Structure of the ubiquitin-interacting motif of S5a bound to the ubiquitin-like domain of HR23B

Kenichiro Fujiwara¹, Takeshi Tenno²,1, Kaoru Sugasawa³, Jun-Goo Jee⁴,1, Izuru Ohki⁵, Chojiro Kojima⁶, Hidehito Tochio¹, Hidekazu Hiroaki¹, Fumio Hanaoka³,7, Masahiro Shirakawa¹,4,*

¹ Graduate School of Integrated Science, Yokohama City University, 1-7-29 Suehiro, Tsurumi, Yokohama, Kanagawa 230-0045

² Graduate School of Science and Engineering, Ehime University, 3 Bunkyo, Matsuyama, Ehime 790-8577

³ Cellular Physiology Laboratory, Discovery Research Institute, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198

⁴ Genomic Sciences Center, RIKEN, 1-7-22 Suehiro, Tsurumi, Yokohama, Kanagawa 230-0045

⁵ Department of Structural Biology, Biomolecular Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565-0874

⁶ Graduate School of Biological Sciences, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0101

⁷ Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamada-oka, Suita, Osaka 565-0871, Japan

* Correspondence should be addressed to Masahiro Shirakawa

Tel: +81-45-508-7213 Fax: +81-45-508-7361 e-mail: shirakawa@tsurumi.yokohama-cu.ac.jp

Running title: Structure of a UIM-UbL complex

Key words: proteasome/ubiquitin/ubiquitin-interacting motif/ubiquitin-like domain
SUMMARY

Ubiquitination, a modification in which single or multiple ubiquitin molecules are attached to a protein, serves signalling functions that control several cellular processes. The ubiquitination signal is recognized by downstream effectors, many of which carry a ubiquitin-interacting motif (UIM). Such interactions can be modulated by regulators carrying a ubiquitin-like (UbL) domain, which binds UIM by mimicking ubiquitination. Of them, HR23B regulates the proteasomal targeting of ubiquitinated substrates, DNA repair factors and other proteins. Here we report the structure of the UIM of the proteasome subunit S5a bound to the UbL domain of HR23B. The UbL domain presents one hydrophobic and two polar contact sites for interaction with UIM. The residues in these contact sites are well conserved in ubiquitin, but ubiquitin also presents a histidine at the interface. The pH-dependent protonation of this residue interferes with the access of ubiquitin to UIM and UBA, and its mutation to a smaller residue increases the affinity of ubiquitin for UIM.

INTRODUCTION

Chains or single molecules of ubiquitin can be attached to cellular proteins, giving rise to poly- or monoubiquitination, respectively, which mediates distinct cellular signals (1–3). Polyubiquitination mainly targets proteins for degradation by the proteasome and can modulate diverse biological processes, such as cell cycle progression, apoptosis, antigen presentation and stress response; by contrast, monoubiquitination can act as both a endocytic sorting signal in vesicular transport and a key regulator of transcription, replication and DNA repair (1–5). The versatility of ubiquitination signals is mediated by different downstream regulatory factors, many of which carry a UIM or ubiquitin-associated (UBA) domain, or both (6–8). UIM was first identified in the proteasome subunit S5a and occurs in many proteins involved in rapid protein degradation, or protein trafficking, such as endocytosis (Figure 1b; (7)).
So far, 27 UIM-containing proteins have been identified in humans (9). The S5a UIM requires a string of at least four Lys-48-linked ubiquitins for efficient binding (10). S5a of higher eukaryotes has two UIMs, located near each other in the carboxyl terminus. The C-terminal UIM by itself binds polyubiquitin chains as efficiently as the full-length S5a protein, whereas the N-terminal UIM alone binds polyubiquitin chains less efficiently (10). Nevertheless, a sequence alignment by Young and co-workers revealed that S5a of *S. cerevisiae* (Rpn10p) has only one UIM sequence that corresponds to the N-terminal UIM of higher eukaryotes (Figure 1b; (10)). Recently, it has been shown that *S. cerevisiae* S5a contributes to the targeting of a subset of ubiquitinated substrates to the proteasome, but it is not the sole factor involved in polyubiquitin recognition (11, 12). Subunit S6’ of mammalian proteasome was also shown to recognize the polyubiquitin degradation signal (13). In contrast to S5a UIMs, those of the endocytic factors Hrs, Eps15, Eps15R, and Vps27 are proposed to bind monoubiquitin tags (8, 14–15). The affinity between these UIMs and monoubiquitin seems low: the dissociation constant ($K_d$) for Hrs binding to monoubiquitin is 230–300 µM (16, 17).

Interactions of ubiquitin tags with UIM or UBA of downstream effectors are modulated by proteins that bear UbL domains, some of which are thought to bind UIM by mimicking ubiquitination. Similar to ubiquitin-like modifier proteins such as NEDD8, UbL domains share high sequence homology with ubiquitin (typically 24-33 %) and thus adopt ubiquitin folds, as shown by the structure determination of the UbLs of PLIC-2, Parkin and HR23B (18–20 and 30).

Of the UbL-containing proteins, HR23B is the best characterized and has been shown to target the excision repair factor XPC/Rad4 and the endoplasmic-recticulum-associated deglycosylation enzyme Png1p to the proteasome through interaction of its UbL with the proteasome (21–26). More recently, Rad23 (the *S. cerevisiae* homologue of HR23B) was
proposed to deliver a variety of ubiquitinated cellular proteins to proteasome (27). The interaction between the proteasome and the UbL domain of HR23B is essential for all of these functions. In higher eukaryotes, the C-terminal UIM of S5a has been proposed to be the receptor site for the UbL domain, to which it binds efficiently and with high affinity ($K_d = 3.4 \mu M$; this study, see Results and Discussion); by contrast, the N-terminal UIM does not bind to UbL (22). In yeast, however, S5a does not have a sequence corresponding to this C-terminal UIM (10) and does not bind to Rad23 (11). Instead, the proteasome subunit Rpn1 serves as its receptor. These observations may suggest that the HR23B receptor sites in the proteasome differ between yeast and higher eukaryotes.

The interaction between ubiquitin/UbL and UIM has been studied by mutagenesis and NMR chemical shift perturbation experiments (10, 16, 18, 19, 28, 29 and 30). These experiments all suggest that the conserved hydrophobic patch composed of the side chains of Leu 8, Ile 44, and Val 70 of ubiquitin serves as the binding site for S5a UIM or the proteasome.

To understand the structural basis for this recognition, we have solved the structure of the C-terminal UIM of human proteasome subunit S5a in complex with the UbL domain of human HR23B by solution NMR. A comparison of the UIM binding surface of UbL with the corresponding region of ubiquitin, coupled with extensive mutagenesis studies, indicates that the residues forming these interfaces are well conserved between UbL and ubiquitin, except for His 68 of ubiquitin. This residue can regulate the access of the UIM or UBA to ubiquitin in a pH-dependent manner. The present structural and mutational data enable us to revise the consensus sequence of UIM.

**EXPERIMENTAL PROCEDURES**

**Preparation of proteins and peptides**

We expressed the N-terminal 87 residues of human HR23B, referred to as UbL, and residues
263-307 of S5a, referred to as UIM, as His-tagged proteins in *Escherichia coli* BL21(DE3). Labelled proteins were obtained by growing *E. coli* in synthetic media containing $^{15}\text{NH}_4\text{Cl}$ and $^{15}\text{NH}_4\text{Cl}/^{13}\text{C-}\text{glucose}$, respectively, and were purified chromatographically. UbL has an extra HHHHHH sequence at its C terminus, whereas UIM has an extra GSH sequence at its N-terminus after cleavage by thrombin. The UIM-UbL complex was formed by titrating UIM into UbL, using changes in the amide group resonances of UbL in the $^{15}\text{N}-^{1}\text{HHSQC}$ spectrum to indicate 1:1 stoichiometry. The UIM-UbL complex was in slow exchange on the chemical shift timescale. Samples for NMR measurements typically comprised 1.0 mM UIM-UbL complex in 20 mM potassium phosphate buffer (pH 6.8) and 5 mM KCl. Various combination of isotopically labelled UIM-UbL complex were used to obtain chemical shift assignments and distance restraints.

**NMR spectroscopy**

NMR spectra were acquired at 303 K using Bruker DRX500 or DRX800 spectrometer. For assigning the $^{1}\text{H}$, $^{15}\text{N}$ and $^{13}\text{C}$ resonances, we carried out a series of three-dimensional (3D) triple-resonance experiments (31). The stereospecific assignment of the methyl groups of the leucines and valines was achieved using 15% fractionally $^{13}\text{C}$-labelled protein. Distance restraints were derived from $^{15}\text{N}$- and $^{13}\text{C}$ resolved 3D and 4D NOESY experiments with a mixing time of 100 ms. Intermolecular restraints were derived from the 4D $^{13}\text{C}/^{15}\text{N}$-edited NOESY spectrum of a complex formed from $^{13}\text{C}$-labelled UbL and $^{15}\text{N}$-labelled UIM with a mixing time of 150 ms, and 3D $^{13}\text{C}$-filtered-$^{13}\text{C}$-edited NOESY spectra of a complex formed from one unlabelled component and one $^{13}\text{C}/^{15}\text{N}$-labelled component with a mixing time of 100 ms. For torsion angle ($\phi, \psi$) restraints, the backbone vicinal coupling constants ($^{3}J_{\text{HN}, \text{H}\alpha}$) were determined by means of HNHA and a database search procedure based on
Structure of a UIM-UbL complex

backbone chemical shifts, using the program TALOS (32). pH titration of the unique histidine of ubiquitin, His 68, was carried out by measuring one-dimensional heteronuclear multiple-bond \( ^1H-^{15}N \) correlation (HMBC) spectra of \( ^{15}N \)-labelled human ubiquitin at 303 K.

**Structure calculations**

For structure determination of the UIM-UbL complex, 100 structures were initially calculated by a simulated annealing procedure in CNS (33) and further refined by AMBER 7 (34). All interproton distance restraints were derived from unambiguously assigned NOE cross-peaks. The upper limits of NOE restraints were calibrated according to their intensities using ARIA protocols. For the final steps of the calculations, restraints were included for 25 slowly exchanging backbone amides [2.8-3.4 Å (N-O), 1.8-2.4 Å (H-O)]. In total, 1,630 meaningful NOE restraints (639 intraresidual, 372 sequential, 211 medium range, 305 intramolecular long range and 103 intermolecular) and 50 dihedral angle restraints were used.

Refinement using AMBER 7 consisted of 20-ps molecular dynamics, followed by 1000 steps of energy minimization. To approximate solvent interaction, a generalized Born model was used (35). Force constants were 20 kcal mol\(^{-1}\) Å\(^{-1}\) for distance restraints and 150 kcal mol\(^{-1}\) rad\(^{-2}\) for dihedral angle restraints. Of refined 100 structures, the best 20 structures were selected and analysed using MOLMOL (36), AQUA and PROCHECK-NMR (37) software. No NOE was violated by more than 0.3 Å, and no torsion restraint was violated by more than 4Ú. The hydrogen bond between UIM Ser 294 and UbL Gly 50 was assumed on the basis of both mutagenesis of UIM Ser 294 and observation of the bond in 98 out of the 100 calculated structures.

**Mutational analyses**

Mutant constructs were prepared with the GeneEditor *in vitro* site-directed mutagenesis
system (Promega). The binding of S5a UIM to ubiquitin, HR23B UbL, and their mutants was determined by surface plasmon resonance using a BIACORE-X instrument. The equilibrium binding affinity of GST-UIM or its mutant immobilized on a CM5 chip was analysed by monitoring the change in response units as a function of ubiquitin, UbL or there mutant concentrations ranging from 0.1-500 µM at flow rate of 20 µl/min in buffer containing 10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, and 0.005 % Surfactant P20. The binding of the UBA of yeast Dsk2p to ubiquitin was determined in the same way using a BIACORE-X instrument. A GST-fusion protein that contains Dsk2p residues 328-373 (38) was immobilized onto a CM5 chip.

The effect of mutations on the affinity of S5a UIM for tetraubiquitin was examined by a surface plasmon resonance-based competition assay. Resonance curves were measured for 0.1 µM tetraubiquitin binding to immobilized GST-UIM in the presence of 40 µM Histagged wild type UIM or its point mutants. A buffer containing 10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA and 0.005 % Surfactant P20 was used at flow rate of 20 µl/min. The effect of ubiquitin mutation of His 68 to valine (H68V) for its binding affinity to the UIM of S5a was also analysed by GST-pull down assay. 12 µg of GST-UIM was incubated with 3 µg of wild-type or mutant ubiquitin in 100 µl of 20 mM HEPES (pH 7.9), 100 mM NaCl, 20% glycerol, 0.1% Nonidet P-40 and 200 µg/ml bovine serum albumin. Bound proteins were analysed by GSH-mediated pull-down assay coupled with SDS-polyacrylamide gel electrophoresis and silver staining.

RESULTS
Structure determination

The structure of the UIM of human proteasome subunit S5a in complex with the UbL domain of human HR23B was determined from a total of 1,730 NMR-derived restraints (Figure 2a.
and Table I). The structure of residues 1-75 of UbL and that of the UIM residues 278-296 of S5a are well defined (Figure 2b). By contrast, the terminal residues of UIM have no contact with UbL and are therefore disordered. In the complex, UbL folds into an α/β structure comprising a layer of five-stranded twisted β-sheet, backed by a long helix (Figure 3a). It closely resembles the structures of ubiquitin ((39), PDB code: 1d3z), as indicated by the root-mean-square deviation (r.m.s.d.) of 0.741 Å over 30 Cα coordinates for residues in the regions of secondary structure elements. The UIM-bound UbL of HR23B has a similar structure to that of the unliganded form (30). Therefore, the UbL domain of human HR23B shares the same fold as members of the ubiquitin-like modifier family and other ubiquitin-like domains.

UIM adopts a ‘hook’-like conformation, consisting of an N-terminal loop (residues 278-282) followed by an α helix (residues 283-296), which are flanked on either side by N- and C-terminal unstructured regions (residues 263-277 and 297-307). The helix dominates the binding interface, fitting snugly along strands β3 and β5 of UbL, whereas the N-terminal loop runs alongside the loop that connects helix α1 and strand β3 of UbL, mediating minor interfacial contacts. The helical portion of the UbL-bound UIM of S5a, residues 284-296, adopts a nearly identical fold to the corresponding part of a free UIM of yeast Vps27p (residues 303-315), whose crystal structure was solved as an antiparallel four helix bundle consisting of four molecules (40).

**Recognition of UIM by UbL**

UbL has three principle contact sites with the helix of UIM that mediate a central hydrophobic and two polar interactions. The hydrophobic contact site is defined by the outward facing residues Leu 8, Ile 47, Val 71 and Met 73 of the UbL β-sheet; the side chains of these residues form a pocket that packs tightly around the methyl group of Ala 290 of UIM (Figure 3b and c). Leu 8 also forms hydrophobic interactions with Leu 278 of UIM. Mutation of Leu
8 or Val 71 of UbL, or of Ala 290 of UIM, largely abolishes binding, indicating that the surface complementarity of these residues is important for binding (Table II).

The association of the UIM helix is further stabilized by two polar interactions near both of its ends. Ser 294 of UIM accepts a hydrogen bond from the main chain amide group of Gly 50 of UbL, which adopts a consecutive type IV (residues 47-50) and type I’ (residues 48-51) \( \beta \)-turns (\( \beta \)Tr, Figure 3b and c). This conformation, which is stabilized by an aromaticmethyl stacking interaction between the phenol ring of Tyr 48 and the methyl group of Ala 49, places the amide group of Gly 50 in a suitable position to form an intermolecular hydrogen bond. The significance of this bond can be seen by the total loss of binding that occurs when the hydroxyl group of Ser 294 is replaced by a proton through a serine-to-alanine mutation (Table II).

The other polar interaction engages Glu 283, the amino-terminal residue of the UIM helix, whose side chain carboxyl group can be in a position to make an electrostatic interaction with the side chain amino group of Lys 45 of UbL (Figure 3b). Alanine substitution of Glu 283 moderately decreases binding, suggesting that this charge complementarity is important for binding (Table II).

Comparison of UIM binding sites between UbL and ubiquitin

Interestingly, all of the major UIM contact sites in UbL are conserved in ubiquitin. Residues Leu 8, Ile 44, Val 70 and His 68 of ubiquitin create a hydrophobic surface that resembles the hydrophobic contact site on UbL (Figure 3d). Previous mutational analyses have shown that, except for His 68, these ubiquitin residues are critical for proteasomal degradation and for binding to S5a, and are essential for life in yeast (28, 41, 42).

The polar contact sites in UbL are also conserved in ubiquitin. The 45Phe-Ala-Gly47 segment of ubiquitin adopts a consecutive type IV (residues 44-47) and type I (residues 45-48) \( \beta \)-turns identical to that seen in
the 48Tyr-Ala-Gly50 segment of UbL with an r.m.s.d. of 0.12 Å across the main chain heavy atoms. Thus, the amide group of Gly 47 of ubiquitin is in the same position as Gly 50 of UbL, which forms a key hydrogen bond with Ser 294 of UIM in the complex.

Finally, the positively charged side chain of Arg 42 in ubiquitin is located in the same position as that of Lys 45 of UbL, and thus possibly mediates an electrostatic interaction with the conserved acidic residue at position 283 of UIM (Figure 3d). Arg 42 of ubiquitin has been shown to be essential for yeast viability (42).

The striking conservation of UIM-binding sites between ubiquitin and UbL suggests that the contact sites and mode of ubiquitin binding to UIM are similar to those observed in our structure. We confirmed that the effects of mutations in UIM are similar for UIM binding to UbL and UIM binding to tetraubiquitin: in other words, mutating Glu 283, Ala 290 or Ser 294 in UIM impairs its binding to tetraubiquitin (Figure 4). The importance of Ser 294 of UIM was also shown by a previous mutagenesis study (10).

The conserved UIM binding site of UbL and ubiquitin

These shared features of the UIM-binding mode of UbL and ubiquitin raise the question of why UIM binds only weakly to monoubiquitin. The dissociation constant ($K_d$) defining the interaction between S5a UIM and monoubiquitin is 273 μM at pH 7 (Figure 5b), and that between Hrs and monoubiquitin is 230-300 μM (16, 17). A detailed comparison of the surface of UbL and ubiquitin shows that the largest difference is a protrusion on ubiquitin caused by the bulky imidazole ring of His 68 (Figure 3d). This protrusion from the otherwise smooth ubiquitin surface may interfere with UIM binding, as can be seen in a structure of the S5a UIM-ubiquitin complex modelled by best-fit superposition of the coordinates of ubiquitin to those of UbL in the lowest-energy structure of the UIM-UbL complex: the imidazole of His 68 encounters steric hindrance with the bound UIM (Figure 3d). Consistent with this,
substitution of the UbL residue at the same position, Val 71, with histidine impairs UbL’s ability to bind UIM (Table II).

To test whether the inability of monoubiquitin to bind S5a UIM is indeed due to steric hindrance, we replaced His 68 of ubiquitin with valine, the amino acid that is found in this position in UbL, and examined its effect on binding to UIM. This mutation increased the binding affinity, suggesting that the protrusion caused by His 68 may inhibit ubiquitin’s access to UIM to some extent (Figure 5a). This result suggests that the UIM-binding mode of ubiquitin is similar to that of UbL, consistent with the mutational data.

To determine whether His 68 regulates ubiquitin binding, we examined the effect of protonation of its imidazole ring on the ability of ubiquitin to bind to UIM. The binding affinity correlated well with the pH-dependent protonation state of the His 68 imidazole ring, as monitored by $H^\delta_2$ and $H^\varepsilon_1$ chemical shifts (Figure 5b). By contrast, this pH dependency was not observed for binding of the H68V mutant. Therefore, the His 68 side chain probably may have a regulatory role in the binding of ubiquitin to UIM.

We then examined whether this histidine of ubiquitin might regulate the binding of ubiquitin to UBA, another ubiquitin-binding motif that has been found in 79 human proteins (6, 9, 43). For this, we measured the binding affinity between ubiquitin and UBA from yeast Dsk2p at various pH values. It displayed a similar pH dependency to that between ubiquitin and UIM, suggesting that His 68 of ubiquitin also regulates the binding of UBA to ubiquitin (Figure 5c). This is consistent with recent structure determination of a CUE-ubiquitin complex (44). The CUE domain shows structure similarity with UBA domains, and shares a common binding site on ubiquitin with UIM.

**DISCUSSION**
UIM interface of UbL

The UIM interface of UbL determined by our structure determination and mutagenesis studies is consistent with results from previous mutagenesis and chemical shift perturbation experiments of ubiquitin and UbLs (10, 16, 18, 19, 28, 29). All of those experiments suggest that the conserved hydrophobic patch composed of the side chains of Leu 8, Ile 44, and Val 70 of ubiquitin is important for UIM or proteasome binding. The UIM-UbL complex reveals that the hydrophobic patch mainly functions as a pocket for the methyl group of Ala 290, which is conserved within UIM sequences (Figure 1b; (7)).

In addition to the contact mediated by the conserved hydrophobic patch, our complex structure shows that two polar interactions mediated by Glu 283 and Ser 294 of UIM contribute to the UIM-UbL association. While the hydrogen bond through Ser 294 is crucial for the binding, the electrostatic interaction between the side chains of Glu 283 of UIM and Lys 45 of UbL seems to be less important, because the substitution of Glu 283 of UIM by alanine caused only a moderate effect in the binding of UIM to either UbL of HR23B or tetraubiquitin (Table II and Figure 4). Notably, glutamic acid is well conserved at this position in UIM sequences that have been either shown or implied to bind a ubiquitin tag, but not in the N-terminal UIM of human S5a or yeast S5a (Figure 1b). Therefore, the weak interaction between human HR23B and the N-terminal UIM, and that between yeast Rad23 and S5a, may be attributed to the absence of glutamic acid at this position in their UIM sequences. Although the N-terminal UIM of human S5a binds to polyubiquitin chains much more weakly than does the C-terminal UIM, it still binds (10). This observation suggests that the electrostatic interaction mediated by Glu 283 of UIM makes a non-essential contribution to the UIM-ubiquitin interaction.

The contact area of the UIM-UbL complex is 474 Å², as defined by the calculated change in the solvent-accessible surface area of UbL upon UIM binding. The area is comparable to the
average values of 517 ± 83 Å² and 617 ± 66 Å² calculated for peptide complexes of ten SH3 domains and nine SH2 domains, respectively. In the previous chemical shift perturbation experiments of the UbL domains from PLIC-2 and HR23B, NMR signals of residues close to but outside (as well as at) the interface of our UIM-HR23B complex structure also exhibited substantial changes upon binding to S5a (18 and 30). For example, the signals of Lys 79, Lys 82, Ile 102, and Lys 103 of PLIC-2 exhibited substantial chemical shift perturbations upon S5a binding (18). However, the corresponding residues of the UbL domain of HR23B, Lys 51, Asn 54, Val 74, and Thr 75, make no direct contribution to the UIM interface of HR23B UbL. The non-involvement of Thr 75 in UIM binding was confirmed by mutagenesis: its substitution by alanine causes no apparent change in the affinity of UbL for UIM (Table II). We assume that the NMR signals of PLIC-2 residues in the proximity of the interface were probably affected by local conformational changes or long-range electrostatic effects on S5a binding. Such effects are often seen in chemical shift perturbation experiments (45), and were observed in the interaction between the UbL of HR23B and UIM (30). The UIM-UbL interaction mode proposed by Ryu et al. based on chemical shift perturbation experiments (30) is different from that observed in our complex structure. They assumed that residues from Glu 283 to Gln 296 or Gly 297 adopts a helix, based on mainchain chemical shift values of the UbL-bound UIM, that Glu 284 and Glu 285 of this helix locate near Lys 48 and Lys 51 of UbL, and that Leu 295 is positioned near the hydrophobic patch including Leu 8 of UbL. These interactions would place the helix of UIM on the UbL surface in an opposite orientation to that observed in our structure.

**Role of His 68 of ubiquitin in UIM binding**

Our data suggest that the protonation of His 68 of ubiquitin elicits a pH-dependent interference of the access of ubiquitin to UIM and UBA. In yeast, the substitution of this residue with alanine has been shown to lead to, albeit weak, sensitivity to cold in yeast growth (42).
The in vivo function of the observed pH dependency of UIM and UBA binding of ubiquitin mediated by His 68 is unclear. Ubiquitin tags serve as a sorting signal for vesicular transport via endosomes, the trans-Golgi network and multivesicular bodies (4, 46). The pH-sensitive change in their binding to downstream effectors or modulators is reminiscent of cargo adaptors that are involved in vesicular transport, such as the KDEL and mannose 6-phosphate receptors, which bind and dissociate cargo molecules in a pH-dependent manner and thereby deliver them unidirectionally between cellular compartments at various luminal pH values (47). However, there is no evidence that ubiquitin tags move into the acidic luminal space of vesicles. Alternatively, it may be possible that acidic membrane components affect the protonation state of this histidine in monoubiquitin tags and thus mediate their binding to UIMs of endocytic factors. Many endocytic factors with a UIM also contain ENTH, VHS or FYVE domains that are located near the UIMs in terms of primary structure (4, 46). These domains bind to acidic lipids, such as phosphoinositides and phosphatidic acid, and thereby bring the adjacent UIMs into proximity with negatively charged membrane components, which could create a low pH environment near the membrane. Proteins, such as bacterial colicin A and pheromone-binding protein from the silk moth, are suggested to undergo conformational transitions at or near negatively charged lipid interfaces (48, 49, 50). These observations raise the intriguing possibility that His 68 of ubiquitin may serve as a pH-sensor to regulate the access of ubiquitin to UIM and UBA of downstream effectors.

His 68 likely also regulate the conformation of Lys 48-linked polyubiquitin chains. The conformation of diubiquitin switches from an open to closed state as the pH increases from 4.5 to 6.8 (51). The crystal structure of diubiquitin indicates that the hydrophobic patch at the UIM interface also serves as the inter-subunit interface, which includes the side chain of His 68 (52). The same surface has been suggested to function as the inter-subunit interface in the closed conformation in solution (51). Thus, the protein surface including the hydrophobic patch and His 68 functions as the inter-subunit interface of diubiquitin at higher pH, but not
at low pH, as occurs in ubiquitin interactions with both UIM and UBA (51). Our pH-titration experiment showed that the pH dependencies of the $H^{\varepsilon}_1$ and $H^\delta_2$ chemical shifts of both subunits of diubiquitin are similar to that of monoubiquitin, shown in Fig. 5b, and thus the histidine side chains of diubiquitin is mostly protonated at pH 4.5 and mostly deprotonated at pH 6.8 (Footnotes 1). These observations raise the possibility that His 68 is involved in regulating the inter-subunit association, as well as in regulating ubiquitin-UIM and ubiquitin-UBA interactions. Notably, the same surfaces of the distal two ubiquitin subunits of tetraubiquitin function as inter-subunit interfaces (51).

This histidine is conserved in the UbL domains of PLIC-2 and Parkin. A triple point-mutation in the UbL domain of PLIC-2 (I75A, A77S and H99A) abolished its binding to the proteasome (18). The mutated residues correspond to Ile 44, Ala 46 and His 68 of ubiquitin, all of which are thought to be important for UIM binding. Therefore, the detrimental effect of the triple-mutation of PLIC-2 cannot be attributed solely to substitution of the histidine.

**UIM consensus sequence**

The structure of the UIM-UbL complex, along with the binding data from the S5a UIM, UbL and ubiquitin mutants presented here, enables us to revise the consensus sequence for ubiquitin binding by UIM. On the basis of the sequence alignment of S5a and endocytic factors (7), it was previously proposed to be $\phi$-x-x-Ala-x-x-Ser-x-x-Ac, where $\phi$ and Ac denote a large hydrophobic and an acidic residue, respectively. Our data now show that, whereas Glu 283, Ile 287, Ala 290 and Ser 294 of S5a contribute to UbL and ubiquitin recognition, Gly 297, which is located at the position of the C-terminal Ac in the proposed consensus, makes no contact with UbL and is disordered in the structure of the complex.

Residues shown to be important for UbL binding by our data are highly conserved in the UIMs of endocytic factors that have been either shown or implied to bind a ubiquitin tag.
Structure of a UIM-UbL complex

(Figure 1b). The most important residues, alanine and serine at positions 290 and 294 in S5a, respectively, are conserved among these factors, and a hydrophobic residues at position 287 is also well conserved. In contrast, the conservation of the acidic residue at position 283 is lower than those in UIM sequences (7). This residue seems to make a smaller contribution than the others, as revealed by our mutagenesis data (Table II and Figure 4; see ‘UIM interface of UbL’ section). These observations allow us to revise the consensus sequence of UIM to (Ac)-x-x-x-φ-x-x-Ala-x-x-x-Ser-x-x-(Ac). Fisher et al. have also shown that glutamic acid cluster at residues 259-262, and residues, A266 and S270, of Hrs UIM are important for its affinity for ubiquitin (see Figure 1b for the sequence of HRS and (37)). They also showed a mutation of E273 of Hrs UIM caused a decreased affinity, although the importance of the corresponding residue in S5a UIM, G297, is not observed in our structure.

Conclusion

We have reported the structure of a UIM bound to a UbL. Our structural and mutational data indicate that the contact sites in UbL are highly conserved in ubiquitin, but ubiquitin also presents a histidine residue at the interface. Thus, this study provides a structural basis for the interaction between ubiquitin and UIM-containing downstream effectors. Future experiments need to be directed at establishing the functional significance of the observed pH dependency of ubiquitin binding to UIM and UBA.

Note

After submission of this manuscript, the solution structures of the UbL of HR23A in complexed with S5a UIM [Mueller, T. and Feigon, J. (2003) *EMBO Journal.*, 22, 4634-4645] and Vps27 UIM-ubiquitin complex [Swanson, K., Kang, R. S., Stamenova, S. D., Hicke, L., Radhakrishnan, I. (2003) *EMBO Journal.*, 22, 4579-4606] were published. The structure of HR23A UbL-UIM is similar to that of HR23B-UIM shown in this paper. Possibility of the interference of ubiquitin His 68 for UIM binding was also discussed based on the structure. In
contrast, Swanson et al. showed that alanine substitution of His 68 of ubiquitin impairs ubiquitin’s binding to the N-terminal UIM of Vps27.
References

Structure of a UIM-UbL complex


Structure of a UIM-UbL complex


Footnotes

1 T. Tenno, unpublished data

Abbreviations: UIM, ubiquitin-interacting motif; UbL, ubiquitin like; HR23B, human homolog of RAD23 B;

Structure Data deposition: The coordinates of structure have been deposited in the Protein Data Bank. [PDB code 1UEL]

Acknowledgements
We thank A. Ohno for providing the UBA domain of Dsk2p, H. Kobayashi for providing the expression plasmid for the UBA domain of Dsk2p, and K. Iwai for providing the ubiquitin-activating enzyme (E1). We thank K. Tanaka for discussion. This work is supported by grants to M.S. and H.T. from the Japanese Ministry of Education, Science, Sports and Culture.
FIGURE LEGENDS

Figure 1

Sequence alignment of UbL and UIM. a, Sequences of human HR23B UbL and human ubiquitin. b, Sequence alignment of the C-terminal UIM of human proteasome subunit S5a with UIMs of endocytic factors that have been either shown or implied to bind a ubiquitin tag (upper), and of the N-terminal UIMs of human and Saccharomyces cerevisiae S5a (Rpn10p) (lower). The sequences are from human (hs), mouse (mm), Drosophila melanogaster (dm) and Saccharomyces cerevisiae (sc). In a and b, identical residues are highlighted in black, and homologous residues are highlighted in grey. Secondary structural elements of HR23B UbL and S5a UIM are indicated. Residues shown to be important for the complex formation through structural or mutational analyses are marked with asterisks.

Figure 2

Structure determination of the UIM-UbL complex. a, Selected region of three-dimensional 13C-filtered-13C-edited NOESY spectra of a complex formed from 13C/15N-labelled UbL and unlabelled UIM with a mixing time of 100 ms, depicting intermolecular NOEs. b, Stereoview of the best-fit superposition of the 20 final structures of the complex between UIM (black) and UbL (blue). Residues 263-277 and 297-307 of UIM are omitted from all figures for clarity.

Figure 3

Structure of the UIM-UbL complex. a, Ribbon diagram of the lowest-energy structure. UIM and UbL are colored orange and light blue, respectively. Secondary structural elements of
UbL are indicated. The 48Tyr-Ala-Gly50 segment of UbL, which forms βT conformation, is shown in green. b, Surface representation of the binding sites of UbL bound to UIM. Hydrophobic and charged residues are shown in yellow and red, respectively, and the βT segment is shown in green. The side chains of UIM residues that interact with UbL are also shown. c, Schematic diagram of the contacts between UIM and UbL. The main chain of UIM is shown in red. The side chains of UIM and UbL residues that form the interface are shown in black and green, respectively. Hydrophobic contacts are indicated by blue dotted lines, and hydrogen bonds and charged interactions are indicated by red broken lines. Me denotes a methyl group. d, Conserved UIM-binding sites in ubiquitin, deduced from the structure of the UIM-UbL complex, are shown on a model of the complex between human ubiquitin and UIM. Hydrophobic and charged residues are coloured yellow and red, respectively, on the surface representation of human ubiquitin. The βT segment, 45Phe-Ala-Gly47, is shown in green. His 68, which causes the protrusion, is shown in blue. The model was constructed by best-fit superposition of the coordinates of human ubiquitin (PDB code: 1d3z) to those of UbL in the lowest-energy structure of the UIM-UbL complex.

**Figure 4**

The effect of mutations on the affinity of S5a UIM for tetraubiquitin examined by a surface plasmon resonance-based competition assay. Resonance curves were measured for 0.1 µM tetraubiquitin binding to immobilized GST-UIM in the presence or absence of 40 µM His-tagged wild-type UIM or its point mutants are shown. The graph indicates that the mutants of UIM, S294A, A290S, E283A, L278S, have weaker binding affinities for tetraubiquitin than the wild-type UIM does.
Figure 5
The UIM-ubiquitin interaction. a, Mutation of His 68 to valine (H68V) enhances the binding affinity of ubiquitin to the UIM of S5a, shown by GST-pull down assay. Bound proteins were analysed by GSH-mediated pull-down assay coupled with SDS-polyacrylamide gel electrophoresis and silver staining. The input lane represents 10% of the wild-type ubiquitin used in the experiment. b, Correlation between the affinity of ubiquitin for UIM and the protonation of ubiquitin His 68. Shown is the pH dependency of the interaction between UIM and wild-type or H68V ubiquitin, and the $^1$H$\varepsilon$1 chemical shift of His 68 (green). The pKa value of His 68, obtained by fitting the H$\varepsilon$1 and H$\delta$2 chemical shift data at pH range from 4.1 to 8.2, is 5.5. c, The pH dependency of interaction between UBA from yeast Dsk2p and ubiquitin.
Structure of a UIM-UbL complex

Tables
Table I
Experimental restraints and structural statistics of final 20 structures

<table>
<thead>
<tr>
<th>Total number of distance restraints</th>
<th>1630</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intraresidue (</td>
<td>i-j</td>
</tr>
<tr>
<td>Sequential (</td>
<td>i-j</td>
</tr>
<tr>
<td>Medium range (1&lt;</td>
<td>i-j</td>
</tr>
<tr>
<td>Intramolecular long range (</td>
<td>i-j</td>
</tr>
<tr>
<td>Intermolecular</td>
<td>103</td>
</tr>
<tr>
<td>Number of hydrogen bonds</td>
<td>25</td>
</tr>
<tr>
<td>Number of dihedral angle restraints</td>
<td>50</td>
</tr>
</tbody>
</table>

Energies§
- Mean AMBER energy (kcal/mol) -5172 ± 14
- Mean van der Waals energy (kcal/mol) -937 ± 12
- Mean constraint energy (kcal/mol) 18 ± 1

Mean deviations from ideal geometry
- Bond lengths (Å) 0.0100 ± 0.0001
- Bond angles (°) 2.31 ± 0.01

Violation statistics
- Number of distance violation > 0.2 Å 0.6 ± 0.5
- Maximum distance violation (Å) 0.27
- Number of dihedral angle violation > 2 ° 0.3 ± 0.4
- Maximum dihedral angle violation (°) 3.64

r.m.s.d. from the average structure*
- Backbone (Å) 0.34 ± 0.06
- Heavy atoms (Å) 0.78 ± 0.06

PROCHECK statistics*
- Residues in most favoured regions (%) 91.4
- Residues in allowed regions (%) 7.9
- Residues in generously allowed regions (%) 0.6
- Residues in disallowed regions (%) 0.1

§ Generalized Born model of AMBER 7 was used
* Residues 1-75 of HR23B and 278-296 of S5a were used
Structure of a UIM-UbL complex

Table II
Dissociation constants of S5a UIM binding to wild-type and mutant HR23B UbL

<table>
<thead>
<tr>
<th>HR23B UbL</th>
<th>$K_d$ (µM)</th>
<th>Fold decrease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>3.4 ± 0.3</td>
<td></td>
</tr>
<tr>
<td>L8A</td>
<td>&gt;1,000</td>
<td>&gt;300</td>
</tr>
<tr>
<td>I47A</td>
<td>16.6 ± 3.1</td>
<td>4.9</td>
</tr>
<tr>
<td>V71H</td>
<td>&gt;1,000</td>
<td>&gt;300</td>
</tr>
<tr>
<td>M73V</td>
<td>1.8 ± 0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>T75A</td>
<td>5.1 ± 0.7</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Dissociation constants of HR23B UbL binding to wild-type and mutant S5a UIM

<table>
<thead>
<tr>
<th>S5a UIM</th>
<th>$K_d$ (µM)</th>
<th>Fold decrease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>3.4 ± 0.3</td>
<td></td>
</tr>
<tr>
<td>L278S</td>
<td>15.8 ± 0.5</td>
<td>4.6</td>
</tr>
<tr>
<td>E283A</td>
<td>12.1 ± 1.0</td>
<td>3.6</td>
</tr>
<tr>
<td>A290S</td>
<td>109 ± 14</td>
<td>32</td>
</tr>
<tr>
<td>S294A</td>
<td>213 ± 24</td>
<td>63</td>
</tr>
</tbody>
</table>

All data are the average and estimated deviation of at least two independent measurements.
Figure 1
Figure 2
Figure 3
Figure 4
Figure 5