Cell

Structural Complexity in Ubiquitin Recognition

J. Wade Harper[1,*] and Brenda A. Schulman[2,*]
[1]Department of Pathology, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA
[2]Howard Hughes Medical Institute, Departments of Structural Biology and Genetics and Tumor Cell Biology, St. Jude Children's Research Hospital, 332 North Lauderdale, Memphis, TN 38105, USA
*Contact: wade_harper@hms.harvard.edu (J.W.H.); brenda.schulman@stjude.org (B.A.S.)
DOI 10.1016/j.cell.2006.03.009

Ubiquitinated proteins are sorted into distinct pathways via association with several classes of ubiquitin binding domain-containing proteins. A virtual explosion in the field of ubiquitin binding proteins has revealed several new classes and interactions with distinct surfaces on ubiquitin, providing a clearer understanding of how sorting of ubiquitinated proteins is achieved.

Protein modification by ubiquitin and ubiquitin-like (Ubl) proteins has emerged as a central mechanism through which cellular pathways are regulated (Pickart, 2004). In this process, the C-terminal glycine of ubiquitin becomes linked to primarily lysine residues in target proteins via an E1-E2-E3 cascade. As ubiquitin itself contains seven lysine residues, multiple molecules of ubiquitin can also become linked to each other to form polyubiquitin chains. Thus, ubiquitination can take many forms which can differentially control the fate of the target protein. The best understood function of ubiquitination is proteolysis, whereby lysine-48 (K48)-linked polyubiquitin chains allow recognition by the 26S proteasome. However, proteins can also be monoubiquitinated or polyubiquitinated through alternative (e.g., K63) linkages, and such modifications are thought to control protein activity or localization (Hicke et al., 2005). The ubiquitination event can be viewed as a molecular zip code, which is used to sort different ubiquitination products to different destinations. Errors in delivery of ubiquitinated proteins to the proteasome or other destinations could be disastrous for cells.

Ubiquitin zip codes are read by several families of ubiquitin binding domain (UBD)-containing proteins that recognize a conserved "hydrophobic pocket" on ubiquitin centered at isoleucine-44 (I44; Hicke et al., 2005). Now, new work—including two papers in this issue of *Cell*—reveal additional molecular mail carriers that sort and process ubiquitinated proteins and provide evidence that additional surfaces of ubiquitin are recognized by different classes of UBDs (Bienko et al., 2005; Hirano et al., 2006; Hoeller et al., 2006; Lee et al., 2006; Mattera et al., 2006; Penengo et al., 2006; Reyes-Turcu et al., 2006). These studies further our understanding of how specificity is achieved.

Two-Faced UIMs: A New Subclass

One important question is whether the code for ubiquitin recognition will be as simple as one UBD = one mode of ubiquitin recognition. Studies of other types of protein-protein interaction motifs have revealed diversity in the ways a given class of protein interaction module can interact with partners, and the fact that distinct ubiquitin binding CUE domains bind ubiquitin in substantially different ways (Figure 1A; Hicke et al., 2005) indicates that such complexity may exist with UBDs. Indeed, recent work on by Hirano et al. (2006) on Hrs, a *u*biquitin-*i*nteraction *m*otif (UIM)-containing protein involved in receptor endocytosis (Hicke et al., 2005), suggests that this is the case. The structure of Hrs-UIM in complex with ubiquitin revealed a new subclass of UIMs. In the structure, the Hrs-UIM binds two ubiquitins, in stark contrast with previously characterized UIM domains which bind one (Figures 1A and 1B). In the Hrs-UIM complex, both ubiquitin molecules (Ub1 and Ub2) use their I44 surface to interact with UIM, but they bind on opposite sides of the helix. Importantly, the interaction surface for Ub1 and Ub2 are very similar, reflecting two conserved ubiquitin-interacting sequences within the single UIM. The two ubiquitin-interacting sequences are displaced by two residues.

Can one-sided and two-sided UIMs be distinguished on the basis of sequence? Hirano et al. (2006) aligned UIM motifs from several proteins and found that the two classes of UIM can be identified by the presence or absence of an internal repeat sequence e-x-e-x-Φ-x-Φ-A-Φ-A-z-S-z-A/S-e (where e is a negatively charged residue, x is any residue, z is a large hydrophobic or polar residue; underlined residues indicate binding site for Ub2; Φ, hydrophobic residue). All UIMs which maintain alanine at position 10 and either serine or alanine at position 14 are expected to form two-sided UIMs. This idea was substantiated via biochemical analysis of three two-sided UIMs (Hrs-UIM, Hsj1-UIM2, and Eps15-UIM2). Importantly, examination of EGF endocytosis in cells expressing mutations in either or both of the UBDs reveals that interaction of both sides of the UIM are required for full Hrs function.

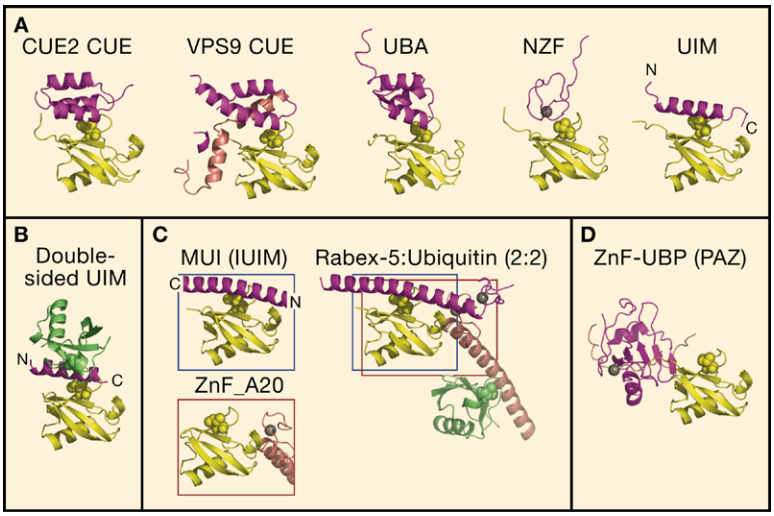

Figure 1. Structures of Various Ubiquitin Binding Domains, Bound to Ubiquitin

(A) Previously characterized ubiquitin binding domains (UBDs; magenta and pink) bound to the I44 hydrophobic patch on ubiquitin (Ub, yellow; I44 is space filled). PDB codes: 1OTR, CUE2; 1P3Q, VPS9 CUE; 1WR1, UBA; 1Q0W, UIM; 1Q5W, NZF. The gray sphere represents the Zn ion.

(B) The double-sided UIM bound to two Ub molecules (Ub1: yellow; Ub2: green). PDB code: 2D3G.

(C) Rabex-5 MIU/IUIM and ZnF_A20 motifs bound to Ub independently and together. Also shown is how a second Ub molecule (green) can bind nearby. PDB code: 2C7N.

(D) The ZnF-UBP domain bound to Ub. PDB code: 2G45.

What is the purpose of two-sided UIMs? One possibility is that having a two-sided UIM provides additional avidity in recognition of targets, particularly those targets that contain either a polyubiquitin chain or contain multiple monoubiquitins. In principle, this could be achieved by having two UIM domains in the same polypeptide. Indeed, many UIM-containing proteins contain tandem UIM domains. However, it is also the case that some proteins that contain two-sided UIMs also have tandem UIMs, sometimes of the one-sided class and sometimes of the two-sided class (Hirano et al., 2006). An alternative possibility is that two-sided and tandem UIMs have evolved to bind ubiquitinated proteins which present multiple ubiquitin molecules in a particular stereotyped architecture with respect to each other. Further studies are required to distinguish these possibilities and to establish any functional differences between one- and two-sided UIMs.

Rabex-5 Structure Reveals Two New UBDs

Two other new UBDs have been identified from studies of the Rabex-5 GTPase exchange factor, the mammalian ortholog of yeast VPS9 (Lee et al., 2006; Mattera et al., 2006; Penengo et al., 2006). A C-terminal CUE-UBD (Hicke et al., 2005) allows recruitment of VPS9 to ubiquitinated cell surface receptors at the plasma membrane or the endosome, facilitating Rab activation. Upon EGF stimulation, the EGF receptor (EGFR) gets ubiquitinated, which recruits Rabex-5 to the plasma membrane (Penengo et al., 2006). However, it was not clear how Rabex-5 is recruited to ubiquitin because Rabex-5 lacks strong sequence identity with VPS9 in its C terminus. Recent work revealed the presence of two new functionally distinct and independent UBDs in the N terminus of Rabex-5: (1) a linear peptide sequence that resembles a UIM (described below), and (2) the ZnF_A20 motif (Lee et al., 2006; Penengo et al., 2006). The first motif has been named "MIU" (*m*otif *i*nteracting with *u*biquitin) or "IUIM" (*i*nverted *UIM*). Both the MIU/IUIM and the ZnF_A20 sequences, which are capable of binding both monomeric ubiquitin and K48-linked chains, are conserved in Rabex-5 orthologs as well as in proteins with unrelated functions.

The Rabex-5 MIU/IUIM interacts with ubiquitin's I44 hydrophobic patch (Figure 1C). However, the sequence orientation of the MIU/IUIM is inverted with respect to UIM, such that the ubiquitin C terminus is directed toward the C terminus of the MIU/IUIM domain (Figures 1A and 1C). Consistent with structural data, mutation of the MIU/IUIM's central alanine or ubiquitin's I44 dramatically reduces binding (Lee et al., 2006; Penengo et al., 2006). The MIU/IUIM's N terminus contains an additional helical turn, providing additional contacts with ubiquitin that increase its affinity for ubiquitin by ~10-fold relative to most UIMs.

The ZnF_A20 motif represents another new UBD, distinct from two other known zinc binding UBDs, the NZF and ZnF-UBP motifs (Hicke et al., 2005; Figure 1). The ZnF_A20 motif interacts with the D58-centered hydrophobic face of ubiquitin and represents the first UBD interacting with a surface of ubiquitin other than the I44 patch, a finding supported by mutational data (Lee et al., 2006; Penengo et al., 2006).

UBD-containing proteins often have multiple UBDs of different classes, raising the question of how these different UBDs function together or independently to recognize ubiquitinated targets. The Rabex-5 studies provide the first structural insight into combinatorial recognition of ubiquitin by multiple UBDs in a single protein. Although the structures presented contain one molecule of Rabex-5 per ubiquitin, this apparent 1:1 stoichiometry appears to be a result of crystal packing in which higher-order interactions are observed. Indeed, both studies present biophysical data indicating that both the MIU/IUIM and ZnF_A20 domains bind ubiquitin independently (Lee et al., 2006; Penengo et al., 2006). However, Penengo et al. (2006) provide evidence that two molecules of Rabex-5 can interact with a single ubiqui-