Structures of proteases for ubiquitin and ubiquitin-like modifiers

Byung Hak Ha1,2 & Eunice EunKyeong Kim1

1Life Sciences Division, Korea Institute of Science and Technology, 2School of Biological Sciences, Seoul National University, Seoul, Korea

Post-translational modifiers can alter the function of proteins in many different ways. The conjugation of ubiquitin (Ub) and ubiquitin-like modifiers (Ubls) to proteins has been shown to be especially crucial in regulating a variety of cellular processes including the cell cycle, growth control, quality control, localization and many more. It is a highly dynamic process and involves a number of enzymes called E1, E2 and E3. Ub and Ubls are removed from the target proteins by deubiquitinating enzymes (DUBs) or Ubl-specific proteases (ULPs), thereby deconjugation can act as an additional level of control over the ubiquitin-conjugation system. In addition, DUBs and ULPs are responsible for activating Ub and Ubls from their inactive corresponding precursor forms. Here we review recent progress in molecular details of these deconjugating enzymes of Ubls. [BMB reports 2008; 41(6): 435-443]

Ubiquitin and ubiquitination

Ubiquitin (Ub) is a 76 residue-long protein with a well-defined \( \alpha/\beta \) fold that is highly conserved in all eukaryotes but absent in bacteria or archaea. It is synthesized as an inactive precursor, and needs to be processed by deubiquitin enzymes (DUB) to expose the C-terminal glycine. The C-terminus of Ub is then covalently ligated to the \( \varepsilon \)-amino group of lysine residues on target proteins by the sequential action of three classes of enzymes. Ubiquitin modification of target proteins within cells plays a variety of important roles in regulation of biological processes. Most well characterized among these is the elimination of abnormal proteins, by attaching a signal for their recognition and degradation by the 26S proteasome (1, 2). It regulates a much wider range of cell processes, including endocytosis, vesicular trafficking, cell-cycle control, stress response, DNA repair, signaling, transcription and gene silencing (3-5).

Considerable progress has been made in the understanding of Ub conjugation and its role in regulating proteins through degradation or modification. The conjugation of Ub to the target substrate requires the sequential action of three classes of enzymes: an ubiquitin-activating enzyme (E1), an ubiquitin-conjugating enzyme (E2), and an ubiquitin protein ligase (E3). The E1 activates C-terminus of Ub, which results in ATP-dependent thioester formation of the E1 active site cysteine. Ub is then transferred to the active site cysteine of the E2 conjugating enzyme and finally conjugated to its substrate by an E3 ligase, which confers substrate specificity. This three-step mechanism commonly initiates all known ubiquitination reactions, so proteins can be modified on a single or multiple lysine residues by a single ubiquitin or by ubiquitin oligomers (Fig. 1).

Ubs by themselves may also be ligated to additional Ub molecules to form poly-Ub chains, thereby Ub may be attached to a target protein as a monomer or as a polyubiquitin chain, and the nature of the polyubiquitination appears to specify different fates for their target proteins. Amongst the seven lysine residues of ubiquitin Lys48 and Lys63 are the best characterized. The mono-ubiquitination leads to altered trafficking in multiple pathways (6), whereas the canonical Lys48-linked chains usually, but not always, signal proteasome proteolysis. On the other hand, Lys63-linked polyubiquitin chains appear to have functions distinct from targeting proteins to the proteolysis by the proteasome. Lys63-linked chains are believed to act as signaling molecules in diverse cellular pathways including endocytosis, stress response, and DNA repair (7).

Ubiquitin-like modifiers

Since the discovery of ubiquitin in mid 1970s, there have been several proteins found that are related to ubiquitin. They are grouped as two separate classes: ubiquitin-domain proteins (UDPs) and ubiquitin-like modifiers (Ubls). UDPs share sequence and structural homologies with Ub, but they are not conjugated to proteins. Instead they appear to function as adaptors by binding to ubiquitin or Ubls (8). Ubls, on the other hand, not only share structural homologies with Ub regardless of sequence homology, but also have glycine at the C-terminal where the carboxyl group gets attached to the lysine residue of substrates via isopeptide bond formation (Fig. 2). The substrate attachment is achieved via a series of conjugating enzymes. Furthermore, they are produced as precursors and get activated by a specific protease like Ub. Below is a brief summary of these Ubls.
Structures of proteases for Ub and Ubls
Byung Hak Ha and Eunice Eunkyoong Kim

Fig. 1. Generalized maturation, conjugation and de-conjugation of Ub and Ubls. Ub and Ubls are processed from a precursor form to expose the C-terminal glycine by a protease. They are then conjugated to substrates by a series of enzymes called E1, E2 and E3. The modified substrates carry out a variety of biological processes depending on the type of modifications. Once their role is done, the Ub and Ubls are removed from the substrate by specific enzymes and recycled.

Fig. 2. Structures of Ub and Ubls. (a) A ribbon diagram of Ub, Nedd8, SUMO and Ufm1. They all have a general secondary structure of ββαββ and tertiary structures β-grasp fold of ubiquitin. (b) Overlay of Ub and Ubls. The α-helices are shown in red while the β-strands are shown in green.

IFN-stimulated gene 15 (ISG15)
ISG15, also known as ubiquitin cross-reactive protein, is a 17-kDa protein whose expression is highly induced upon IFN stimulation (9). Unlike Ub and other Ubls, it consists of two domains, each of which bears only about 30% sequence similarity to Ub but has a similar structure (10). The conjugation of ISG15 involves a dedicated E1 (UbeL1), but seems to converge with the Ub conjugation pathway at the level of a specific E2 enzyme. Targets include STAT1, Serpina3G/Sp12A, JAK1, MAPK3/ERK1, PLCγ1, EIF2AK2/PKR, MX1/MXA and RIG-1, and is deconjugated by USP18/UBP43. It is a positive regulator of interferon-related immune response and is potentially involved in cell growth and differentiation (11, 12).

Neuronal-precursor cell-expressed developmentally down-regulated protein 8 (NEDD8)
A set of genes was discovered to be down regulated in neural precursor cells during the development of the murine brain. A 9-kDa protein with Nedd8 shares almost 60% sequence identity with Ub (13), but despite such high level of sequence identity, Nedd8 requires its own E1 (APPBP1-Uba) and E2 (Ubc12) and E3 (Dcn1) for attachment to the cellular targets (14). The known substrates are cullins, p53, and Mdm2 (15). Its conjugation to the cullin subunits of the SCF ubiquitin E3 complexes, except APC2, which contributes to the regulation of SCF mediated protein degradation of substrates such as p27, IκBα, and Hif (16-19). E3 ubiquitin ligase Mdm2, a negative regulator of p53, mediates the conjugation of NEDD8 to p53, resulting in an inhibition of transcriptional activity (20).

F-adjacent transcript-10 (FAT10)
FAT10, reported in 1995, is 18-kDa protein encoded in the major histocompatibility complex, and is synergistically inducible by tumor necrosis factor alpha and gamma interferon. It is composed of two ubiquitin-like domains and possesses a free C-terminal diglycine motif that is required for the formation of FAT10 conjugates. FAT10 can interact directly with the 26S proteasome via either one of its Ubl domains and hence provides a Ub-independent road for protein degradation (21, 22).
Small ubiquitin-like modifier (SUMO)
SUMO functions in a manner similar to Ub in that it is bound to target proteins. However, unlike Ub, SUMO is involved in a variety of cellular processes such as nuclear transport, transcriptional regulation, DNA repair, replication, apoptosis, and protein stability by altering the subcellular localization or by antagonizing other modifications (23-25). Although there is a single type of SUMO known as Smt3 in yeast (26), there are four genes expressed in vertebrates (27) and they share less than 50% sequence identity with each other. As in ubiquitination, sumoylation results in the formation of an isopeptide bond between the C-terminal Gly residue of the SUMO and the ε-amino group of a Lys residue in the target protein, and it is mediated by Uba2-Aos1 (E1), Ubc9 (E2) and Siz1, Siz2, Mtm21 (E3). Lysine residues that act as acceptors for SUMO modification are usually located in a SUMO-modification consensus motif, ψKxE (where ψ is a large hydrophobic residue and x is any residue) which is directly recognized by Ubc9 (28, 29).

Autophagy (Atg)
Studies on a autophagy-defective mutant in yeast yielded two Ubls, namely Atg8 and Atg12 and they were later confirmed to be present in human cells (30). The 15.1-kDa Atg12 and the 14.3-kDa Atg8 are involved in the membrane process of autophagy. In the case of Atg8 the substrate is a phospholipid called phosphatidylethanolamine (PE), instead of a protein, but conjugation to PE by Atg7 (E1) and Atg3 (E2) is carried out by a mechanism similar to ubiquitination. The formation of the Atg8-PE conjugate is essential for autophagy, the bulk degradation process of cytosolic components by the vacuolar/lysosomal system (31).

Ubiquitin-fold modifier 1 (Ufm1)
Ufm1 is one of the most recently discovered Ubls (32). It shares only 16% sequence identity with Ub. It does display the same tertiary structure as Ub, but it lacks a cluster of acidic residues (33). Although its biological function has not yet been fully understood since the substrates are being identified, Uba5 (E1), and Ufc1 (E2) have been identified as the enzymes responsible for conjugation of Ufm1. Since Ufm1, Uba5 and Ufc1, are conserved in metazoa and plants but not in yeast, potential roles of Ufm1 modification in various multi-cellular organisms are of special interest. Two novel proteases specific for Ufm1 have been identified (34).

Five distinct subfamily of deubiquitin enzymes
Ub and Ubls are produced as inactive precursor proteins, and they are activated by specific proteases to expose the glycine residues at the C-terminal which are necessary for the isopeptide bond linkage by Lys of the target proteins. Also specific proteases are required to remove Ub and Ubls from the target proteins. In other words, the free Ub and Ubls are controlled and these proteases play central roles in maintaining a sufficient pool of free Ub/Ubls and in the reversible processing of Ub/Ubl conjugation. The proteases are called deubiquitinating enzyme (DUB) and Ubl specific protease (ULP). Thus far, about 100 human DUBs have been identified, and based on the sequence similarities, they have been classified into five distinct subfamilies: ubiquitin-specific protease (USP), ubiquitin C-terminal hydrolase (UCH), Otubain protease (OTU), Machado-Joseph disease protease (MJD), and ubiquitin-specific JAMM domain containing metalloproteases (Table 1). In addition to DUBs, a number of ULPs have been identified (35,36 and the references therein). The overall structures and the active sites that are representative of each class are shown in Fig. 3.

Ubiquitin-specific protease (USP)
Among the DUBs, USP is the largest and most diverse. Their size varies from 300 to 800 amino acids with the catalytic domain being about 40 kDa with conserved cysteine and histidine boxes. The function and the structure of the N- and C-terminal extension domains are still unknown, but they are thought to have a regulatory function.

The crystal structure of the catalytic domain of HAUSP (herpesvirus-associated ubiquitin-specific protease or USP7), a representative USP, reveals three-domain architecture comprised of three domains, 'Fingers, Palm and Thumb' (37). The Thumb consists of eight α helices, while Palm contains eight central β strands and two helices. The fingers comprised of four β strands in the center and two at the tip is about 30 Å from the Palm-Thumb scaffold. In the complex structure of Ub aldehyde, Ub binds at the Finger region, with its C terminus bound in deep catalytic cleft with a thiohemiacetal linkage formed between the aldehyde group and the side chain of Cys223. Binding of Ub-aldehyde induces a drastic conformational change (almost 10 Å rearrangement) in the active site that realigns the catalytic triad residues for catalysis as well as ordering of two surface loops.

The overall architecture of the catalytic core domain of USP7 appears to be conserved in other members of USP family such as USP2 (38), USP8 (39), USP14 (40) and deubiquitinating enzyme from severe acute respiratory syndrome coronavirus (41). In the case of USP14, the r.m.s.d. is only 1.7 Å for 238 Cα atoms with USP7 despite low sequence identity (14%). The BL2 loop in USP14 directly contributes to the coordination of Ub-aldehyde. The catalytic residues are Cys223-His464-Asp481 arranged in similar manner as in papain. The BL2 loop is indicated with blue in Fig. 3. In this case the catalytic active site is blocked by two surface loops which gets translocated upon Ub binding and allows the C-terminal of Ub to bind at the active site (40). Recently reported structure of murine cytomegalovirus M48 in a covalent complex with an Ub-vinylmethylester shows a papain-like fold with the active site cysteine forming a thioether linkage to the suicide substrate. The Ub core interacts with an exposed β hairpin loop of M48USP. The active site triad is composed of Cys23-
Table 1. Deubiquitinating enzymes and UBL specific proteases

| Modifier | Protease | Solved structure (PDB ID) | Reference |
|----------|----------|---------------------------|-----------|
| Ub       | USP      | USP2 (2HD5)               | 38        |
|          |          | USP7/HAUSP (1NB8, 1NB6F)  | 37        |
|          |          | USP6 (2GFO, 2GWF, 2FZP)   | 39        |
|          |          | USP14 (2AYO, 2AYN)        | 40        |
|          |          | PIpro (viral USP, 2FE8)   | 41        |
|          |          | M48<sup>169</sup>(viral USP, 2TQ) | 42      |
| UCH      | Yuh1 (1CMX) |                         |           |
|          | UCHL1 (2ETL) |                         |           |
|          | UCHL3 (1UCH, 1XD3) |                     |           |
| OTU      | Otubain-1 (3BY4) |                        | 49        |
|          | Otubain-2 (1TFF) |                        | 47        |
|          | A20 (2VFJ) |                           | 48        |
| MJD      | Ataxin-3 (2AGA, 1YZB) |                    | 51, 52    |
| JAMM/MPN | A8JAMM (1R5X) |                        | 54        |
|          | Mov34 (2O95) |                           | 56        |
| SUMO     | Ulp1,2 | Ulp1 (1EUV) | 66 |
| SENPs    | SENP1 (2CKH, 2CKG, 2G4D, 2Y0, 2Y1) | 67, 68, 69 |
|          | SENP2 (1TH0, 1TGZ, 2IO0, 2IO1, 2IO2, 2IO3) | 70, 71 |
| NEDD8    | NEDP1/DEN1 | NEDP1/DEN1 (2BK8, 2BKQ, 1XT9) | 72, 73 |
| ISG15    | UBP43 |                                |           |
| Atg8     | Atg4 | Atg4B (2CY7, 2DI1) | 74, 75 |
| Ulm1     | USP1,2 | USP1 (2ZB4) | 76 |

PDB ID: Protein Data Bank (http://www.rcsb.org/)

Asp156-His158 instead of the canonical Cys-His-Asp arrangement (42).

**Ubiquitin C-terminal hydrolase (UCH)**

The UCH family is generally small with 20-30kDa, and they appear to cleave Ub off relatively small protein substrates of up to 20-30 amino acids. The sequences are well-conserved across species with approximately 40% identities. The crystal structure of human UCH-L3 is comprised of central antiparallel β-sheet flanked by α-helices on both sides (43, 44) resembling the papain cysteine protease (Fig. 3). The active site catalytic triad of Cys95, His169 and Asp184 and the oxyanion hole residue Gln89 are thought to parallel the mechanism established for the papain family of cysteine proteases. The structure of UCH-L1 (45) and yeast Yuh1 (46) are similar. Mutations in UCH-L1 gene have been reported to be linked to Parkinson’s disease, and abnormal overexpression has been reported to relate to several forms of cancer. The structures of Yuh1 with Ub-aldehyde and UCH-L3 with the Ub-vinylmethylster, suggest that there are fairly extensive interaction between the protease and Ub, and a 20-residues long loop near the active site (called ‘active-site crossover loop’) plays an important role in substrate binding. The structure of UCH-L1 also suggests an induced conformational rearrangement of the active site in order for hydrolytic activity (46).

**Otubain protease (OTU)**

OTUs are recently identified family of DUBs that belong to the ovarian tumor superfamily of proteins with conserved residues around the catalytic cysteine and histidine residues. Several OTU-domain containing proteins such as otubain1 and 2, cezanne, and A20 have been identified. The crystal structure of otubain-2, a OTU domain protein with 224 amino acids, revealed a five-stranded β-sheet sandwiched in between a small helical amino-terminal region consisting of α1 and α2 and a large helical region comprised of α3 to α10 (47). The active site containing Cys51, His 224 and Asn226 is formed at the interface of α3 and the loop connecting strands 4 and 5 (Fig. 3). Therefore, it belongs to cysteine protease superfamily. A20 is a negative regulator of NF-κB. The N-terminal OTU domain of 790 amino acid protein resembles the core of otubain-2 with an r.m.s.d. of 2.9 Å for the aligned 129 Ca atoms (48). Based on the surface and biochemical analysis the proximal and distal Ub binding sites are suggested. Based on the structures of OTU domain from yeast Otu-1 complexed with Ub and structure-guided mutagenesis, a novel mode of Ub recognition has been suggested (49).

**Machado-Joseph disease protease (MJD)**

The Josephin domain play an important role in the cellular functions of ataxin-3, which is the protein responsible for the
Fig. 3. Structures of DUBs and ULPs. (a) The ribbon diagram of superimposed USP, UCH, OUT, MJD, JAMM, SENPs, USP1 in Table 1. Light yellow represents a representative standard model to superpose other structures and light blue means Ub/Ubl structure. Detail of structures is listed below in order light yellow, bluewhite, and palecyan color. USP subfamily: HAUSP (1NBF), USP14 (2AYO), PLpro (2FE8), UCH subfamily: UCH-L3 (1XZ3), Yuh1 (1CMX), UCH-L1 (2ETL), OTU subfamily: OTU1 (3BY4), OTU2 (1TFF), A20 (2VFJ), MJD subfamily: Ataxin3 (1YZB and 2AGA), JAMM subfamily: AItJAMM (1RSX), MPN (2OS5), SENP subfamily: Ulp1 (1EUV), SENP1 (2CKH), NEDP1 (2BKR), USP and Atg4B subfamily: USP1 (2Z86), Atg4B (2CY7). The α-helices are indicated by red and β-strands by yellow arrows. (b) A representative structure of light yellow model in (a) is shown in the red α-helices and darkgreen β-strands ribbon drawing. (c) Active site of the enzymes. The catalytic residues are labeled. Figures are generated using PyMOL (http://pymol.sourceforge.net).
neurodegenerative Machado-Joseph disease which is the most common dominantly inherited cerebella ataxia worldwide. Ataxin-3 is known to bind poly Ub chains containing four or more subunits, most likely through the Ub interacting motifs as well as the N-terminal UDP domain of human Ub- and proteasome binding factors HHR23A and B which are involved in translocating proteolytic substrates to the proteasome and with the valosin-containing protein (50). NMR structures of the Josephin domain of ataxin-3 shows an α/β fold consisting of a six-stranded antiparallel β-sheet flanked by a total of seven α-helices despite primary sequence divergent (51,52). It again resembles the papain superfamily of Cys proteases, with the expected Cys, His, and Asn catalytic triad as seen in some other DUBs. Moreover, using chemical shift perturbations, direct interactions with Ub was mapped.

**JAMM/MPN**

The Rpn11, a subunit of the 19S proteasome lid subcomplex contains highly conserved Jab1/MPN domain-associated metallopeptidase (JAMM) motif (53). It is a 120 amino acid protein module with Glu-(Xaa)5-His-(Xaa)5-His-(Xaa)10-Asp, and is found in archaebacteria, bacteria, and eukaryotes. The structure of the JAMM domain from the archaebacteria protein AfJAMM revealed that it consists of eight-stranded β sheet flanked by a long α helix (α1) between the first and second strand, and a short a helix (α2) between the fourth and fifth strands. The zinc ion coordinated by two histidines, an aspartate, and a water molecule is located in a crevice formed by the convex surface of the β2-β4 sheet and α2 (54). The arrangement of Zn ligands resembles that found in thermolysin (55), but the fold is entirely distinct.

It is interesting to note that the MPN domain of Mov34 is highly homologous to that of AfJAMM domain as indicated by an r.m.s.d. of 1.8 Å for 74 matching Ca atoms, although the ability of Mov34 to bind Ub has not been established. So, it has a typical metalloprotease fold, but it is unable to bind metal ion (56). MPN domain family members comprise subunits of the proteasome, COP9 signalosome and translation initiation factor 3 complexes. Another related structure is the C-terminal domain (residues 2147-2397) of yeast protein called Prp8 which interacts with several spliceosomal proteins. It exhibits low affinity for Ub. Recently reported structure shows that it is composed of a Jab1/MPN-like core with insertions and appendices. In this case the canonical JAMM motif is replaced by Gin-Xaa-Gln-Xaa-His-Xaa-Gln-Xaa-Ser-Glu, and a water molecule was found at the Zn position (57).

**Ubl specific proteases and deconjugating enzymes**

**Specific proteases for SUMO**

SUMO pathway differs from the ubiquitin pathway, and processing and maturation of SUMO are mediated by SUMO-specific proteases. In humans, eight genes of yeast Ulp1 homologus SUMO-specific protease, SENP proteins, have been identified (58-61) while two SUMO-specific proteases, Ulp1 and Ulp2, have been characterized in *Saccharomyces cerevisiae* (62, 63). They share various degree of sequence identities, and they are not all specific for SUMO as SENP8 (also called NEDP1 or DEN1) is specific for NEDD8 (64, 65). SENP1, SENP2 and SENP3 are SUMO-specific proteases, but they have distinct sub-cellular localizations dictated by their non-conserved N-terminal regions. For example SENP1 is localized in nuclear and SENP3 in nucleolar, while SENP6 seems to be localized in the cytoplasm. SENP2, on the other hand, is found at cytoplasmic, nuclear pore, or nuclear body.

Both Ulp1 and SENP proteases are classified as a large group of cysteine proteases. They catalyze cleavage of the peptide bond after the ultimate C-terminal glycine as well as deconjugation from the sumoylated target. Both are catalyzed by a conserved 200-residue Ulp-specific protease catalytic domain. The catalytic domain of Ulp1 shows structural similarity to other canonical cysteine proteases (66). SENP1 and SENP2 also share the structural similarity (67-71). The complex structures of Ulp1-Smt3, SENP2-SUMO1, and SENP1-SUMO1/2 show no dramatic conformational changes as seen in the DUBs when compared to the Ubl free structure. However, there are local rearrangements near the active site. The structures of catalytically inactive Ulp5 of SENP1 and SENP2 bound to SUMO precursors and SUMOylated RanGAP1 (69, 70). Both are productive SENP complexes with substrates, and they show cis-configuration for the scissile peptide bond. For SUMO precursors, this results in a 90° kink in the SUMO C-terminal. A Similar orientation was observed for the SUMOylated RanGAP1. To make such a cis orientation of scissile bond the Trp residue on SENP activity cleft plays a role as the clamp to close and stabilize kinked substrate by a hydrogen-bonding network.

**NEDP1**

NEDP1 has a central five-stranded β-sheet in which the middle strand β-5 is antiparallel to the other four. One face of the sheet packs against two helices (α-2 and α-7), and the axes of both helices are parallel to each other and to the β-strands, resulting in a narrow channel (72). This is homologous to the Ulp1 protease in yeast with an r.m.s.d. of 1.3 Å for 140 Ca atoms. Binding of NEDD8 induces a dramatic conformational change in a flexible loop centered on Gin96 in NEDP1 that swings over the C-terminus of NEDD8, locking it into an extended β-structure optimal for catalysis. Even though the complex structure between NEDP1-NEDD8 and Ulp1-SUMO is very similar, a comparison of two complex structures reveals several differences in the position, size, and charge of interaction residues such as two positively charged patches in Ulp1, and uniformly negatively charged patches of NEDP1. Moreover, the difference of C-terminus 72nd residue of NEDD8 (Ala) and Ub (Arg) contributes significantly to the ability of NEDP1 to discriminate between them. The Arg will perturb the interface between NEDD8 and NEDP1 due to electrostatic and van der Walls repulsion with Arg42 and Arg74 (72).
Atg4
A novel cysteine protease, Atg4, is responsible for processing and deconjugation of Atg8 from PE (31). Despite no obvious sequence homology with known proteases, the structure of human Atg4B exhibits a classical papain-like fold with an additional small \(\alpha/\beta\) fold domain called “short fingers domain” (73). This domain is thought to be the binding site for Atg8. The active site cleft in the crystal structure is masked by a loop (residue 259-262), suggesting there would be a conformational change upon substrate binding (73, 74). In this enzyme, the catalytic triad is formed by Cys74, Asp278 and His280 which is somewhat different from the canonical catalytic triad.

UiSP
Two novel cysteine proteases, UiSP1 and UiSP2, are reported for the processing of Ufm1 (34). Again, the crystal structure of UiSP1 revealed a papain-like structure with Cys53, Asp175 and His177 forming the catalytic triad and Tyr41 that participates in the formation of the oxyanion hole (75). The catalytic residues are posed similar to that of papain except that the aspartate and the histidine are from “Asp-Pro-His” box. In fact, similar configuration of the active site residues are observed in Atg4B (73,74) and M48USP (42) suggesting that they form a new subfamily of the cysteine protease superfamily. The binding study carried out by NMR and ITC suggests substantial interaction between the protease and Ufm1.

Conclusion remarks
There have been great advances in the identifying and understanding of the ubiquitin and ubiquitin-like proteins. As seen above, most DUBs are cysteine proteases and only a few are metallopeptases. In the case of cysteine proteases they seem to share many common features especially at the active site despite distinct overall structures. In particular, they have similar orientation of Cys-His-Asp triad as well as Gln/Asn or Tyr stabilizing the oxyanion intermediate and perhaps a loop or a hydrophobic residue acting as a clamp.

Acknowledgements
This work was supported by the KIST Institutional Grant and the Functional Proteomics Center, the 21C Frontier Research & Development Program from Korea Ministry of Science and Technology.

REFERENCES
1. Ciechanover, A., Finley, D. and Varshavsky, A. (1984) Ubiquitin dependence of selective protein degradation demonstrated in the mammalian cell cycle mutant ts85. Cell 37, 57-66.
2. Finley, D., Ciechanover, A. and Varshavsky, A. (1984) Thermolability of ubiquitin-activating enzyme from the mammalian cell cycle mutant ts85. Cell 37, 43-55.
3. Hicke, L. (2001) A new ticket for entry into budding vesicles-ubiquitin. Cell 106, 527-530.
4. Huang, T. T. and D’Andrea, A. D. (2006) Regulation of DNA repair by ubiquitylation. Nat. Rev. Mol. Cell Biol. 7, 323-334.
5. Di Fiore, P. P., Polo, S. and Hofmann, K. (2003) When ubiquitin meets ubiquitin receptors: a signalling connection. Nat. Rev. Mol. Cell Biol. 4, 491-497.
6. Hicke, L. and Dunn, R. (2003) Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. Annu. Rev. Cell Dev. Biol. 19, 141-172.
7. Wu, C. J., Conze, D. B., Li, T., Srinivasula, S. M. and Ashwell, J. D. (2006) Sensing of Lys 63-linked polyubiquitination by NEMO is a key event in NF-kappaB activation. Nat. Cell Biol. 8, 398-408.
8. Jentsch, S. and Pyrowolakis, G. (2000) Ubiquitin and its kin: how close are the family ties? Trends in Cell Biol. 10, 335-342.
9. Korant, B. D., Blomstrom, D. C., Jonak, G. J. and Knight, E. Jr. (1984) Interferon-induced proteins. Purification and characterization of a 15,000-dalton protein from human and bovine cells induced by interferon. J. Biol. Chem. 259, 14835-14839.
10. Narasimhan, J., Wang, M., Fu, Z., Klein, J. M., Haas, A. L. and Kim, J. J. (2005) Crystal structure of the interferon-induced ubiquitin-like protein ISG15. J. Biol. Chem. 280, 72356-72365.
11. Zhao, C., Denison, C., Huibregtse, J. M., Gygi, S. and Krug, R. M. (2003) Human ISG15 conjugation targets both IFN-induced and constitutively expressed proteins functioning in diverse cellular pathways. Proc. Natl. Acad. Sci. U. S. A. 102, 10200-10205.
12. Malakhova, O. A., Yan, M., Malakhov, M. P., Yuan, Y., Ritchie, K. J., Kim, K. I., Peterson, L. F., Shuai, K. and Zhang, D.E. (2003) Protein ISGylation modulates the JAK-STAT signaling pathway. Genes Dev. 17, 453-460.
13. Kumar, S., Yoshida, Y. and Noda, M. (1993) Cloning of a cDNA which encodes a novel ubiquitin-like protein. Biochem. Biophys. Res. Commun. 195, 393-399.
14. Hochstrasser, M. (1998) There’s the Rub: a novel ubiquitin-like modification linked to cell cycle regulation. Genes Dev. 12, 901-907.
15. Xinorodimas, D. P., Saville, M. K., Bourdon, J. C., Hay, R. T. and Lane, D. P. (2004) Mdm2-mediated NEDD8 conjugation of p53 inhibits its transcriptional activity. Cell 118, 83-97.
16. Podust, V. N., Brownell, J. E., Gladysheva, T. B., Luo, R. S., Wang, C., Coggins, M. B., Pierce, J. W., Lightcap, E. S. and Chau, V. (2000) A Nedd8 conjugation pathway is essential for proteolytic targeting of p27Kip1 by ubiquitination. Proc. Natl. Acad. Sci. U. S. A. 97, 4579-4584.
17. M., Parent, L. A., Coggins, M. B., Pierce, J. W., Podust, V. N., Luo, R. S., Chau, V. and Palombella, V. J. (2000) Nedd8 modification of cul-1 activates SCF(betaTrCP)-dependent ubiquitination of IkappaBalpha. Mol. Cell. Biol. 20, 2326-2333.
18. Ohh, M., Kim, W. Y., Moslehi, J. J., Chen, Y., Chau, V., Read, M. A. and Kaelin, W. G., Jr. (2002) An intact NEDD8 pathway is required for Cullin-dependent ubiquitylation in mammalian cells. EMBO Rep. 3, 177-182.

http://lbmbreports.org
19. Hori, T., Osaka, F., Chiba, T., Miyamoto, C., Okabayashi, K., Shimbara, N., Kato, S. and Tanaka, K. (1999) Covalent modification of all members of human cullin family proteins by NEDD8. *Oncogene* **18**, 6829-6834.

20. Xirodimas, D. P., Saville, M. K., Bourdon, J. C., Hay, R. T. and Lane, D. P. (2004) Mdm2-mediated NEDD8 conjugation of p53 inhibits its transcriptional activity. *Cell* **118**, 83-97.

21. Rassi, S., Schmidtke, G. and Grottrup, M. (2001) The ubiquitin-like protein FAT10 forms a covalent conjugates and induces apoptosis. *J. Biol. Chem.* **276**, 35334-35443.

22. Hopp, M. S., Kalveram, B., Rassi, S., Grottrup, M. and Schmidtke, G. (2005) FAT10, a ubiquitin-independent signal for proteasome degradation. *Mol. Cell. Biol.* **25**, 3483-3491.

23. Girdwood, D., Bumpass, D., Vaughan, O. A., Thain, A., Anderson, L. A., Snowden, A. W., Garcia-Wilson, E., Perkins, N. D. and Hay, R. T. (2003) P300 transcriptional repression is mediated by SUMO modification. *Mol. Cell* **11**, 1043-1054.

24. Seeler, J. S. and Dejean, A. (2003) Nuclear and unclear functions of SUMO. *Nat. Rev. Mol. Cell Biol.* **4**, 690-699.

25. Muller, S., Ledl, A. and Schmidt, D. (2004) SUMO: a regulator of gene expression and genome integrity. *Oncogene* **23**, 1998-2008.

26. Meluh, P. B. and Koshland, D. (1995) Evidence that the MIF2 gene of Saccharomyces cerevisiae encodes a centromere protein with homology to the mammalian centromere protein CENP-C. *Mol. Cell. Biol.* **6**, 793-807.

27. Bohren, K. M., Nadkarni, V. J., Song, H., Gabbay, K. H. and Owerbach, D. (2004) A M55V polymorphism in a novel SUMO gene (SUMO-4) differentially activates heat shock transcription factors and is associated with susceptibility to type I diabetes mellitus. *J. Biol. Chem.* **279**, 27233-27238.

28. Bernier-Villamor, V., Sampson, D. A., Matunis, M. J. and Lima, C. D. (2002) Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell* **108**, 345-356.

29. Sampson, D. A., Wang, M. and Matunis, M. J. (2001) The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. *J. Biol. Chem.* **276**, 21664-21669.

30. Ichimura, Y., Kirisako, T., Takao, T., Satomi, Y., Shimonishi, Y., Ishihara, N., Mizushima, N., Tanida, I., Komimori, E., Ohsumi, M., Noda, T. and Ohsumi, Y. (2000) A ubiquitin-like system mediates protein lipidation. *Nature* **408**, 488-492.

31. Ohsumi, Y. (2001) Molecular dissection of autophagy: two ubiquitin-like systems. *Nat. Rev. Mol. Cell Biol.* **2**, 211-216.

32. Komatsu, M., Chiba, T., Tatsuki, K., Iemura, S., Tanida, I., Okazaki, N., Ueno, T., Kominami, E., Natsume, T. and Tanaka, K. (2004) A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier. *EMBO J.* **23**, 1977-1986.

33. Sasakawa, H., Sakata, E., Yamauchi, Y., Komatsu, M., Tatsuki, K., Kominami, E., Tanaka, K. and Kato, K. (2006) Solution structure and dynamics of Ufm1, a ubiquitin-fold modifier 1. *Biochem. Biophys. Res. Commun.* **343**, 21-26.

34. Kang, S. H., Kim, G. R., Seong, M., Baek, S. H., Seol, J. H., Bang, O. S., Ovaa, H., Tatsuki, K., Komatsu, M., Tanaka, K. and Chung, C. H. (2007) Two novel ubiquitin-fold modifier 1 (Ufm1)-specific proteases, UfSP1 and UfSP2. *J. Biol. Chem.* **282**, 5256-5262.

35. Love, R. R., Catic, A., Schlcker, C. and Ploegh, H. L. (2007) Mechanisms, biology and inhibitors of deubiquitinating enzymes. *Nat. Chem. Biol.* **3**, 697-705.

36. Nijman, S. M., Luna-Vargas, M. P., Veld, A., Brummelkamp, T. R., Dirac, A. M., Sixma, T. K. and Bernard, R. (2005) A genomic and functional inventory of deubiquitinating enzymes. *Cell* **123**, 773-786.

37. Hu, M., Li, P., Li, M., Li, W., Yao, T., Wu, J. W., Gu, W., Cohen, R. E. and Shi, Y. (2002) Crystal structure of a UBP-family deubiquitylating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell* **111**, 1041-1054.

38. Renatus, M., Parrado, S. G., D’Arcy, A., Edflof, U., Gerhardt, B., Hassiepen, U., Pierrat, B., Riedl, R., Vinzenz, D., Worpenberg, S. and Kroemer, M. (2006) Structural basis of ubiquitin recognition by the deubiquitinating protease USP2. *Structure* **14**, 1293-302.

39. Avvakumov, G. V., Walker, J. R., Xue, S., Finerty, P. J. Jr., Mackenzie, F., Newman, E. M. and Dhe-Paganon, S. (2006) Amino-terminal dimerization, NRD1-threonine interaction, and inhibited catalytic domain conformation of the ubiquitin-specific protease 8 (USP8). *J. Biol. Chem.* **281**, 38061-38070.

40. Hu, M., Li, P., Song, L., Jeffrey, L. P., Chenova, T. A., Baker, S. C., Stevens, R. C. and Mesecar, A. D. (2006) Severe acute respiratory syndrome coronavirus papain-like protease: structure of a viral deubiquitinating enzyme. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5717-5722.

41. Ratia, K., Saikatendu, K. S., Santarsiero, B. D., Barrett, N., Baker, S. C., Stevens, R. C. and Mesecar, A. D. (2006) Severe acute respiratory syndrome coronavirus papain-like protease: structure of a viral deubiquitinating enzyme. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 677-687.

42. Schleier, C., Weihofen, A. W., Frijns, E., Kattenhorn, L. M., Gautel, R. and Ploegh, H. L. (2007) Structure of a herpesvirus-encoded cytostatic protease reveals a unique class of deubiquitinating enzymes. *Mol. Cell* **25**, 487-497.

43. Misaghi, S., Galardy, P. J., Meester W. J. N., Ovaa, H., Ploegh, H. L. and Gaudet R. (2005) Structure of the ubiquitin hydrolase UCH-L3 complexed with a suicide substrate. *J. Biol. Chem.* **280**, 1512-1520.

44. Das, C., Hoang, Q. Q., Keinbring, C. A., Luchansky, J. S., Oh, R. K., Ray, S. S., Lansbury, P. T., Ring, D. and Petsko, G. A. (2006) Structural basis for conformational plasticity of the Parkinson’s disease-associated ubiquitin hydrolase UCH-L1. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4575-4680.

45. Johnston, S. C., Frijns, E., Kattenhorn, L. M., Gautel, R. and Ploegh, H. L. (2007) Structure of a ubiqutin-conjugating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell* **111**, 1041-1054.

46. Renatus, M., Parrado, S. G., D’Arcy, A., Edflof, U., Gerhardt, B., Hassiepen, U., Pierrat, B., Riedl, R., Vinzenz, D., Worpenberg, S. and Kroemer, M. (2006) Structural basis of ubiquitin recognition by the deubiquitinating protease USP2. *Structure* **14**, 1293-302.

47. Avvakumov, G. V., Walker, J. R., Xue, S., Finerty, P. J. Jr., Mackenzie, F., Newman, E. M. and Dhe-Paganon, S. (2006) Amino-terminal dimerization, NRD1-threonine interaction, and inhibited catalytic domain conformation of the ubiquitin-specific protease 8 (USP8). *J. Biol. Chem.* **281**, 38061-38070.
OTU domain and mechanistic insights into deubiquitination. Biochem. J. 409, 77-85.

49. Messick, T. E., Russell, N. S., Iwata, A. J., Sarachan, K. L., Shiekhattar, R., Shanks, J. R., Reyes-Turcu, F. E., Wilkinson, K. D. and Marmorstein, R. (2008) Structural basis for ubiquitin recognition by the Otu1 ovarian tumor domain protein. J. Biol. Chem. 283, 11038-11049.

50. Doss-Pepe, E. W., Stenroos, E. S., Johnson, W. G. and Madura, K. (2003) Ataxin-3 interactions with rad23 and valosin-containing protein and its associations with ubiquitin-chains and the proteasome are consistent with a role in ubiquitin-mediated proteolysis. Mol. Cell Biol. 23, 6469-6483.

51. Nicastro, G., Menon, R. P., Asino, L., Knowles, P. P., McDonald, N. Q. and Pastore, A. (2005) The solution structure of the Josephin domain of ataxin-3: structural determinants for molecular recognition. Proc. Natl. Acad. Sci. U. S. A. 102, 10493-10498.

52. Mao, Y., Senic-Matuglia, F., Di Fiore, P. P., Polo, S., Hochedon, M. E. and De Camilli, P. (2005) Deubiquitinating function of ataxin-3: insights from the solution structure of the Josephin domain. Proc. Natl. Acad. Sci. U. S. A. 102, 12700-12705.

53. Verma, R., Aravind, L., Oania, R., McDonald, W. H., Yates, J. R., Koonin, E. V. and Deshaies, R. J. (2002) Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. Science 298, 611-615.

54. Ambroggio, X. I., Rees, D. C. and Deshaies, R. J. (2004) JAMM: a metalloprotease-like zinc site in the proteasome and signalosome. PloS Biol. 2, 113-118.

55. Kurniawan, K., Ouyang, J., Lee, H. Y., Catic, A., Ploegh, H. and Gill, G. (2006) The SUMO-specific protease SENPS is required for cell division. Mol. Cell Biol. 26, 4489-4498.

56. Gong, L. and Yeh, E. (2006) Characterization of a family of nucleolar sumo-specific proteases with preference for sumo-2 or sumo-3. J. Biol. Chem. 281, 15869-15877.

57. Li, S. J. and Hochstrasser, M. (1999) A new protease required for cell-cycle progression in yeast. Nature 398, 246-251.

58. Li, S. J. and Hochstrasser, M. (2000) The yeast ULP2 (SMT4) gene encodes a novel protease specific for the ubiquitin-like Smt3 protein. Mol. Cell. Biol. 20, 2367-2377.

59. Yeh, E. T., Gong, L. and Kamitani, T. (2000) Ubiquitin-like Smt3 protein. J. Biol. Chem. 278, 28892-28900.

60. Di Bacco, A., Ouyang, J., Lee, H. Y., Catic, A., Ploegh, H. and Gill, G. (2006) The SUMO-specific protease SENPs is required for cell division. Mol. Cell Biol. 26, 4489-4498.

61. Gong, L. and Yeh, E. (2006) Characterization of a family of nucleolar sumo-specific proteases with preference for sumo-2 or sumo-3. J. Biol. Chem. 281, 15869-15877.