(44)
<b>Binary switch – Allostery (59)</b>	WAS	466-LVGALMHVMQK-476	WAS	P21-Rho- binding domain	Binding of the small GTPase CDC42 to WAS allosterically relieves an auto-inhibitory intramolecular interaction, thereby activating WAS.	(48, 92)
<b>Binary switch – Pre-translational (172)</b>	AMOT	239- <b>PPEY</b> -242	YAP1	WW domain	This splice isoform-specific WW-binding motif of the p130 isoform of AMOT binds to the first WW domain of the transcription co-activator YAP1, sequestering it in the cytosol.	(52, 53)
<b>Cumulative switch – Rheostatic (11)</b>	TP53	50- <b>IEQWFTE</b> -56	GTF2H1	TFIIH p62 subunit, N- terminal domain	Successive phosphorylation of TP53 enhances its affinity for GTF2H1, an interaction involved in	(33)

					activation of transcription initiation and elongation by TP53.	
<b>Pre-assembly switch – Composite binding site formation (17)</b>	CKS1B	Cyclin-dependent kinase regulatory subunit	SKP2	Leucine-rich repeats	Binding of the phosphorylated CDK inhibitor CDKN1B to the SCF <sup>Skp2</sup> ubiquitin ligase complex requires association of the F-box protein SKP2 with the regulatory protein CKS1, since this results in the formation of a composite binding site for CDKN1B.	(55)
	CDKN1B	184-VEQ(T)PKK-190	CDK2	Protein kinase domain		
	CDKN1B	183-SVEQpTPKK-190	CKS1B-SKP2			
<b>Avidity-sensing switch (24)</b>	CD3Z	69-NQLpYNEL-75	ZAP70	SH2 domain <sup>2</sup>	Dual phosphorylation of the ITAM motif in CD3Z results in high-avidity binding to the tandem SH2 domains of ZAP70.	(93, 94)
	CD3Z	83-pYDVL-86	ZAP70	SH2 domain <sup>2</sup>		
<b>Specificity switch – Competition (10)</b>	DAG1	889-PPPY-892	DMD	WW domain	The WW-binding motif for DMD and the SH3-binding motif for GRB2 on DAG1 overlap, making their interactions mutually exclusive and competitive.	(40, 56, 57)
	DAG1	888-SPPPYVP-894	GRB2	SH3 domain		
<b>Specificity switch – Altered binding specificity (23)</b>	ITGB3	767-TANNPLYK-774	TLN1	PTB domain (IRS-1 type)	Phosphorylation of Y773 in ITGB3 switches the specificity of ITGB3 from TLN1 to DOK1, with a 2-fold decrease of the affinity for TLN1 and close to a 400-fold increase of the affinity for DOK1, resulting in negative regulation of integrin function.	(39)
	ITGB3	767-TANNPLpY-773	DOK1	PTB domain (IRS-1 type)		
<b>Specificity switch – Motif hiding (12)</b>	ADAM22	831-RSNpSWQ-836 / 854-RSNpSTE-859	YWHAB	14-3-3 protein	Phosphorylation-induced binding of YWHAB to ADAM22 blocks ER	(58)

	ADAM22	829- <b>RPR</b> -831 / 850- <b>RFR</b> -852 / 852- <b>RPR</b> -854	COPB1	Adaptin N terminal region	retention motifs in ADAM22 and regulates its transport to the membrane.	
<b>Specificity switch – Domain hiding (17)</b>	NEDD4L	465- <b>RSLSSP</b> -470	YWHAH	14-3-3 protein	Phosphorylation of the ubiquitin ligase NEDD4L induces binding to YWHAH, which prevents interaction of the NEDD4L WW domains with the WW-binding motif of SCNN1G. As a result, SCNN1G is not marked for degradation and SCNN1G-mediated sodium currents increase.	(61)
	SCNN1G	624- <b>PPKY</b> -627	NEDD4L	WW domain		

<sup>1</sup>Proteins (from human unless specified otherwise) are defined by their gene names. Full names are: Dab1 (mouse): Disabled homolog 1; Src (mouse): Neuronal proto-oncogene tyrosine-protein kinase Src; WAS: Wiskott-Aldrich syndrome protein; CDC42: Cell division control protein 42 homolog; AMOT: Angiomotin; YAP1: Yorkie homolog; TP53: Cellular tumor antigen p53; GTF2H1: General transcription factor IIIH subunit 1; CKS1B: Cyclin-dependent kinases regulatory subunit 1; SKP2: S-phase kinase-associated protein 2; CDK2: Cyclin-dependent kinase 2; CDKN1B: Cyclin-dependent kinase inhibitor 1B; CD3Z: T-cell surface glycoprotein CD3  $\zeta$  chain; ZAP70: Tyrosine-protein kinase ZAP-70; DAG1: Dystroglycan; DMD: Dystrophin; GRB2: Growth factor receptor-bound protein 2; ITGB3: Integrin  $\beta$ -3; TLN1: Talin-1; DOK1: Docking protein 1; ADAM22: Disintegrin and metalloproteinase domain-containing protein 22; YWHAB: 14-3-3 protein  $\beta/\alpha$ ; COPB1: Coatamer subunit  $\beta$ ; NEDD4L: E3 ubiquitin-protein ligase NEDD4-like; YWHAH: 14-3-3 protein  $\eta$ ; SCNN1G: Amiloride-sensitive sodium channel subunit  $\gamma$ .

<sup>2</sup>ZAP70 contains two SH2 domains, each of which binds to one phosphorylated tyrosine residue within the ITAM motif.

## Supplementary Materials for

### **The switches.ELM Resource: A Compendium of Conditional Regulatory Interaction Interfaces**

Kim Van Roey, Holger Dinkel, Robert J. Weatheritt, Toby J. Gibson, Norman E. Davey\*

\*Corresponding author. E-mail: [davey@embl.de](mailto:davey@embl.de)

Published 2 April 2013, *Sci. Signal.* **6**, rs7 (2013)  
DOI: 10.1126/scisignal.2003345

#### **This PDF file includes:**

Fig. S1. Visual output of switch instances in switches.ELM.

Fig. S2. Visual output of the switches.ELM exploratory analysis tool.

Fig. S3. Switch instances in switches.ELM mapped to the ErbB signaling pathway annotated in KEGG.

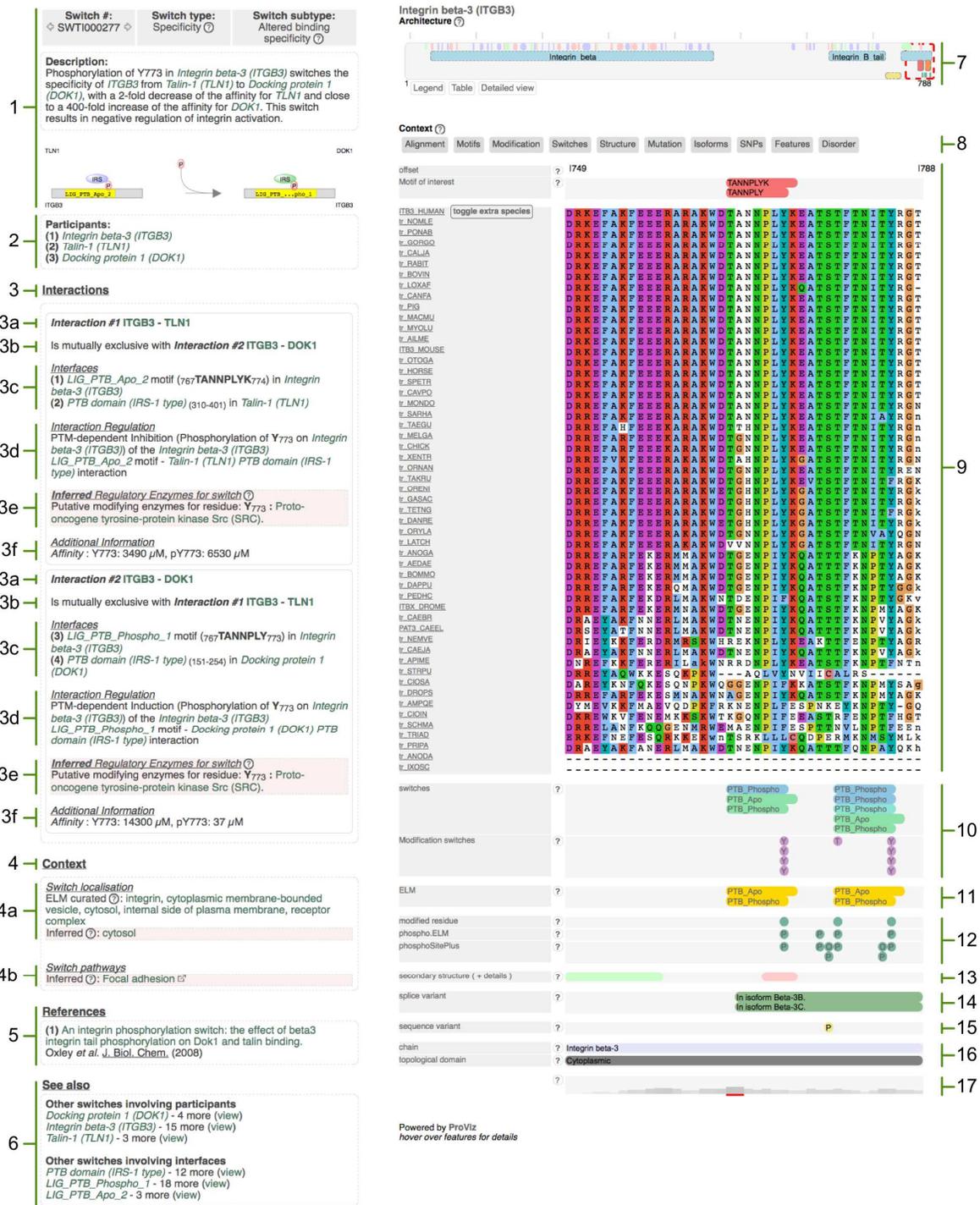
#### **Other Supplementary Material for this manuscript includes the following:**

(available at [www.sciencesignaling.org/cgi/content/full/6/269/rs7/DC1](http://www.sciencesignaling.org/cgi/content/full/6/269/rs7/DC1))

Table S1 (Microsoft Excel format). Curated data in the first version of switches.ELM.

Table S2 (Microsoft Excel format). Inferred data in the first version of switches.ELM.

Table S3 (Microsoft Excel format). Putative switch mechanisms for experimentally validated functional motifs curated in the ELM database.



**Fig. S1. Visual output of switch instances in switches.ELM. (1)** Type, text description and graphical presentation of the switch instance. **(2)** List of interaction participants. **(3)**

List of interactions constituting the switch instance. **(3a)** Interaction named by its two participants (may be numbered by sequence of binding events). **(3b)** List of mutually exclusive or required interactions. **(3c)** List of binding interfaces that mediate the interaction. **(3d)** Description for how an interaction is regulated. **(3e)** Curated and inferred modifying enzymes involved in regulation of the switch instance. **(3f)** Additional information. **(4)** Contextual information. **(4a)** Curated and inferred localization data. **(4b)** Curated and inferred pathways involving the switch. **(5)** List of relevant publications describing experimental data for the switch instance. **(6)** List of other switch instances involving any of the participants or interaction interfaces. **(7)** Diagram of whole protein architecture with annotated features. **(8)** Toggle categories of information. **(9)** Expandable alignment of the region of interest. **(10)** Validated motifs and PTMs that participate in a curated switch instance. **(11)** Validated motifs that are curated in ELM. **(12)** Validated modification sites. If available, experimentally mutated sites are shown. **(13)** Structural information. **(14)** Protein isoforms. **(15)** SNPs. **(16)** Additional sequence features. **(17)** IUPred disorder scores.

**Protein:** Cyclin-dependent kinase inhibitor 1 (CDKN1A)  
**Species:** Homo sapiens  
**Alignment:** View full protein

**Query Motif:** 144QxxMxxFYxS153  
 Is this a LIG\_CRL4\_Cdt2\_1 motif?  
 Is this a LIG\_CRL4\_Cdt2\_2 motif?  
 Is this a LIG\_PCNA motif?  
 Is this a MOD\_phos\_NEK2 motif?

**Annotated switches**  
 This protein contains 3 annotated motif switches (show)

**13 Putative Switches** | 12 Experiments | 0 Warnings

**Putative Switches**

**Specificity switch** Overlapping/Adjacent motifs experimental  
 Validated LIG\_CRL4\_Cdt2\_1 (145-157) 3 shared residues  
 Validated LIG\_CYCLIN\_1 (155-159) Flanking  
 Validated LIG\_PCNA (144-153) 4 shared residues  
 46 predicted - toggle regular expression hits.

**Altered physicochemical compatibility**  
**Modification on motif**  
 153 (S) - Phosphoserine; by DYRK1B, Phosphoserine

**Modification in motif**  
 145 (T) - Phosphothreonine; by PKB\_group, Phosphothreonine; by PKC\_alpha, Phosphothreonine; by PIM-1, Phosphothreonine; by PKB\_group, Phosphothreonine; by PKB\_group, Phosphothreonine; by PKB\_group, Phosphothreonine; by PKA; PKB/AKT1, PIM1 and PIM2  
 146 (S) - Phosphoserine; by PKB\_group, Phosphoserine; by PKC\_alpha, Phosphoserine; by PIM-1, Phosphoserine; by PIM-1, Phosphoserine; by PKC 153 (S) - Phosphoserine; by DYRK1B, Phosphoserine

**Modification flanking motif**  
 130 (S@+14) - Phosphoserine; by MAPK\_group, Phosphoserine; by CDK2, Phosphoserine, Phosphoserine; by CDK2, Phosphoserine, Phosphoserine  
 141 (K@-3) - Acetylation, Ubiquitination  
 154 (K@0) - Acetylation, Ubiquitination  
 160 (S@+6) - Phosphoserine; by PKC; in vitro  
 161 (K@+7) - Acetylation, Ubiquitination  
 163 (K@+9) - Acetylation, Ubiquitination

**Rheostatic switch** Multiple modifications flanking motif  
 Modified at 130, 141, 145, 146, 153, 154, 160, 161, 163

**Pre-translational switch** In non-constitutive exon  
 None

**Auto-inhibitory switch** Possible intra-molecular binding  
 None

**Motif hiding**  
 Predicted LIG\_14-3-3\_2\_1 shared residues

**Architecture**

**Context**

Alignment | Motifs | Modification | Switches | Structure | Mutation | Isoforms | SNPs | Features | Disorder

offset: 1126 | Motif of interest: QxxMxxFYxS | 1164

CDKN1A\_HUMAN | toggle extra species

CDKN1A\_MOUSE | toggle extra species

switches: PCNA\_PIPBox, PCNA\_PIPBox

Modification switches: PIP-box K+4 motif, Nuclear localization signal

short sequence motif: PIP-box K+4 motif, Nuclear localization signal

ELM (+ toggle predicted hits)

ELM regex: UBX, APH1, GxGD, NDR, NDR, PCNA, NDR, NDR, SH3, WW\_Fin1, USP7, PKK, PCSK, PCNA, PCSK, PCSK, PCSK, phos\_NEK2, PCSK\_SK1, CYCLIN, Crl4\_Cdt2, CRL4\_Cdt2, ER, CRL4\_Cdt2, NLS\_MonoCo-CRL4\_Cdt2, phos, PKA, PKA, PKB, NLS\_Bipartite, NLS\_Bipartite, NLS\_Bipartite, NLS\_MonoEd

modified residue: ?

phospho.ELM: ?

phosphoSitePlus: ?

mutagenesis site: ?

secondary structure (+ details): ?

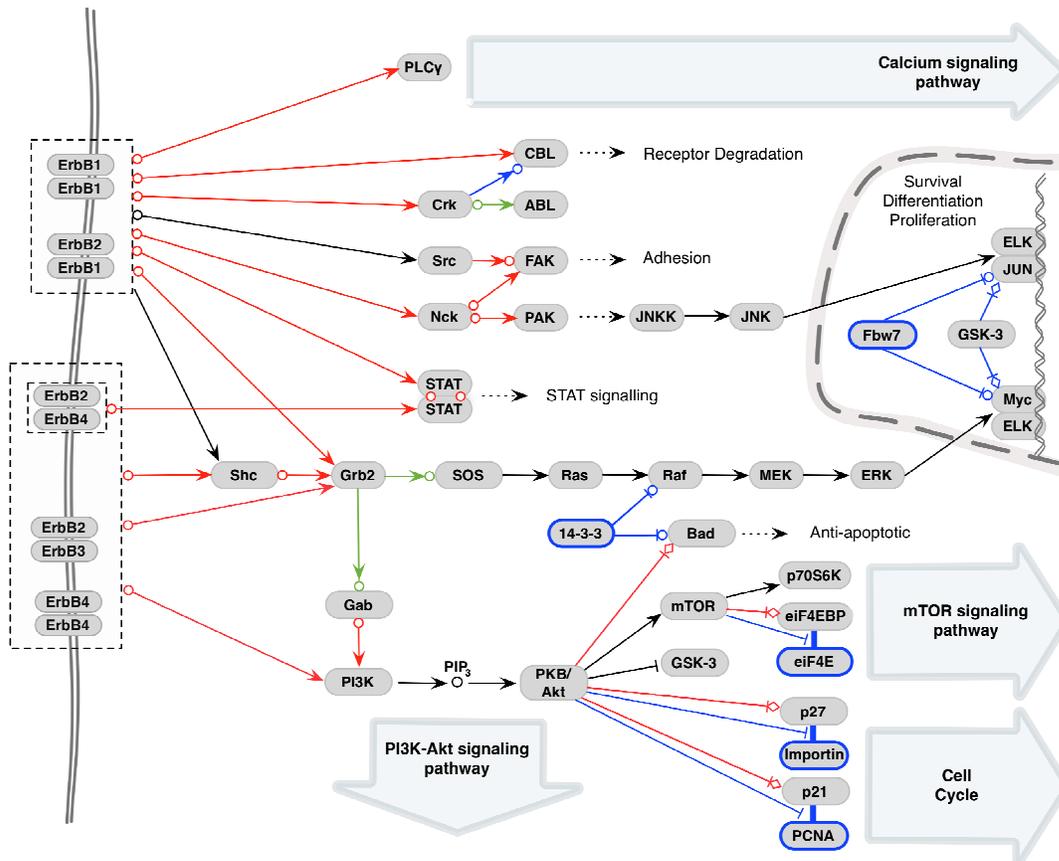
sequence variant: ?

chain: ? Cyclin-dependent kinase inhibitor 1

Powered by ProViz  
 hover over features for details

**Fig. S2. Visual output of the switches.ELM exploratory analysis tool.** (1) General information, including name and species of the query protein, sequence of the query motif, possible matches to known ELM motif classes, and curated switches involving this protein. (2) Select buttons to view either a list of putative switches for the user-submitted motif, links to publications providing experimental evidence (for curated switches involving this protein, mutagenesis data available for this protein region, and structural information for the protein linked to PDB), or warnings concerning the structural context of the motif. (3) List of putative switches regulating the user-submitted motif, including available experimental data for modification sites that might be involved. Hovering over the question

marks will show switch type definitions. Data extracted from external resources for the aligned protein region is identical to those shown in the visual output of a curated switch instance (fig. S1), except for the additional prediction of SLiMs based on the regular expressions for motif classes in ELM.



**Fig. S3. Switch instances in switches.ELM mapped to the ErbB signaling pathway annotated in KEGG.** Green connections denote motif-mediated interactions, while red connections indicate motif-mediated interactions regulated by a switching mechanism and curated in the switches.ELM resource. Blue connections and proteins are not annotated in the original pathway in KEGG. However, they are motif-mediated interactions that are regulated by a switching mechanism, are curated in the switches.ELM, and were inferred to be involved in the ErbB pathway.