Crystal Structure of a PFU-PUL Domain Pair of *Saccharomyces Cerevisiae* Doa1/Ufd3

RIEKO NISHIMASU[1, 2], HIROFUMI KOMORI[3, 4], YOSHIKI HIGUCHI[3, 4], HIROSHI NISHIMASU[5,*] and HIDEKAZU HIROAKI[2,*]

[1]*Division of Molecular Pharmacology and Pharmacogenomics, Department of Biochemistry and Molecular Biology, Kobe University Graduate School of Medicine, Kobe, Japan;*

[2]*Division of Structural Biology, Department of Biochemistry and Molecular Biology, Graduate School of Medicine, Kobe University Graduate School of Medicine, Kobe, Japan;*

[3]*Department of Life Science, Graduate School of Life Science, University of Hyogo and Himeji Institute of Technology, Hyogo, Japan;*

[4]*RIKEN SPring-8 Center, Sayo-gun, Hyogo, Japan;*

[5]*Division of Structure Biology, Department of Basic Medical Science, Institute of Medical Science, The University of Tokyo, Tokyo, Japan;*

Doa1/Ufd3 is involved in ubiquitin (Ub)-dependent cellular processes in *Saccharomyces cerevisiae*, and consists of WD40, PFU, and PUL domains. Previous studies showed that the PFU and PUL domains interact with Ub and Hse1, and Cdc48, respectively. However, their detailed functional interactions with Doa1 remained elusive. We report the crystal structure of the PFU-PUL domain pair of yeast Doa1 at 1.9 Å resolution. The conserved surface of the PFU domain may be involved in binding to Ub and Hse1. Unexpectedly, the PUL domain consists of an Armadillo (ARM)-like repeat structure. The positively charged concave surface of the PUL domain may bind to the negatively charged C-terminal region of Cdc48. A structural comparison of Doa1 with Ufd2 revealed that they share a similar ARM-like repeat, supporting a model in which Doa1 and Ufd2 compete for Cdc48 binding and may dictate the fate of ubiquitinated proteins in the proteasome pathway.

Doa1 (degradation of alpha2 1, also known as Ufd3 for ubiquitin fusion degradation 3) is implicated in various ubiquitin (Ub)-dependent cellular processes in *Saccharomyces cerevisiae*. Doa1 was originally identified in a mutant deficient in degrading the short-lived *MAT*a2 transcriptional repressor (12). Subsequently, Doa1 was shown to affect Ub-dependent proteolysis (9, 14) and resistance to volatile anesthetics (15). The *doa1 Δ* cells have a reduced amount of free Ub, which is partially complemented by overexpression of Ub (9, 14). Doa1 interacts with Cdc48 (mammalian p97/VCP), an AAA-ATPase that acts as a molecular chaperone for ubiquitinated protein substrates. Cdc48 is involved in several Ub-mediated pathways, such as membrane fusion and proteasomal degradation (9, 22). Cdc48 also interacts with the E4 multiubiquitinating enzyme Ufd2, which extends the Ub chains of ubiquitinated substrates to promote degradation. Doa1 and Ufd2 bind to Cdc48 in a mutually exclusive manner, whereas Doa1 and the deubiquitinating enzyme Otu1 can

simultaneously bind to Cdc48 (28). In addition, the *doa1* Δ phenotypes are reminiscent of defects observed in mutants of deubiquitinating enzymes (19, 32). These observations suggested a model in which the balance between Doa1, Ufd2, and Otu1 may determine whether a ubiquitinated substrate is further multiubiquitinated for degradation or deubiquitinated and/or released for other purposes (28).

The increased sensitivity of *doa1* Δ cells to DNA damaging agents suggested that Doa1 is also involved in the DNA damage response (11). Recent work implicated Doa1 in recycling Ub from the proteasome pathway into the DNA damage response pathways, since neither PCNA (proliferating cell nuclear antigen) nor histone H2B underwent damage-induced ubiquitination in the *doa1* Δ cells, and Lys48-linked trimeric Ub accumulated in the *doa1* Δ cells (20). Moreover, Doa1 plays a role in sorting ubiquitinated membrane proteins into multivesicular bodies (MVBs) (26); GFP-Ub is accumulated within the vacuoles of *doa1* Δ cells. In this context, Doa1 binds to a Vps27-Hse1 (yeast orthologues of mammalian Hrs-STAM) complex, which serves as a Ub-sorting receptor at the surface of early endosomes.

Doa1 homologues are widely distributed among eukaryotes. The *Schizosaccharomyces pombe* homologue Lub1 (low ubiquitin content 1) participates in Ub homeostasis and stress responses associated with Ub-dependent proteolysis (24). As observed in *S. cerevisiae*, the *lub1* Δ cells contain low levels of Ub and display hypersensitivity to both UV and temperature. A *Candida albicans* homologue of Doa1 is also associated with Ub-dependent proteolysis and is involved in multiple cellular processes (17). PLAA (phospholipase A2 activating protein) may be a mammalian homologue of Doa1 (2), although it is unknown whether PLAA participates in Ub-mediated pathways. Doa1 complements *lub1* Δ phenotypes (24), and a Doa1-PLAA chimeric protein binds to Ub and complements *doa1* Δ phenotypes (22), indicating that Doa1 is functionally conserved among eukaryotes.

Doa1 is a 715-residue protein consisting of three domains: an N-terminal WD40 domain, a central PFU (PLAA family ubiquitin binding) domain, and a C-terminal PUL (PLAA, Ufd3, and Lub1) domain (Figure 1A). While the WD40 domain is homologous to the WD40 proteins present in many eukaryotes (23), the PFU and PUL domains are unique to Doa1 homologues. Although WD proteins comprise a seven-bladed β-propeller repeat structure and mediate protein-protein interactions, the structure and function of the Doa1 WD40 domain have remained unclear. Rumpf *et al.* reported that the WD40 and PFU domains are both required for Ub binding (28), whereas other groups found that the PFU domain alone is sufficient for Ub binding (22, 26, 29). In addition, the PFU domain binds to the SH3 domain of Hse1 in the MVB pathway (26). The PUL domain binds to the C-terminal region of Cdc48 in the proteasome pathway (9, 22, 28, 37). However, the mechanism by which Doa1 achieves its biological functions by interacting with its partner proteins has not been identified, partially because of the lack of structural information.

Here, we present the crystal structure of the PFU-PUL domain pair of yeast Doa1. The structure revealed that the PFU domain adopts a novel ββββαα fold, whereas the PUL domain forms an unexpected Armadillo (ARM)-like repeat structure. Our findings provide structural insights into the mechanism by which Doa1 participates in Ub-mediated pathways through interactions with various proteins.

MATERIALS AND METHODS

Sample preparation

The gene encoding the region encompassing the PFU and PUL domains of *S. cerevisiae* Doa1 (residues 354-715) was amplified by PCR from the *S. cerevisiae* genome, and was

cloned into the *Nde*I-*Eco*RI sites of the pET-28 vector (Novagen). The inserted sequence was verified by DNA sequencing. The N-terminal His-tagged PFU-PUL protein was overexpressed in *E. coli* C43(DE3) (Novagen). The transformed cells were grown at 37°C in 2.5 L LB medium containing 0.1 g/L ampicillin, induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 4 h, and then harvested by centrifugation. The cell pellet was resuspended in lysis buffer, consisting of 50 mM Na-phosphate (pH 7.6), 0.3 M NaCl, 10 mM imidazole, 0.1 mM DTT, and 0.1 mM phenylmethylsulfonyl fluoride (PMSF), lysed by sonication, and clarified by centrifugation. The resultant supernatant was purified by chromatography on Ni-NTA (Qiagen), HiLoad 16/60 Superdex 200 (GE Healthcare), and Q Sepharose (GE Healthcare) columns. The purified protein was dialyzed against 10 mM Tris-HCl (pH 8.0) and concentrated to 30 mg/mL. The selenomethionine (SeMet)-labeled protein was expressed in *E. coli* B834(DE3) (Novagen) and was purified using the same protocol as for the native protein.

Crystallization

Crystallization was performed at 20°C, using the hanging drop vapor diffusion method. Crystals of the SeMet-labeled PFU-PUL protein were grown by mixing 1 ml of the protein solution (16 mg/mL) and 1 ml of the reservoir solution [0.1 M Tris-HCl (pH 7.5), 28% PEG3350, and 0.1 M ammonium sulfate].

Data collection and structure determination

X-ray diffraction data were collected at 100 K on beamline BL41-XU at SPring-8 (Hyogo, Japan). Diffraction data were processed using the HKL2000 program (HKL Research, Inc.). The structure was determined by the MAD method, using the 1.9 Å resolution data from a SeMet-labeled crystal. All four selenium atoms were located, and initial phases were calculated using SHARP/autoSHARP (35), followed by automated model building using ARP/wARP (18). The model was further manually built using COOT (7) and was refined using Refmac (3). Structural figures were prepared using PyMol (6).

RESULTS

Crystallography

We expressed the region encompassing the PFU and PUL domains (residues 354-715) of *S. cerevisiae* Doa1 in *Escherichia coli* and purified it to homogeneity. Neither the full-length nor the WD40 domain of Doa1 was expressed as a soluble protein in *E. coli* (data not shown). The crystal structure of the PFU-PUL domain pair was solved by the multiwavelength anomalous dispersion (MAD) method using a SeMet-labeled protein, and was refined at 1.9 Å resolution to an R_{work} of 19.6% (R_{free} of 21.9%) with good stereochemistry (Table I). The crystal belongs to the space group $P6_1$, with one molecule in the asymmetric unit. No electron density was observed for the N-terminal region of the PFU domain (residues 354-375) and the region between the PFU and PUL domains (residues 440-463), probably because these regions are flexible and disordered in the crystal structure. The refined model comprises residues 376-439 and 464-715, and 214 water molecules.

Overall structure

The structure revealed that the PFU domain (residues 376-439) consists of two antiparallel β-sheets (β1-β2 and β3-β4), followed by two α-helices (α1-α2) and one 3_{10}-helix (η1), whereas the PUL domain (residues 464-715) comprises 16 α-helices (α3-α18) and three 3_{10}-helices (η2-η4) (Figure 1B). Consistent with our gel-filtration results (data not shown), the PFU-PUL domain pair exists as a monomer in the crystal, suggesting that Doa1 may function as a monomer in solution. Since the region between the two domains

Table I. Data collection and refinement statistics

Dataset	SeMet		
	Peak	Edge	Remote
Data collection			
Wavelength, Å	0.9789	0.9792	0.964
Space group	$P6_1$		
Cell dimensions			
a, b, c, Å	103.1, 103.1, 72.1		
Resolution, Å	50.-1.9 (1.97-1.90)	50-2.0 (2.07-2.00)	50-2.1 (2.18-2.10)
No. of unique reflections	34,591 (3,476)	29,724 (2,961)	25,774 (2,555)
Redundancy	11.2 (11.2)	11.2 (11.2)	11.1 (11.2)
Completeness, %	100 (100)	100 (100)	100 (100)
$I/\sigma I$	49.7 (8.9)	47.5 (8.0)	46.3 (7.3)
R_{sym}, %	6.2 (25.6)	5.6 (31.1)	5.7 (36.3)
Refinement			
Resolution, Å	50-1.9 (1.95-1.90)		
No. of reflections	32,712 (2,422)		
R_{work}, %	19.5		
R_{free}, %	21.9		
No. of atoms			
Protein	2,543		
Water	214		
Average B-factor, Å^2			
Protein	26.2		
Water	34.0		
Rmsd from ideal values			
Bond lengths, Å	0.009		
Bond angles, °	1.1		
Ramachandran plot, %			
Most favored	95.1		
Allowed	4.9		
Disallowed	0		
PDB ID code	3L3F		

Values in parentheses are for the highest resolution shell.

(residues 440-463) is disordered, it is difficult to identify the spatial arrangement between the PFU and PUL domains unambiguously. One PUL domain contacts one PFU domain in the asymmetric unit, and four PFU domains are related by crystallographic symmetry. Thus, five arrangement patterns between the PFU and PUL domains are possible. The buried surface areas between the two domains in these five arrangements range between 121 and 968 Å^2. The domain arrangement shown in Figure 1B seems the most likely, based on the largest buried surface area (968 Å^2) and the relatively low B-factors for residues at the interface. In this arrangement, the α6-α7 and α8-α9 loop regions of the PUL domain interact with the β1-β2 and β3-β4 loop regions and helix α1 of the PFU domain (Figure 1C). Tyr549 on the α6-α7 loop interacts with Glu394, Lys397, Pro398, and Leu400 on the β3-β4 loop. Glu583 and Asn584 on the α8-α9 loop hydrogen bond with Glu381, and Lys401 and Tyr421, respectively. The main-chain carbonyl group of Lys546 on helix α6 hydrogen bonds with the

side chain of Arg420 on helix α1. However, it is unknown whether the spatial arrangement between the PFU and PUL domains is fixed in solution, because most of the residues at the domain interface are not conserved between yeast Doa1 and human PLAA (Figure 2). Thus, we speculate that the region connecting the PFU and PUL domains is intrinsically flexible, so that the two domains could function independently.
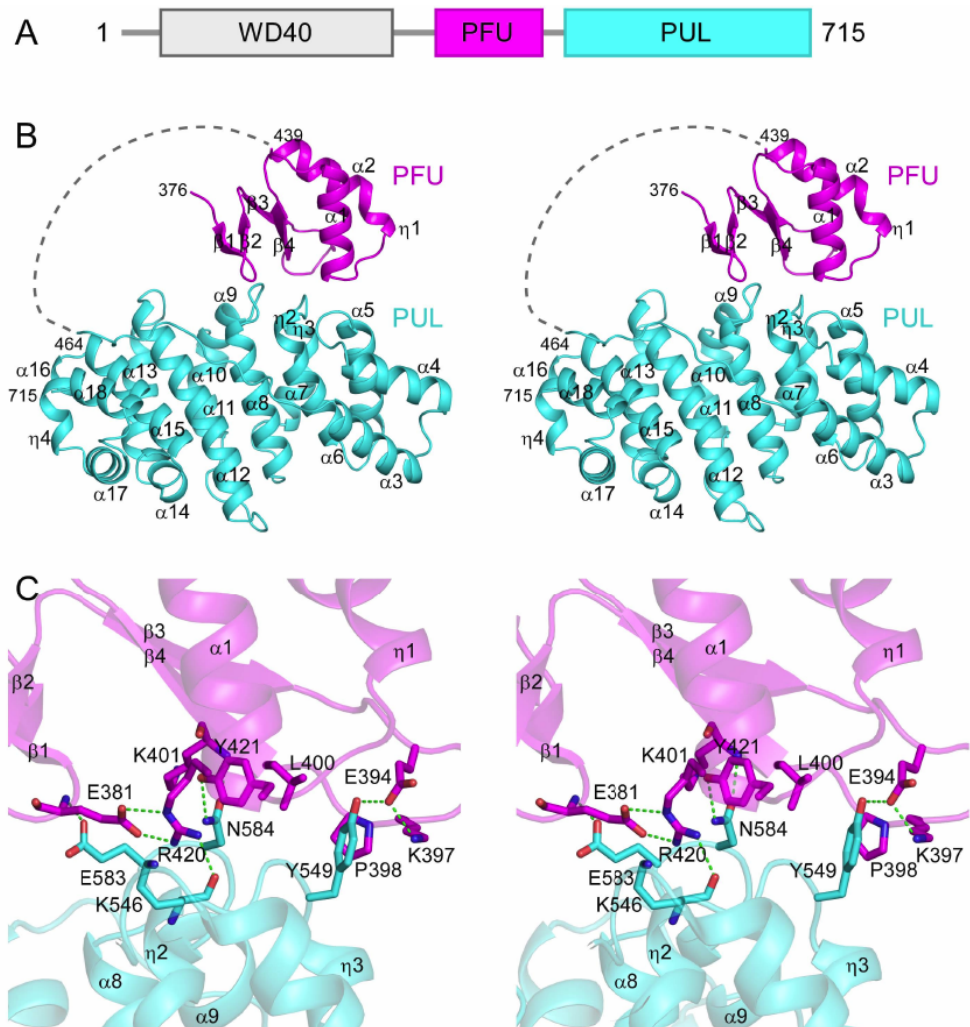

Figure 1. Overall structure.

(A) Domain structure of Doa1. The WD40, PFU and PUL domains are colored gray, magenta and cyan, respectively. (B) Overall structure of Doa1 (stereo view). The PFU and PUL domains are colored magenta and cyan, respectively. The disordered region between the two domains is shown as a dashed line. (C) Interaction between the PFU the PUL domains(stereo view).Residues at the domain interface are shown in stick representations. Hydrogen bonds are shown as dashedlines.
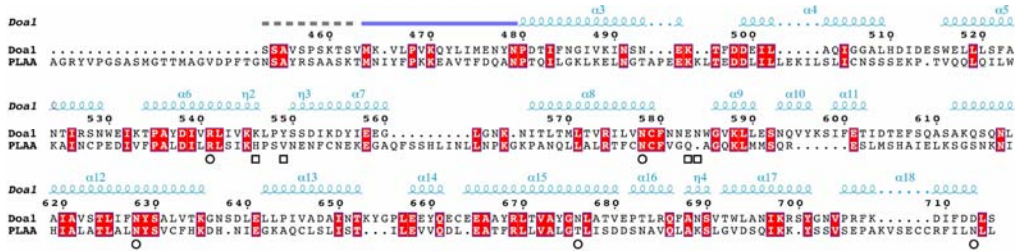
Figure 2. Structure-based sequence alignment of yeast Doa1 and human PLAA.
Conserved residues are highlighted with a red background. The secondary structure of Doa1, colored as in Fig. 1, is shown above the alignment. The disordered region in the crystal structure is indicated by a dashed line. The figure was prepared using 3D-Coffee (25) and ESPript (10).

The PFU domain

The PFU domain consists of a βββββαα fold. An NMR solution structure of the PFU domain of human PLAA was reported quite recently (8). The PFU domain of human PLAA (PDB ID code 2K8B) also consists of a βββββαα fold and shares structural similarity with that of yeast Doa1 (RMSD for 63 Cα atoms of 1.8 Å) (Figure 3A), indicating that the PFU domains of Doa1 homologues are structurally conserved. Conserved hydrophobic residues (Val391, Ile393, Leu402, Phe417, Val431, and Phe434) form a hydrophobic core and stabilize the architecture of the PFU domain (Figure 3A). These conserved residues also contribute to the formation of a hydrophobic pocket on the molecular surface of the PFU domain (Figure 3B). A Dali search (13) revealed that the structure of the PFU domain is partially related to that of Ub. For example, residues 12-64 of the Doa1 PFU domain superimposed on residues 1-62 of human Ub (PDB ID code 1SIF) with an RMSD for 47 Cα atoms of 2.2 Å. The β3, β4, α1, and η1 of the Doa1 PFU correspond to the β1, β2, α1, and η1 of human Ub, respectively. However, unlike Doa1, Ub consists of a ββααββαβ fold, showing that the PFU domain adopts the novel fold characteristic of the Doa1 homologues.

The PUL domain

The PUL domain consists of six ARM-like repeats comprising different numbers of α-helices (Figure 4A). The α-helical repeat structure was unanticipated, because the PUL
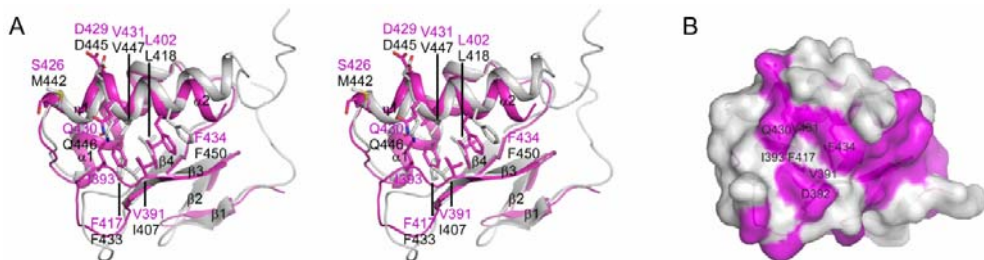

Figure 3. Structure of the PFU domain.
(A) Superimposition of the Doa1 PFU domain and the PLAA PFU domain (PDB ID code 2K8B) (stereo view). The PFU domains of Doa1 and PLAA are colored magenta and gray, respectively. The secondary structures of Doa1 are labeled. (B) Surface representation of the PFU domain of Doa1. Conserved residues between Doa1 and PLAA are colored magenta.
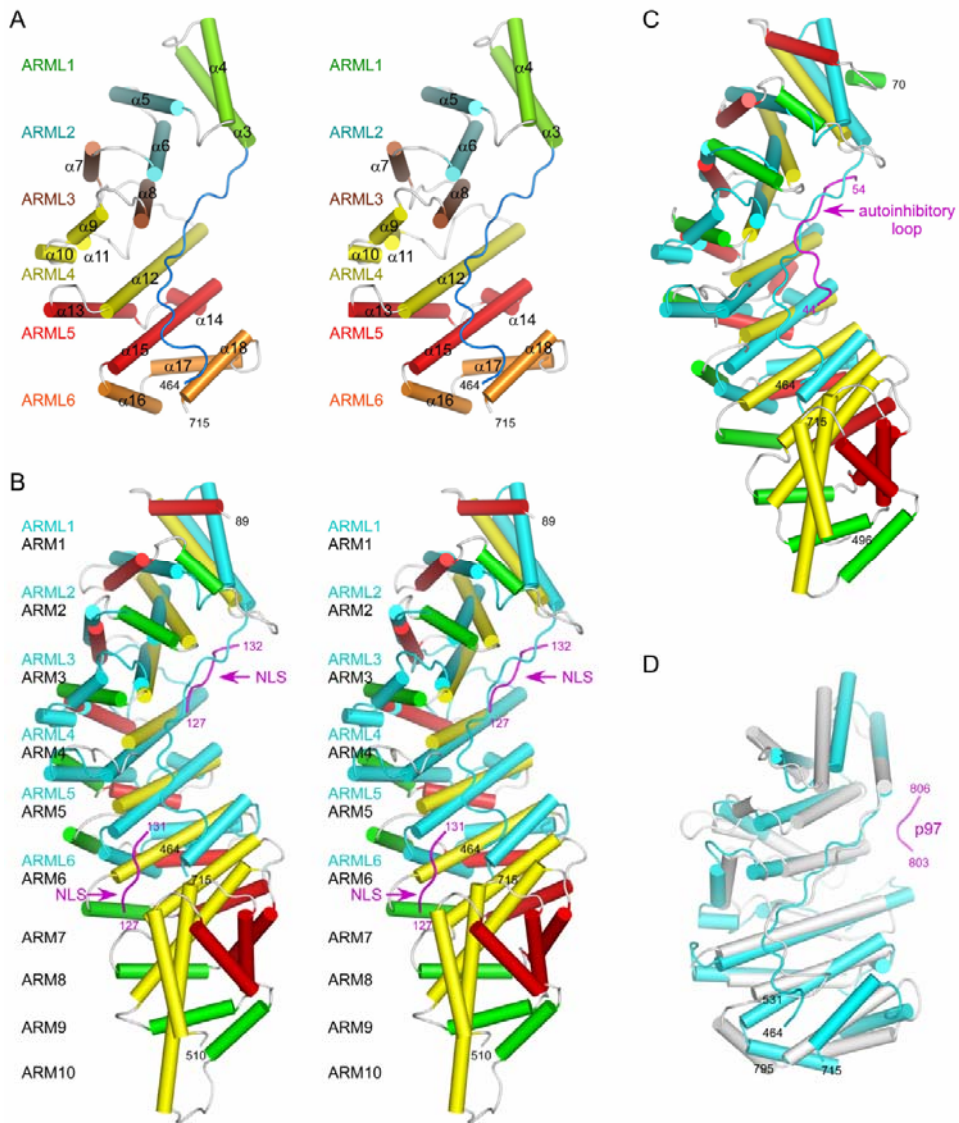
Figure 4. Structure of the PUL domain.
(A) Structure of the PUL domain of Doa1 (stereo view). ARM-like repeats are shown in different colors. The N-terminal loop is colored blue. (B) Superimposition of the Doa1 PUL domain and yeast Kapa in complex with the SV40 NLS peptide (PDB ID code 1BK6) (stereo view). The PUL domain of Doa1 is colored cyan. Helices H1, H2, and H3 of Kapα are colored green, red, and yellow, respectively. The NLS peptides are colored magenta. (C) Superimposition of the Doa1 PUL domain and mouse importin a with the internal autoinhibitory loop. The PUL domain of Doa1 is colored cyan. Helices H1, H2, and H3 of importin a are colored green, red, and yellow, respectively. The internal autoinhibitory loop is colored magenta. (D) Superimposition of the Doa1 PUL domain and the PLAA PUL domain (PDB ID code 3EBB). The PUL domains of Doa1 and PLAA are colored cyan and gray, respectively. The p97 peptide bound to PLAA is colored magenta.

domain does not share detectable sequence homology with known α-helical repeat proteins. The ARM-like repeats 1-3, 4, and 5-6 comprise two, four, and three α-helices, respectively. These α-helices pack together through hydrophobic interactions, forming a crescent-shaped extended structure. Helices α3, α6, α8, α12, α15, and α18 define the concave surface, whereas helices α4-5, α7, α9-11, α13-14, and α16-17 define the convex surface.

A Dali search (13) revealed that the PUL domain shares structural similarity with karyopherin α (Kapα, also known as importin α), which contains ten ARM repeats and mediates the nuclear import of proteins containing nuclear localization signals (NLSs) (4,31). The canonical ARM repeat comprises three helices (H1, H2, and H3) (Figure 4B). The H2 and H3 helices pack together in an antiparallel manner, with the H1 helix roughly perpendicular to the H2 and H3 helices. The ten ARM repeats of Kapα stack through hydrophobic interactions, forming an elongated right-handed superhelix of helices. Residues 481-711 of the PUL domain superimposed on residues 103-319 of yeast Kapα bound to the NLS of the simian virus 40 large T-antigen (SV40 NLS, 127-KKKRKV-132) (PDB ID code 1BK6) (4) with a Z-score of 12.2 and an RMSD for 188 Cα atoms of 4.0 Å (Figure 4B). Residues 469-711 of the PUL domain also superimposed on residues 44-313 of mouse importin α (PDB ID code 1IAL) (16) with a Z-score of 11.2 and an RMSD for 199 Cα atoms of 3.8 Å (Figure 4C). Although the repeats of Doa1 comprise different numbers of α-helices, the ARM-like repeats 1-6 of Doa1 roughly superimposed on the ARM repeats 1-6 of Kapα. In particular, helices α16, α17, and α18 of Doa1 ARM-like repeat 6 superimposed well on helices H1, H2, and H3 of Kapα ARM repeat 6, respectively. However, there are notable differences in the canonical ARM repeats of Kapα and the ARM-like repeats of Doa1. The ARM repeats consist of about 40 residues that contain a consensus sequence motif, resulting in a regular α-helical repeat structure comprising three α-helices. The ARM repeats of yeast Kapα share a very similar structure and superimposed well on each other, with an RMSD for aligned Cα atoms of 0.9 Å on average (4). In contrast, the ARM-like repeats of Doa1 consist of helices of various lengths that lack an obvious consensus sequence motif, resulting in an irregular α-helical repeat structure comprising 2-4 α-helices. These observations suggest that the canonical ARM repeats of Kapα and the ARM-like repeats of Doa1 originated from different ancestors, but acquired a similar α-helical repeat structure through convergent evolution.

The PUL domain of human PLAA complexed with a peptide consisting of the C-terminal four residues of p97 (PDB ID code 3EBB) has been deposited in the Protein Data Bank. The PUL domain of human PLAA also consists of six ARM-like repeats, comprising 17 α-helices (Figure 4D). Although the PUL domain of yeast Doa1 shares only 22% sequence similarity with that of human PLAA (Figure 2), the two structures superimposed well, with a Z-score of 21.9 and an RMSD for 236 Cα atoms of 3.2 Å (Figure 4D), suggesting a conserved function of Doa1 homologues among eukaryotes.

The N-terminal loop of the PUL domain

The N-terminal 15 residues of the PUL domain (residues 464-479) are well ordered in the electron density and interact with the concave surface of the PUL domain (Figure 5A). Several residues on the concave surface hydrogen bond to the backbone of the N-terminal loop. Arg541 on helix α6 hydrogen bonds to the backbone carbonyl group of Asn477. Asn578, Asn628, and Asn676 on helices α8, α12, and α15 form bidentate hydrogen bonds with the backbone amide and carbonyl groups of Met475, Leu473, and Val469, respectively. Asp713 on helix α18 hydrogen bonds to the backbone amide groups of Val466 and Leu467. In addition, the side chains of Tyr478, Glu476, Met475, Tyr472, Lys470, Val469, and
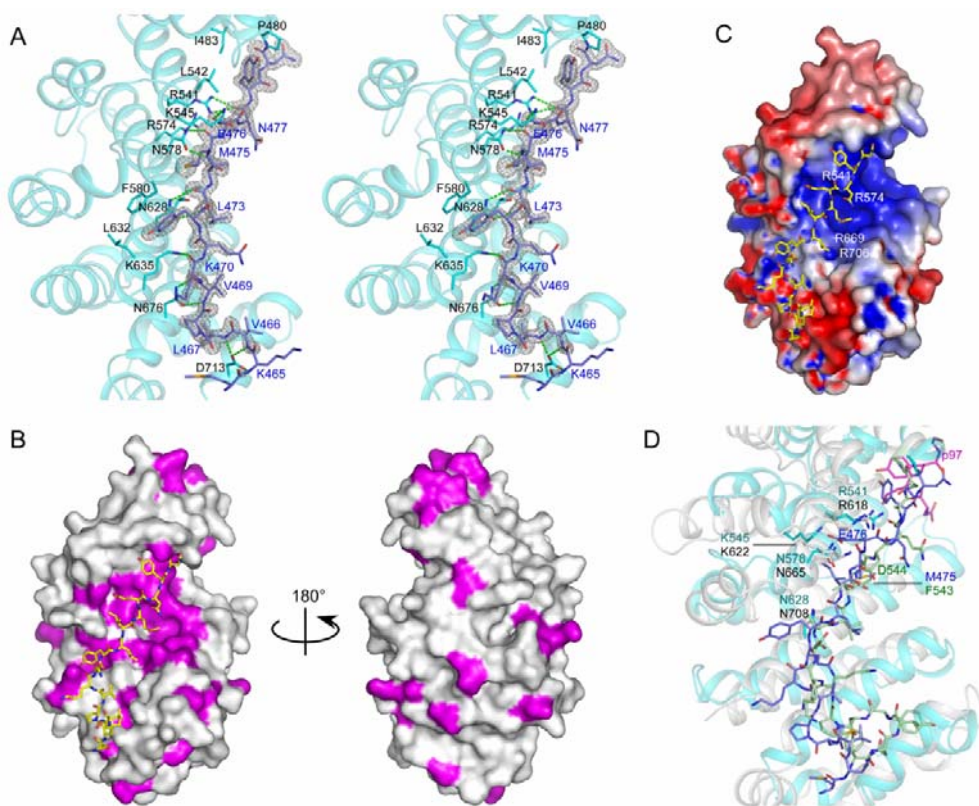
Figure 5. Interaction between the N-terminal loop and the concave surface of the PUL domain.
(A) The F_O - F_C omit map for the N-terminal loop contoured at 4 σ is shown as a gray mesh. The N-terminal loop and the ARM-like repeats of the PUL domain are colored blue and cyan, respectively. Hydrogen bonds are shown as dashed lines. (B) Surface representation of the concave (Left) and convex (Right) surfaces of the Doa1 PUL domain. The N-terminal loop is shown in stick representations with the carbon atoms colored yellow. Conserved residues between Doa1 and PLAA are colored magenta. (C) The electrostatic surface potential of the PUL domain, calculated using APBS (1). Positive and negative surface potentials are colored blue and red, respectively. The N-terminal loop is shown in stick representations with the carbon atoms colored yellow. (D) Superimposition of the PUL domains of Doa1 and PLAA. The concave surfaces of Doa1 and PLAA are colored cyan and gray, respectively. The N-terminal loops of Doa1 and PLAA are shown in stick representations with carbon atoms colored blue and green, respectively. The peptide consisting of the C-terminal four residues of p97 bound to PLAA is colored magenta.

Leu467 on the N-terminal loop interact with the concave surface of the PUL domain. The average B-factor for the loop region (27 Å^2) is comparable to that for the rest of the protein (26 Å^2), suggesting that these interactions are not simply crystal-packing interactions. While most of the residues on the N-terminal loop are not conserved, many conserved residues (Pro480, Ile483, Arg541, Leu542, Lys545, Arg574, Asn578, Phe580, Ile620, Ala621, Thr624, Asn628, and Lys635) are located on the concave surface (Figures 2 and 5B) In addition, this surface is basic, due to the presence of conserved positively charged residues (Arg541, Arg574, Arg669, and Arg706) (Figure 5C). These observations suggest the
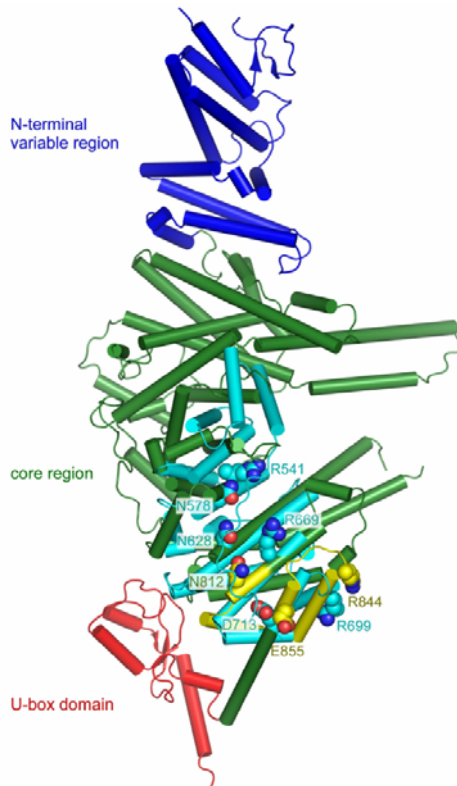
Figure 6. Structural similarity between Doa1 and Ufd2. The PUL domain of Doa1 and the core region of Ufd2 are superimposed. The PUL domain of Doa1 is colored cyan. The N-terminal variable region, core region, and U-box domain of Ufd2 are colored blue, green and red, respectively. The region that may be important for Cdc48 binding is colored yellow. Residues that could be involved in binding to the C-terminal region of Cdc48 are shown in space-filling representations.

functional importance of the interactions between the N-terminal loop and the concave surface of the PUL domain.

In the crystal structure of the human PLAA (PDB ID code 3EBB), the corresponding loop region (residues 464-479) also interacts with the concave surface of the PUL domain (Figure 5D). Notably, Arg618, Asn665, and Asn708 of PLAA, which correspond to Arg541, Asn578, and Asn628 of Doa1, hydrogen bond to the backbone of the N-terminal loop. Asp544 (Glu476 of Doa1) forms a salt bridge with Lys622 (Lys545 of Doa1). Phe543 (Met475 of Doa1) interacts with Arg661, Cys664, Ala701, and Thr704 (Arg574, Val577, Ala621, and Thr624 of Doa1). These observations support the notion that the interactions between the N-terminal loop and the concave surface of the PUL domain are physiologically relevant.

DISCUSSION

We have solved the crystal structure of the PFU-PUL domain pair of yeast Doa1. The structure revealed that the PFU domain forms a unique βββββαα fold and the PUL domain has an unexpected, irregular ARM-like repeat structure. The structure also showed the interactions between the N-terminal loop and the concave surface of the PUL domain. The present structure provides possible structural explanations for previous biochemical and genetic data showing that Doa1 interacts with various binding partners and participates in Ub-mediated cellular processes. Our present structure is structurally significant in the following points; (i) this is the first crystal structure of the PFU-PUL domain pair with a flexible linker, in which PFU-PUL interaction was observed; (ii) this is the first crystal structure of the PFU domain in high resolution; and (iii) the autoinhibition mechanism of the

loop bound to the PUL domain was reproduced in the different crystal form of the Doa1-PUL domain alone (36).

Biochemical studies showed that the PFU domain of Doa1 is involved in Ub binding (22, 26, 28) and that the F417D/F434D double mutant, in which the conserved Phe417 and Phe434 are replaced by Asp, cannot bind to Ub (22). In the present structure, Phe417 contributes to stabilize the architecture of the domain, whereas Phe434 participates in forming a putative Ub-binding hydrophobic pocket (Figure 3). It is also possible that the mutation of Phe417 to Asp resulted in the structural perturbation. Fu *et al.* reported that the PFU domain (residues 386-465) of human PLAA weakly binds to Ub, through Met442 and Asp445 on helix α2 of the PFU domain interacting to Ile44 and Arg42 on the Ile44−centered patch of Ub, respectively (8). However, it should be noted that the affinity between the wild-type PLAA PFU domain and Ub is very low; the K_d value of the wild-type for Ub is 1.8 mM. Moreover, Met442 and Asp445 in human PLAA are not conserved in yeast Doa1 (Figure 3A), thereby suggesting the yeast Doa1-PFU domain is not sufficient for Ub binding, while the WD-PFU domain pair is necessary. This notion is consistent with the previous biochemical data showing that the fragment encompassing both of the WD and PFU domains (residues 1-494) binds to Ub, but neither the fragment containing only the WD domain (residues 1-288) nor the fragment lacking the WD domain (residues 288-715) could (28). However, Mullally *et al.* and Ren *et al.* reported that the fragments encompassing only the PFU domain (residues 350-450 and residues 301-465) are sufficient for Ub binding, respectively (22, 26). These discrepancies may stem from differences in their experimental conditions.

Yeast two-hybrid screens suggested that Doa1 could interact with the SH3 domain of Hse1 (33). In addition, a recent study showed that Doa1 directly interacts with Hse1 and plays a role in helping to process ubiquitinated membrane proteins for sorting into MVBs (26). GST pull-down experiments revealed that the conserved region (residues 433-445) within the Doa1 PFU domain is critical for binding to the Hse1 SH3 domain. This was confirmed with a mutant Hse1 containing an Ala replacement of a Trp residue on the predicted hydrophobic surface of the SH3 domain, which could not bind Doa1 (26). In our PFU structure, residues 433-439 are located on helix α2 and residues 440-445 are disordered (Figure 3A). Phe434 on helix α2, together with other conserved residues, forms a hydrophobic pocket on the molecular surface (Figure 3B). This hydrophobic pocket could be involved in the interaction with the Trp residue of the Hse1 SH3 domain.

Biochemical studies showed that the PUL domain of Doa1 binds to the C-terminal region of AAA-ATPase Cdc48 (9, 22, 28). Recently, the C-terminal 10 residues of Cdc48 were reported to be sufficient for Doa1 binding (37). A structural comparison of Doa1 with Kapα suggests that the concave surface of the PUL domain could serve as a binding site for Cdc48. In the structure of yeast Kapα, the SV40 NLS binds in an extended conformation to the two sites on the concave surface formed by the ARM repeats, through a complex combination of interactions (Figure 4B) (4): (i) the orientation of the NLS is defined by hydrogen bonding interactions between the backbone of the NLS and the conserved Asn residues of Kapα; (ii) the aliphatic portions of the Lys side chains of the NLS interact with the conserved Trp residues of Kapα; and (iii) the Lys side chains of the NLS form salt bridges with the conserved acidic residues of Kapα. In the present structure of Doa1, the conserved residues on the concave surface interact with the N-terminal loop of the PUL domain (Figure 5A and B). Notably, the two bidentate hydrogen bonding interactions between the invariant Asn residues (Asn578 and Asn628) and the backbone of the loop in Doa1 are reminiscent of those between the invariant Asn residues (Asn157 and Asn199) and the backbone of the

SV40 NLS in Kapα. In addition, several residues on the concave surface (Glu665, Tyr668, Arg669, and Arg706) are highly conserved. The negatively charged concave surface of Kapα is involved in the binding to the positively charged SV40 NLS (4). The positively charged concave surface of the PUL domain is complementary to the negatively charged C-terminal region of Cdc48 (Figure 5C). The crystal structure of p97 formed a hexameric ring with its flexible C-terminal region, which is located outside the ring and exposed to the solvent (5), thereby allowing Doa1/PLAA to bind the C-terminal region of Cdc48/p97 without steric hindrance. Taken together, these observations suggest that the acidic C-terminal region of Cdc48 binds in an elongated conformation to the basic concave surface of the PUL domain. The conserved residues (Arg541, Asn578, and Asn628) of Doa1 could hydrogen bond with the backbone of Cdc48. Notably, the N-terminal loop regions of yeast Kapα (54-KRRNV-58) and mouse importin α (49-KRRNV-53), which share sequence similarity with the NLS, bind to the NLS-binding site and play a role in the autoinhibitory regulation of nuclear import (21). The crystal structure of mouse importin a, with the N-terminal autoinhibitory loop bound to the NLS-binding site, revealed the structural basis for regulating nuclear import (Figure 4C) (16). Hence, it seems likely that the N-terminal loop of the PUL domain may have a regulatory role. We speculate that the present crystal structure of the PUL domain may represent a possible autoinhibited form, and the C-terminal region of Cdc48 could bind in place of the N-terminal loop to the concave surface.

Previous studies indicated that in yeast, Doa1/Ufd3 and Ufd2 compete for Cdc48 binding and may determine the fate of ubiquitinated substrates (28). The crystal structure of *S. cerevisiae* Ufd2 showed that it consists of an N-terminal variable region (residues 1-187), a core region (residues 188-879), and a U-box domain (residues 884-947) (Figure 6) (34). The core region contains irregular ARM-like repeats comprising 31 α-helices and shares structural similarity with yeast Kapα. Biochemical studies suggested that Ufd2 binds to Cdc48 through electrostatic interactions (30), and yeast two-hybrid experiments indicated that the C-terminal region, the Ufd2 core region (residues 808-856), is important for Cdc48 binding (27). Within this region, the side chains of the conserved Arg844 and Glu855 residues are exposed to the solvent. Based on these observations, Tu *et al.* speculated that the core region of Ufd2 interacts with Cdc48, using Arg844 and Glu855 (34). A structural comparison of the PUL domain of Doa1 with the core region of Ufd2 revealed that they commonly have a similar irregular ARM-like repeat structure (Figure 6). The PUL domain of Doa1 (residues 552-715) superposed onto the core region of Ufd2 (residues 614-862) with an RMSD of 4.1 Å for 148 Cα atoms (Figure 6). Notably, Arg699 and Asp713 of Doa1 are located at positions equivalent to Arg844 and Glu855 of Ufd2, respectively. In addition, the highly conserved Asn812 of Ufd2 appears to be equivalent to Asn628 of Doa1. These observations in part provide a possible structural explanation for the fact that Doa1 and Ufd2 compete for Cdc48 binding (27). However, it appears that the concave surface of the Ufd2 core region lacks the residues that correspond to Arg541, Asn578, and Asn628 of Doa1 and could hydrogen bond to the backbone of Cdc48. This might imply that the Cdc48 binding sites of Doa1 and Ufd2 only partially overlap. Consistent with this, pull-down assays suggested that Doa1 might have higher Cdc48 binding affinity that Ufd2 (28). Further details of Doa1 and Ufd2 binding to Cdc48 will be required to fully understand the mechanism by which they compete for Cdc48 binding.

During the preparation of this manuscript, the crystal structure of the PUL domain (residues 464-715) of yeast Doa1/Ufd3 was reported (36), which is essentially identical to that of the PUL domain we presented here. In addition, isothermal titration calorimetry (ITC) and mutagenesis experiments showed that the PUL domain binds to a peptide consisting of

the C-terminal 14 residues of Cdc48, and that the conserved residues (Asn578, Arg541, and Arg669) of the PUL domain are involved in Cdc48 binding. These results are consistent with our idea that the positively charged concave surface of the PUL domain serves as a binding site for Cdc48, and that Arg541 and Asn578 are involved in hydrogen bonding interactions with the Cdc48 backbone. As also observed in our structure, the N-terminal 20 residues cover the concave surface of the PUL domain in their structure. Based on this observation and the analogy with Kapα, Zhao *et al.* also proposed a regulatory role for the N-terminal loop of the PUL domain (36). The structure of Doa1 in complex with full-length Cdc48 will clarify the mechanisms by which Doa1 recognizes the C-terminal residues of Cdc48 and the possible regulatory role for the N-terminal loop.

REFERENCES

1. **Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A.** 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA **98:** 10037-10041.
2. **Beatty, B. G., Qi, S., Pienkowska, M., Herbrick, J. A., Scheidl, T., Zhang, Z. M., Kola, I., Scherer, S. W., and Seth, A.** 1999. Chromosomal localization of phospholipase A2 activating protein, an Ets2 target gene, to 9p21. Genomics **62:** 529-532.
3. **CCP4.** 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. **50:** 760-763.
4. **Conti, E., Uy, M., Leighton, L., Blobel, G., and Kuriyan, J.** 1998. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. Cell **94:** 193-204.
5. **Davies, J. M., Brunger, A. T., and Weis, W. I.** 2008. Improved structures of full-length p97, an AAA ATPase: implications for mechanisms of nucleotide-dependent conformational change. Structure **16:** 715-726.
6. **Delano, W. L.** 2010. The PyMol molecular viewer. (http://www.pymol.org/).
7. **Emsley, P., and Cowtan, K.** 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. **60:** 2126-2132.
8. **Fu, Q. S., Zhou, C. J., Gao, H. C., Jiang, Y. J., Zhou, Z. R., Hong, J., Yao, W. M., Song, A. X., Lin, D. H., and Hu, H. Y.** 2009. Structural basis for ubiquitin recognition by a novel domain from human phospholipase A2-activating protein. J. Biol. Chem. **284:** 19043-19052.
9. **Ghislain, M., Dohmen, R. J., Levy, F., and Varshavsky, A.** 1996. Cdc48p interacts with Ufd3p, a WD repeat protein required for ubiquitin-mediated proteolysis in Saccharomyces cerevisiae. EMBO J. **15:** 4884-4899.
10. **Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F.** 1999. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics **15:** 305-308.
11. **Hanway, D., Chin, J. K., Xia, G., Oshiro, G., Winzeler, E. A., and Romesberg, F. E.** 2002. Previously uncharacterized genes in the UV- and MMS-induced DNA damage response in yeast. Proc. Natl. Acad. Sci. USA **99:** 10605-10610.

12. **Hochstrasser, M., and Varshavsky, A.** 1990. In vivo degradation of a transcriptional regulator: the yeast alpha 2 repressor. Cell **61:** 697-708.

13. **Holm, L., Kaariainen, S., Rosenstrom, P., and Schenkel, A.** 2008. Searching protein structure databases with DaliLite v.3. Bioinformatics **24:** 2780-2781.

14. **Johnson, E. S., Ma, P. C., Ota, I. M., and Varshavsky, A.** 1995. A proteolytic pathway that recognizes ubiquitin as a degradation signal. J. Biol. Chem. **270:** 17442-17456.

15. **Keil, R. L., Wolfe, D., Reiner, T., Peterson, C. J., and Riley, J. L.** 1996. Molecular genetic analysis of volatile-anesthetic action. Mol. Cell. Biol. **16:** 3446-3453.

16. **Kobe, B.** 1999. Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha. Nat. Struct. Biol. **6:** 388-397.

17. **Kunze, D., MacCallum, D., Odds, F. C., and Hube, B.** 2007. Multiple functions of DOA1 in Candida albicans. Microbiology **153:** 1026-1041.

18. **Langer, G., Cohen, S. X., Lamzin, V. S., and Perrakis, A.** 2008. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat. Protoc. **3:** 1171-1179.

19. **Leggett, D. S., Hanna, J., Borodovsky, A., Crosas, B., Schmidt, M., Baker, R. T., Walz, T., Ploegh, H., and Finley, D.** 2002. Multiple associated proteins regulate proteasome structure and function. Mol. Cell **10:** 495-507.

20. **Lis, E. T., and Romesberg, F. E.** 2006. Role of Doa1 in the Saccharomyces cerevisiae DNA damage response. Mol. Cell Biol. **26:** 4122-4133.

21. **Moroianu, J., Blobel, G., and Radu, A.** 1996. The binding site of karyopherin alpha for karyopherin beta overlaps with a nuclear localization sequence. Proc. Natl. Acad. Sci. USA **93:** 6572-6576.

22. **Mullally, J. E., Chernova, T., and Wilkinson, K. D.** 2006. Doa1 is a Cdc48 adapter that possesses a novel ubiquitin binding domain. Mol. Cell Biol. **26:** 822-830.

23. **Neer, E. J., Schmidt, C. J., Nambudripad, R., and Smith, T. F.** 1994. The ancient regulatory-protein family of WD-repeat proteins. Nature **371:** 297-300.

24. **Ogiso, Y., Sugiura, R., Kamo, T., Yanagiya, S., Lu, Y., Okazaki, K., Shuntoh, H., and Kuno, T.** 2004. Lub1 participates in ubiquitin homeostasis and stress response via maintenance of cellular ubiquitin contents in fission yeast. Mol. Cell Biol. **24:** 2324-2331.

25. **Poirot, O., Suhre, K., Abergel, C., O'Toole, E., and Notredame, C.** 2004. 3DCoffee@igs: a web server for combining sequences and structures into a multiple sequence alignment. Nucleic Acids Res. **32:** W37-40.

26. **Ren, J., Pashkova, N., Winistorfer, S., and Piper, R. C.** 2008. DOA1/UFD3 plays a role in sorting ubiquitinated membrane proteins into multivesicular bodies. J. Biol. Chem. **283:** 21599-21611.

27. **Richly, H., Rape, M., Braun, S., Rumpf, S., Hoege, C., and Jentsch, S.** 2005. A series of ubiquitin binding factors connects CDC48/p97 to substrate multiubiquitylation and proteasomal targeting. Cell **120:** 73-84.

28. **Rumpf, S., and Jentsch, S.** 2006. Functional division of substrate processing cofactors of the ubiquitin-selective Cdc48 chaperone. Mol. Cell **21:** 261-269.

29. **Russell, N. S., and Wilkinson, K. D.** 2004. Identification of a novel 29-linked polyubiquitin binding protein, Ufd3, using polyubiquitin chain analogues. Biochemistry **43:** 4844-4854.

30. **Saeki, Y., Tayama, Y., Toh-e, A., and Yokosawa, H.** 2004. Definitive evidence for Ufd2-catalyzed elongation of the ubiquitin chain through Lys48 linkage. Biochem.

Biophys. Res. Commun. **320:** 840-845.

31. **Stewart, M.** 2007. Molecular mechanism of the nuclear protein import cycle. Nat. Rev. Mol. Cell Biol. **8:** 195-208.

32. **Swaminathan, S., Amerik, A. Y., and Hochstrasser, M.** 1999. The Doa4 deubiquitinating enzyme is required for ubiquitin homeostasis in yeast. Mol. Biol. Cell **10:** 2583-2594.

33. **Tong, A. H., Drees, B., Nardelli, G., Bader, G. D., Brannetti, B., Castagnoli, L., Evangelista, M., Ferracuti, S., Nelson, B., Paoluzi, S., Quondam, M., Zucconi, A., Hogue, C. W., Fields, S., Boone, C., and Cesareni, G.** 2002. A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. Science **295:** 321-324.

34. **Tu, D., Li, W., Ye, Y., and Brunger, A. T.** 2007. Inaugural Article: Structure and function of the yeast U-box-containing ubiquitin ligase Ufd2p. Proc. Natl. Acad. Sci. USA **104:** 15599-15606.

35. **Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G.** 2007. Automated structure solution with autoSHARP. Methods Mol. Biol. **364:** 215-230.

36. **Zhao, G., Li, G., Schindelin, H., and Lennarz, W. J.** 2009. An Armadillo motif in Ufd3 interacts with Cdc48 and is involved in ubiquitin homeostasis and protein degradation. Proc. Natl. Acad. Sci USA **106:** 16197-16202.

37. **Zhao, G., Zhou, X., Wang, L., Li, G., Schindelin, H., and Lennarz, W. J.** 2007. Studies on peptide:N-glycanase-p97 interaction suggest that p97 phosphorylation modulates endoplasmic reticulum-associated degradation. Proc. Natl. Acad. Sci. USA **104:** 8785-8790.