



Use of the LC3B-fusion technique for biochemical and structural studies of proteins involved in the N-degron pathway

Received for publication, September 2, 2019, and in revised form, December 19, 2019. Published, Papers in Press, January 9, 2020, DOI 10.1074/jbc.RA119.010912

Leehyeon Kim (김이현)^{1,2}, Do Hoon Kwon (권도훈)¹, Jiwon Heo (허지원), Mi Rae Park (박미래), and Hyun Kyu Song (송현규)³

From the Department of Life Sciences, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, South Korea

Edited by George N. DeMartino

The N-degron pathway, formerly the N-end rule pathway, is a protein degradation process that determines the half-life of proteins based on their N-terminal residues. In contrast to the well-established *in vivo* studies over decades, *in vitro* studies of this pathway, including biochemical characterization and high-resolution structures, are relatively limited. In this study, we have developed a unique fusion technique using microtubule-associated protein 1A/1B light chain 3B, a key marker protein of autophagy, to tag the N terminus of the proteins involved in the N-degron pathway, which enables high yield of homogeneous target proteins with variable N-terminal residues for diverse biochemical studies including enzymatic and binding assays and substrate identification. Intriguingly, crystallization showed a markedly enhanced probability, even for the N-degron complexes. To validate our results, we determined the structures of select proteins in the N-degron pathway and compared them with the Protein Data Bank–deposited proteins. Furthermore, several biochemical applications of this technique were introduced. Therefore, this technique can be used as a general tool for the *in vitro* study of the N-degron pathway.

The N-degron pathway (1), formerly the N-end rule pathway, governs the regulation of the *in vivo* half-life of a protein based on its N-terminal residues (2–4). It is a universal ubiquitin (Ub)-dependent⁴ proteolytic system that exists in both prokaryotes and eukaryotes (5–7). The N-terminal degradation signal, called N-degron, is recognized by the E3 Ub ligase

N-recognin, which targets the substrates bearing N-degron for proteasome-mediated degradation. N-degrons are generated by various cellular processes including proteolytic cleavages, hierarchical enzymatic steps, and post-translational modifications. In eukaryotes, four N-degron pathways have been reported to date: the Arg/N-degron pathway (2, 7), the Ac/N-degron pathway (8, 9), the Pro/N-degron pathway (10), and, very recently, the formyl-Met/N-degron pathway (11). Most of the key molecules in these pathways have been identified and studied at the cellular level, especially using a yeast system (8, 10, 11). Numerous N-degron substrates have been reported, and the associations between the N-degron pathway and human diseases have received great attention (12–14). Moreover, it was recently reported that cross-talk occurs between the N-degron and autophagy degradation pathways (15–17).

In vitro biochemical and structural studies of proteins in the N-degron pathway have been relatively limited (16, 18–21). Large quantities of proteins are required for quantitative analysis of the binding affinity and kinetic parameters of enzymes and for high-resolution structural determination. Additionally, N-degron pathways, especially the well-characterized Arg/N-degron pathway, require several modifying enzymes to generate different N-degrons. Several molecules are available for coordinating such modifications in the cells, but it is difficult to prepare all necessary enzymes to truly mimic the cellular environment in the *in vitro* system. For clear understanding of the recognition of N-degron by N-recognin or other proteins involved in the N-degron pathway, high-resolution structures of such complexes are required. However, the binding affinity between the target protein and N-degron cannot be very high because N-degron has to be released prior to degradation. It is therefore difficult to obtain crystals in the N-degron complex state.

To overcome the above limitations, we developed a new biochemical technique that uses a microtubule-associated protein 1A/1B light chain 3B (LC3B)–fusion tag to generate the desired N-terminal residues in target molecules. LC3B is a Ub structural homolog and is known as an autophagy marker. The LC3B precursor contains several additional residues at the C terminus, which are cleaved by the ATG4B protease to expose C-terminal glycine residue. The mature form of LC3B consists of 120 amino acid residues including a C-terminal glycine, which is covalently modified by phosphatidylethanolamine during autophagy. Using this LC3B-fusion technique, we were able to overexpress and purify the proteins possessing various N-terminal residues,

This work was supported in part by a Korean University Grant and by Samsung Science & Technology Foundation Grant SSTF-BA 1701-14. The authors declare that they have no conflicts of interest with the contents of this article.

This article contains Tables S1 and S2 and Figs. S1–S3.

The atomic coordinates and structure factors (codes 6HLN, 6KGI, 6KGJ, and 6KHZ) have been deposited in the Protein Data Bank (<http://www.pdb.org/>).

¹ These authors contributed equally to this work.

² Recipient of a POSCO Science Fellowship from the POSCO T.J. Park Foundation.

³ To whom correspondence should be addressed: Dept. of Life Sciences, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, South Korea. Tel.: 82-2-3290-3457; E-mail: hksong@korea.ac.kr.

⁴ The abbreviations used are: Ub, ubiquitin; AtPRT6, PRT6 from *A. thaliana*; DUB, deubiquitylating enzymes; LC3B, microtubule-associated protein 1A/1B light chain 3B; MR, molecular replacement; NTAQ1, N-terminal glutamine amidohydrolase; PRT6, proteolysis 6; SAD, single-wavelength anomalous dispersion; SEC-MALS, size-exclusion chromatography column coupled with a multiangle light scattering device; PDB, Protein Data Bank; L-GFP, leucylated GFP; TCEP, tris[2-carboxyethyl]phosphine.

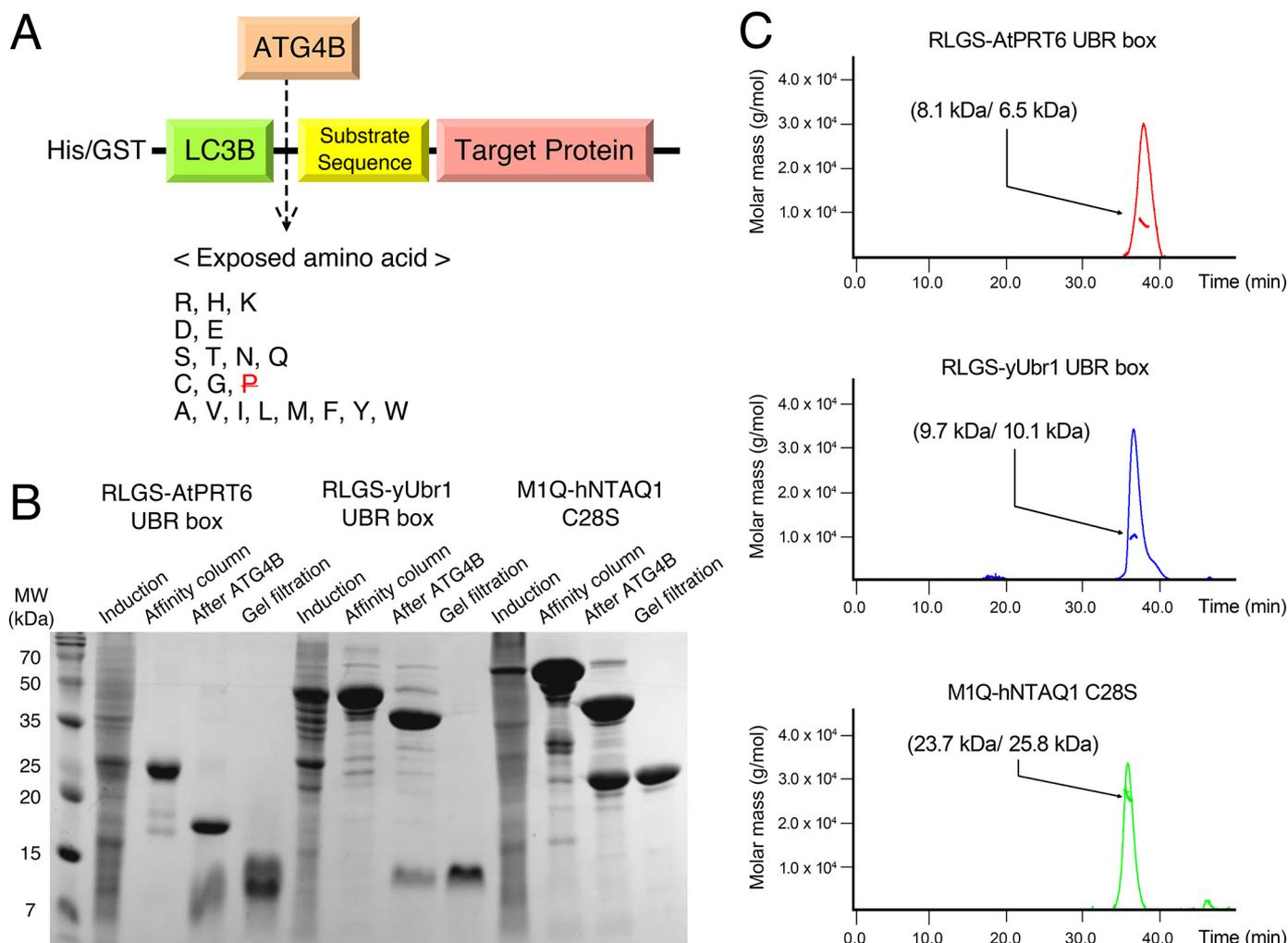


Figure 1. Overall concept and sample preparation of the proteins involved in the N-degron pathway. A, a schematic drawing of the method. ATG4B cleaves the peptide bond between the C terminus of LC3B and substrate sequences except for at a proline residue. B, overproduction and purification of the N-degron pathway proteins (RLGS-AtPRT6 UBR box, His affinity; RLGS-yUbr1 UBR box, GST affinity; and M1Q-hNTAQ1 C28S, GST affinity). C, the oligomeric states were analyzed by SEC-MALS. The horizontal line represents the measured molecular mass, which is indicated by an arrow with the theoretically calculated (Calc) and experimental (MALS) molecular mass values in parentheses (Calc/MALS). MW, molecular mass.

including the specifically modified target molecules involved in the Arg/N-degron pathway. This system therefore provides great advantages for the biochemical and structural studies of proteins in the N-degron pathway. In particular, the attachment of N-degron residues to the target molecules using this technique could be a standard approach for studying the structures of protein complexes.

Results

Generation of proteins bearing various N-terminal residues

The LC3B tag was used for the expression of recombinant proteins as fusion proteins and the ATG4B protease was used to cleave the peptide bond present after the C-terminal glycine of LC3B in a promiscuous manner (Fig. 1A). We took advantage of this feature and created a vector that allows for the desired amino acid to be exposed after Gly-120 of LC3B (Fig. S1) and to facilitate purification, hexahistidine residue, or GST protein was added to the N terminus of the recombinant proteins. All potentially exposed substrate amino acids could be used except for proline. The original concept for this expression system came from the crystallization strategy used for LC3B-substrate

chimeras at either the N or C terminus of LC3B, which is widely used for studying the protein complexes (22, 23). Our group has had success using this strategy for the biochemical and structural studies of human p62/SQSTM1, a major selective autophagy adaptor, using a chimera of the p62 ZZ domain with N-degron residues (16); we are now attempting to make this technique a general research tool for the study of the N-degron pathway. Toward this goal, multiple test cases have been performed to study multiple proteins from different kingdoms. One of these was the E3 Ub ligase Ubr1, which was first identified in yeast and is the most well-characterized N-recognin of the N-degron pathway (24), and recognizes type 1 N-degrons through its UBR box (18). The second one was the Ub protein E3 ligase proteolysis 6 (PRT6), which is involved in the Arg/N-degron pathway of plants. PRT6 also recognizes basic N-terminal residues with its UBR box (25). Finally, a test case was performed with human N-terminal glutamine amidohydrolase (NTAQ1), which converts N-terminal glutamine to glutamate by removing the amine group and catalyzes the first step of the Arg/N-degron pathway (20).

A tool for *in vitro* study of the N-degron pathway

Table 1

Summary of structures in complex with the N-terminal segment from a symmetry-equivalent molecule

Protein	N-terminal residue	Origin	Space group	PDB code	Remarks
SQSTM1/p62	Arg	<i>H. sapiens</i>	$P 2_1 2_1 2_1$	5YP8	Ref. 16, type 1
SQSTM1/p62	Lys	<i>H. sapiens</i>	$P 2_1 2_1 2_1$	5YPA	Ref. 16, type 1
SQSTM1/p62	His	<i>H. sapiens</i>	$I 2 3$	5YPB	Ref. 16, type 1
SQSTM1/p62	Tyr	<i>H. sapiens</i>	$I 2 3$	5YPE	Ref. 16, type 2
SQSTM1/p62	Trp	<i>H. sapiens</i>	$I 2 3$	5YPF	Ref. 16, type 2
SQSTM1/p62	Phe	<i>H. sapiens</i>	$C 1 2 1$	5YPC	Ref. 16, type 2
SQSTM1/p62	Leu	<i>H. sapiens</i>	$P 2_1 2_1 2_1$	5YPG	Ref. 16, type 2
SQSTM1/p62	Ile	<i>H. sapiens</i>	$P 1 2_1 1$	5YPH	Ref. 16, type 2
SQSTM1/p62	Gly	<i>H. sapiens</i>	$I 2 3$	6KHZ	This study
PRT6	Arg	<i>A. thaliana</i>	$P 4_3 3 2$	6LHN	This study, type 1
Ubr1	Gly	<i>S. cerevisiae</i>	$P 3_2$	3NIS	Ref. 18
Ubr1	Arg	<i>S. cerevisiae</i>	$P 4_3 2_1 2$	6KGI	This study, type 1
UBR1	Gly	<i>H. sapiens</i>	$P 1 2_1 1$	3NY1	Ref. 28
ClpS	Leu	<i>E. coli</i>	$C 1 2 1$	3O2O	Ref. 52, type 2
NTAQ1	Ser	<i>H. sapiens</i>	$P 2_1 2_1 2_1$	4W79	Ref. 20
NTAQ1	Gln	<i>H. sapiens</i>	$I 2 2 2$	6KGJ	This study

We applied the LC3B fusion technique and successfully isolated the proteins listed in Table 1. Based on the sequence of the native p62 substrate BiP/GRP78 (Arg-Glu-Glu-Glu-Asp—), the peptide sequence for Xaa-Glu-Glu-Glu-Asp (where Xaa indicates variable amino acid residues) was attached to the N terminus of the ZZ domain of p62 (16). We generated nine different N-terminal residues (basic type 1: Arg, Lys, and His; hydrophobic type 2: Leu, Phe, Trp, Tyr, and Ile; and test case: Gly) for exposure to the ZZ domain of p62. The UBR boxes of *Arabidopsis thaliana* PRT6 (AtPRT6) and *Saccharomyces cerevisiae* Ubr1 (yUbr1) contained an Arg-Leu-Gly-Ser (RLGS) sequence at their N termini and the first methionine of *Homo sapiens* NTAQ1 (hNTAQ1) was changed to glutamine to produce a recombinant protein. They were overexpressed in *Escherichia coli* as soluble proteins, then loaded onto affinity and ion-exchange columns, and detected as abundant bands on SDS gels stained with Coomassie Blue (Fig. 1B and Fig. S2A). Each purified LC3B-fusion protein was treated with ATG4B at a molar ratio of 1:100 to achieve complete cleavage. After LC3B was excised, the target protein with the desired N-terminal amino acid was loaded onto a size-exclusion chromatography column coupled with a multiangle light scattering device (SEC-MALS) (Fig. 1C and Fig. S2B) to confirm the monomeric state of proteins in solution. The N terminus of each purified protein was verified by N-terminal sequencing (Table S1).

Comparison with similar tagging methods

The exposure of the desired N-terminal residues is also facilitated by Ub and SUMO tags, which are also restricted to making prolyl N termini (26, 27). We compared our new method with these existing tags in several ways as an AtPRT6 target. First, we cloned the RLGS-AtPRT6 UBR box into His-LC3B, His-Ub, and His-SUMO vectors and then examined the bacterial expression. As shown in Fig. 2A, all three fusion proteins were soluble and overexpressed. Interestingly, the single affinity purification using the His-nitrilotriacetic acid column was quite efficient for all three cases; however, the LC3B fusion showed the best results. Usually, the next step in the purification process is the cleavage reaction for removing the tags; therefore, obtaining pure sample is important for reducing unwanted cleavage by potential contaminant proteases in *E. coli*. Second, we compared the expression of full-length pro-

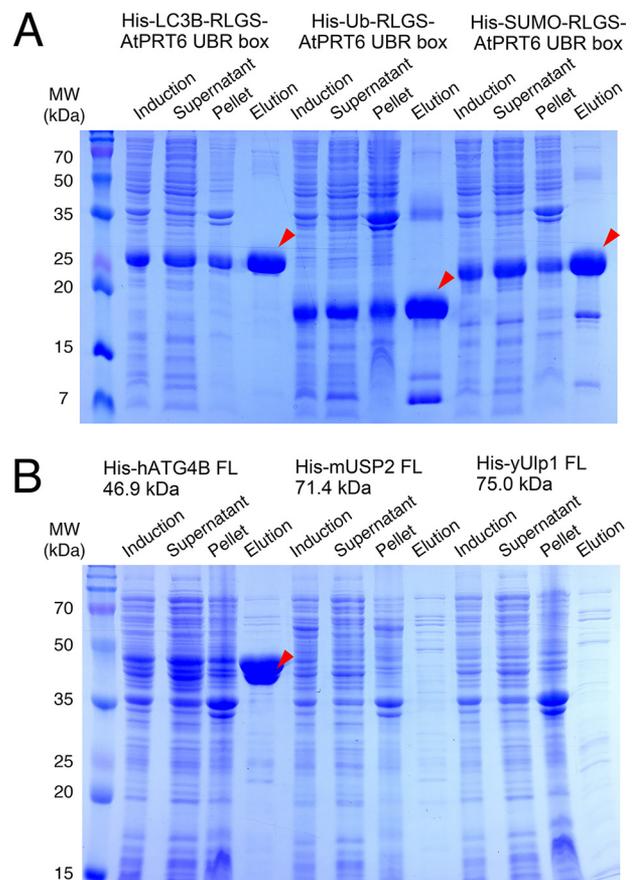


Figure 2. Comparison of LC3B tag with Ub and SUMO tags. A, SDS-PAGE results of the expression and nickel-nitrilotriacetic acid affinity purification of His-LC3B/His-Ub/His-SUMO-tagged RLGS-AtPRT6 UBR box. The target bands are indicated with red arrows. B, SDS-PAGE results of the expression and nickel-nitrilotriacetic acid affinity purification of full-length human ATG4B, mouse USP2, and yeast Ulp1 proteases. Note that the band of His-ATG4B is indicated with a red arrow, whereas those of His-mUSP2 and His-Ulp1 are not clear. MW, molecular mass.

teases for cleaving LC3B, Ub, and SUMO tags. The human ATG4B protease was well-expressed in *E. coli* and was obtained with high purity after single affinity chromatographic purification step (Fig. 2B). However, mouse USP2 and yeast Ulp1 proteases showed lower expression and were not easily purified by the single chromatographic step (Fig. 2B). Instead of these full-length proteases, engineered proteases for Ub and SUMO

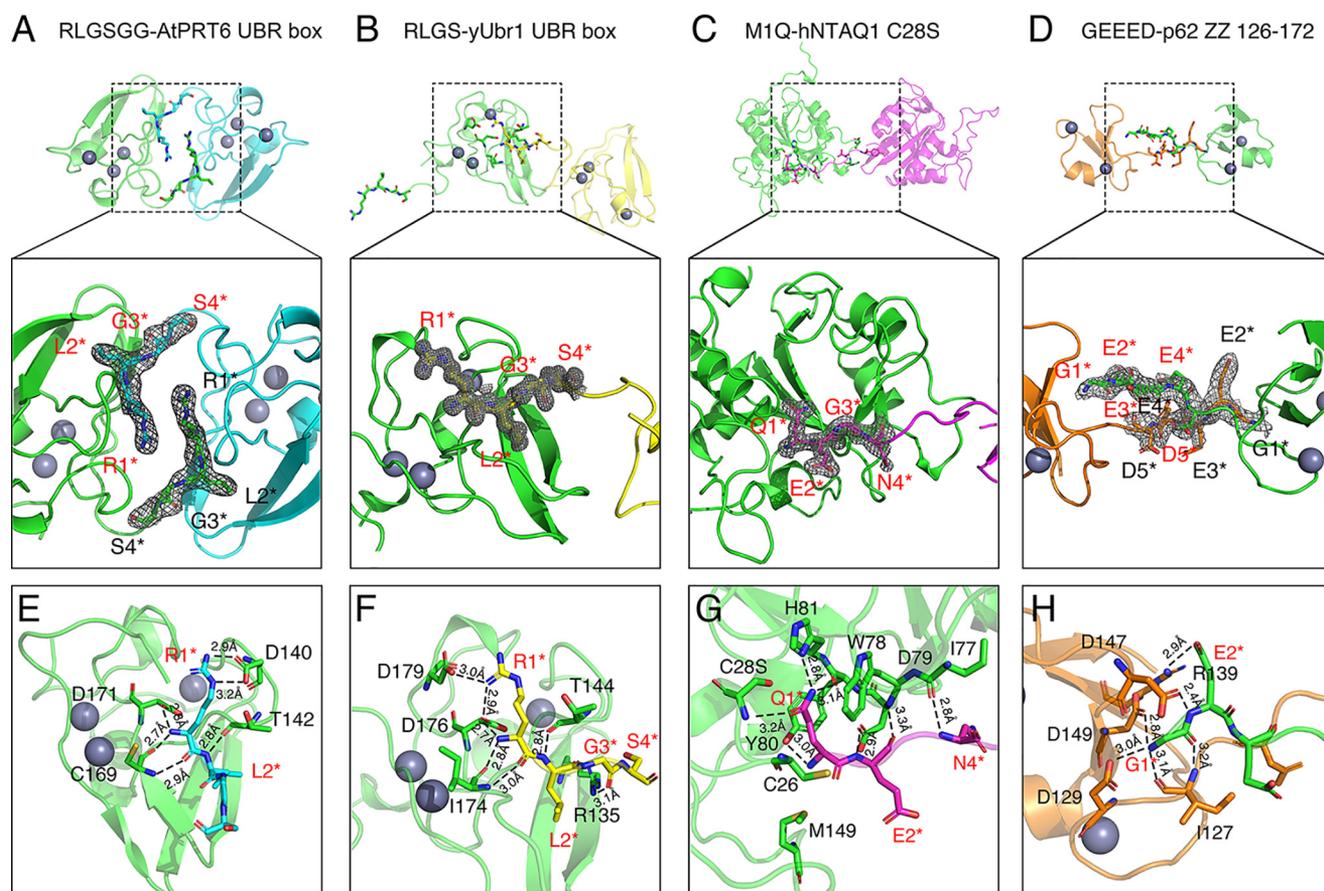


Figure 3. Structures of exemplary proteins. A–D, overall structures of RLGSGG–AtPRT6 UBR box (A), RLGS–yUbr1 UBR box (B), M1Q–hNTAQ1 C28S (C), and GEEED–p62 ZZ (D) with the neighboring molecules in the crystalline lattice. E–H, close-up views of boxed regions in A–D, showing the electron density maps of N-degron segments from the symmetry-equivalent molecules; recognition site of N-degron residues by RLGSGG–AtPRT6 UBR box (E), RLGS–yUbr1 UBR box (F), and M1Q–hNTAQ1 C28S (G), and GEEED–p62 ZZ (H). Interactions between the proteins and N-degron substrates are shown as dotted lines, and the distances are indicated. N-degron residues are labeled (from number 1) with asterisks and colored red for clarity.

cleavage are commercially available (see “Experimental procedures”). We performed time-course cleavage experiments (Fig. S3). All three proteases cleaved the target substrates well. Although, the optimized commercially available proteases cleaved the peptide bond faster, the reaction products of ATG4B reaction might be better. However, the target Ub- and SUMO-fusion proteins were partially cleaved without protease treatment (Fig. S3), probably by contaminant *E. coli* proteases.

Improved crystallization probability

Our protein modification and purification method was found to be particularly useful for crystallization, as indicated by the experiments with the ZZ domain of p62 (16). Originally, crystals of the p62 ZZ domain alone suffered from phasing problems caused by translational noncrystallographic symmetry and diffraction anisotropy, which were overcome by crystallization with the protein containing the substrate sequence at N termini (16). In addition to phasing, the presence of the substrate sequence also increased the resolution to 1.6 Å. Furthermore, the structures of protein complexes that were formed with eight different N-degrons were obtained even though the binding affinity between the ZZ domain and the N-degron peptide is very weak (16). By taking advantage of this method, we determined the crystal structures of RLGS–AtPRT6 UBR box,

RLGS–yUbr1 UBR box, M1Q–hNTAQ1, and GEEED–p62 ZZ, as mentioned above (Fig. 3 and Table 2).

Of the four proteins that crystallized, the structure of the AtPRT6 UBR box has not been previously determined; thus, the structure of the RLGS–AtPRT6 UBR box complex was solved using the single-wavelength anomalous dispersion (SAD) method (Table S2), whereas the remaining crystal structures were solved using molecular replacement (MR). As compared with the other UBR box structures (18, 28, 29), the PRT6 UBR box from *Arabidopsis* also showed a heart-shaped domain with three zinc-coordinated sites (Fig. 3A). In addition, as compared with the structure of the complex formed by the yUbr1 UBR box and RLGS peptide (18), both the yUbr1 UBR box and the AtPRT6 box recognized Arg-1* and Leu-2* residues in a similar manner (where the asterisk indicates the N-degron sequence attached to the N terminus) (Fig. 3E). These properties were also confirmed for the RLGS–yUbr1 UBR box structure (Fig. 3, B and F), indicating that the results obtained using our current method are similar to those obtained from the previous complex structure with the substrate peptide (18). Substrate recognition residues are generally conserved among UBR boxes from other organisms (Fig. 4). In contrast, zinc-coordination residues in the AtPRT6 UBR box are strictly conserved within the human UBR1 UBR box, but not within the yeast Ubr1 (Fig. 4).

Table 2
Data collection and refinement statistics

	RLGSGG–AtPRT6 UBR box (residues 119–189)	RLGS–yUbr1 UBR box (residues 113–194)	M1Q–hNTAQ1 C28S (FL, residues 1–205)	GEEED–p62 ZZ (residues 126–172)
Data collection				
X-ray sources	PAL-5C	PF-17A	PAL-5C	Home source (CuK α , λ = 1.54178 Å)
Space group	$P4_3 3 2$	$P4_3 2_1 2$	$I 2 2 2$	$I 2 3$
Cell dimensions				
a, b, c (Å)	95.4, 95.4, 95.4	60.6, 60.6, 69.8	95.6, 105.8, 140.2	114.0, 114.0, 114.0
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Wavelength (Å)	1.00000	0.90000	0.97957	1.54178
Resolution (Å)	50.00–2.50 (2.54–2.50) ^a	50.00–1.04 (1.06–1.04) ^a	50.00–1.80 (1.84–1.80) ^a	50.00–2.80 (2.85–2.80) ^a
R_{merge}	0.157 (0.941) ^a	0.064 (0.770) ^a	0.080 (0.542) ^a	0.130 (0.860) ^a
$I/\sigma I$	16.6 (1.3) ^a	43.9 (2.1) ^a	25.9 (2.56) ^a	19.8 (2.79) ^a
Completeness (%)	99.2 (98.9) ^a	100.0 (99.9) ^a	100.0 (100.0) ^a	99.9 (100.0) ^a
Redundancy	9.0 (5.2) ^a	12.5 (11.3) ^a	6.7 (6.4) ^a	6.7 (6.8) ^a
Refinement				
Resolution (Å)	38.95–2.50 (2.59–2.50) ^a	30.25–1.04 (1.077–1.04) ^a	35.47–1.80 (1.87–1.80) ^a	40.30–2.80 (2.90–2.80) ^a
No. of reflections	5,514 (528) ^a	62,978 (6,199) ^a	65,202 (6,176) ^a	6,189 (624) ^a
$R_{\text{work}}/R_{\text{free}}$	0.213/0.248	0.189/0.202	0.216/0.218	0.235/0.285
No. of atoms	565	704	3,426	1,409
Protein	562	669	3,270	1,401
Ligand	3	3		8
Water		32	156	
B-factors (Å ²)				
Protein	49.76	16.31	26.70	59.49
Ligand	49.78	16.25	26.68	59.49
Water	45.42	10.26		58.84
Water		18.11	27.02	
Root-mean-square deviations				
Bond lengths (Å)	0.013	0.005	0.031	0.003
Bond angles (°)	1.19	0.77	1.87	0.52
Ramachandran (%)				
Favor	93.24	97.62	97.25	95.58
Allow	5.41	2.38	2.75	4.42
Outliers	1.35	0.00	0.00	0.00
PDB code	6LHN	6KGI	6KGJ	6KHZ

^a Highest resolution shell.

For the hNTAQ1 M1Q mutant, several residues were identified that are involved in recognizing the substrate amino acid Gln* (Fig. 3, C and G). Hydrophilic Ser-28, Tyr-80, and His-81 residues form hydrogen bonds with the substrate Gln N-degron, and hydrophobic Cys-26, Trp-78, and Met-149 residues in hNTAQ1 participate in the interaction (Fig. 3G). In addition to the Gln-1* residue, Glu-2* and Asn-4* were also found to exhibit hydrogen-bonding interactions, as shown in the previously published structure (20). Thus, from these structural data, it was confirmed that the crystals obtained by attaching the substrate sequences to the N termini of given proteins were similar to the crystals obtained through peptide soaking or co-crystallization. This is presumably the result of formation of stable crystal packing by recognizing substrate sequences in the binding pocket.

Applications of current strategy for biochemical study of the N-degron pathway

Our proposed method is applicable not only for determination of N-degron complex structures but also for biochemical characterization of the N-degron pathway *in vitro*. The translational products contain Met residues at their N termini. However, N-degron substrates, especially those in the Arg/N-degron pathway, often possess different N-degron sequences through post-translation modifications, such as proteolytic cleavage or modifications by other members of the Arg/N-degron pathway (2, 30–33). Therefore, using our current technique, N-degron substrates containing diverse amino acid residues found in any steps can be produced.

One such example is the N-terminal arginylation of BiP/GRP78 (hereafter referred to as R-BiP), which is recognized by p62/SQSTM1 (17, 34). The central ZZ domain of p62 appears to play an important role in the recognition of R-BiP. The enzyme ATE1 adds an Arg residue to the modified N terminus of BiP, which functions as a chaperone that binds to the misfolded protein aggregates. Supramolecular complexes between BiP chaperone molecules and protein aggregates are recognized by the p62 oligomer, after which the multiprotein complex is delivered to the autophagosomes where it is ultimately degraded by the lysosomes. The ZZ domain of p62 recognizes only N-terminally arginylated forms of BiP, and thus, it is important to create functional R-BiP for *in vitro* studies. Following LC3B, the arginylated BiP sequence lacking the signal peptide (1–18 amino acids) was cloned to create R-BiP *in vitro*. The affinity between the ZZ domain and R-BiP is typically not strong (16), although they are able to form a tightly bound complex because of additional interactions between Ub and the UBA domain of p62. As a confirmation, gel filtration results showed that only a tiny fraction of R-BiP co-migrates with MBP-PB1-ZZ (Fig. 5A). This tiny fraction of R-BiP protein was weakly detected by Coomassie staining; however, the fraction was clearly detected by a specific R-BiP antibody (Fig. 5A).

Proteins generated using this technique can be used as model proteins for the N-degron pathway studies. One such potential model protein is GFP, which is used for the visualization of protein levels and location, and for the real-time monitoring of degradation by measuring the decrease in the fluorescent intensity. An example of such an application is with leucylated GFP

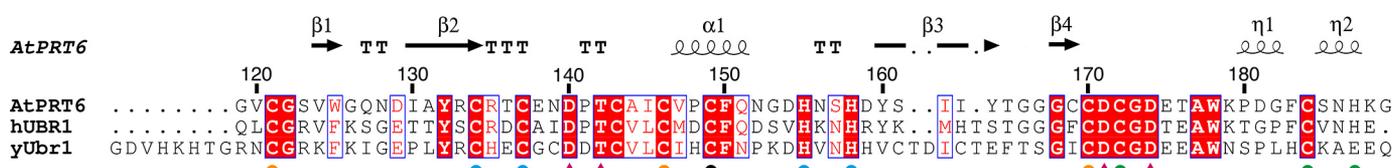


Figure 4. Sequence alignment of the UBR boxes in N-recognins from different organisms. At, *A. thaliana*; h, human; y, yeast. Red shading and empty boxes represent identical and highly conserved residues, respectively. Secondary structure elements and sequence numbers of every 10th residue of the AtPRT6 UBR box are indicated above the sequences. Solid circles below the sequences indicate residues that coordinate zinc ions (orange for Zn1, green for Zn2, sky blue for Zn3, and black for both Zn1 and Zn2). Key determinants for binding of the N-degron are marked with solid magenta triangles below the sequence.

(L-GFP). ClpS plays an important role in the *E. coli* N-end rule pathway (35), in which the N-degron substrates are degraded by the AAA+ chaperone ClpA in complex with the ClpP peptidase (ClpAP) (5, 36–38). As a ClpAP-specific adaptor, ClpS binds strongly to the N-terminal destabilizing residues (especially those of the hydrophobic type 2 N-degrons: Leu, Phe, Trp, and Tyr) and targets the substrate to ClpAP for degradation (35). We prepared a ClpS-conjugated column to identify novel type 2 target proteins in mammalian cells. To confirm whether the ClpS column works properly, L-GFP was generated, purified, and applied to the column (Fig. 5B). When L-GFP flowed through the ClpS column, the column became GFP-colored because of binding with ClpS. As a control experiment, the wt-GFP was applied to the ClpS column, and clearly there was no binding of wt-GFP to the ClpS column. To further confirm this result, we generated the inactive ClpS (D35A/D36A double mutant) column. Both wt- and L-GFP did not bind to the column; therefore, this ClpS column can be used to fish out target proteins from cell extracts.

The first step of hierarchical Arg/N-degron pathway is catalyzed by N-terminal deamidases. Nta1 from yeast catalyzes the conversion of both asparagine and glutamine to aspartate and glutamate, respectively (7, 19). In mammalian systems, NTAQ1 and NTAN1 deamidate Gln and Asn, respectively. We prepared substrate sequence-attached yNta1 and hNTAQ1 enzymes using our current method. The deamidation reaction does not require additional co-factors; thus, enzymatically active yNta1 and hNTAQ1 produced the desired product sequences themselves (Fig. 5C and Table S1). It is interesting to note that the hNTAQ1 C28S mutant possesses residual enzymatic activity, and thus, during the purification procedure, a portion of Gln in the substrate sequence was converted to Glu based on the N-terminal sequencing results (Table S1).

Discussion

The best-known method for exposing desired amino acids is to use a Ub tag (26, 39). Using Ub as a fusion partner has multiple benefits. It enhances yield (40) and can be removed by highly specific proteases known as deubiquitylating enzymes (DUBs) that do not leave additional amino acids at the N termini of proteins of interest (26, 41). However, the Ub fusion method has multiple drawbacks. As shown in Fig. 2B, the full-length USP2 protease was also not well-expressed using *E. coli* in our hand. Most DUBs are quite large and therefore are difficult to purify in large quantities, although methods for modifying DUB enzymes have been developed to address this (42). Additionally, the C-terminal linker of Ub is about 5 amino acids long, and shorter linker lengths lead to lower protease activities.

To get complete cleavage of a fusion protein, a 1:10 molar ratio of DUB to target protein is required (42). Alternatively, the Ub-like protein SUMO can also be used as a tag (43). SUMO-specific proteases are inexpensive to produce (44), and the best-known SUMO protease-1 can be used at a ratio of 1:5,000 to 1:10,000 (27, 45). SUMO fusion proteins also facilitate correct protein folding through the SUMO moiety (46). However, the biggest limitation to the use of SUMO is that occasionally inclusion bodies form, making it difficult to produce properly folded proteins (27). LC3B, the Atg8 mammalian homolog, is a Ub-like modifier involved in the formation of autophagosomal vacuoles or autophagosomes (47, 48). The precursor molecule of LC3B is cleaved by ATG4B (Atg4 in yeast), and then lipidation begins at the last C-terminal amino acid residue. In addition, LC3B can resolve many of the shortcomings of Ub. A large amount of ATG4B protein can be obtained without any particular modifications, and the X-ray structure of LC3B (22) showed that the C-terminal is composed of 12 amino acids, allowing for more efficient protease activity. We have tested three different tags side by side with a target protein, AtPRT6 UBR box (Fig. 2 and Fig. S3). His-tagged versions of LC3B-, Ub-, and SUMO-fusion proteins were all easily purified; however, the purity of the His-LC3B-AtPRT6 UBR box was the best (Fig. 2A). Subsequently, we also compared the expression of tag cleaving proteases, and again, ATG4B was the best (Fig. 2B). However, catalytic domains of Ub and SUMO proteases are commercially available, and they are very efficient in removing tags (Fig. S3, B and C). However, the tags were partially removed by unknown *E. coli* protease(s) without the specific protease treatment (Fig. S3). Therefore, the homogeneity of reaction products of His-LC3B-AtPRT6 UBR box by ATG4B must be better than those of the other tagging systems. In any case, every protein behaves differently when expressed as a fusion protein, and therefore, it would be beneficial to have more tools for fusion-protein overexpression.

Proteins in the N-degron pathway recognize the first specificity-determining residue, nonspecific main-chain atoms, and sometimes additional residues (10, 16, 18, 19, 49). It is thus difficult to identify complex structures because each peptide has a particular desired amino acid (18, 19, 21). It is rational to propose that the presence of any binding partners may help stabilize the flexible nature of an apo structure and, therefore, may be helpful for better crystallization. Because molar ratios are important for co-crystallization, crystallization must be tried at various molar ratios; unfortunately, noncomplex crystals are frequently obtained. During soaking, binding sites are not always accessible through the solvent channels, and some-

A tool for *in vitro* study of the N-degron pathway

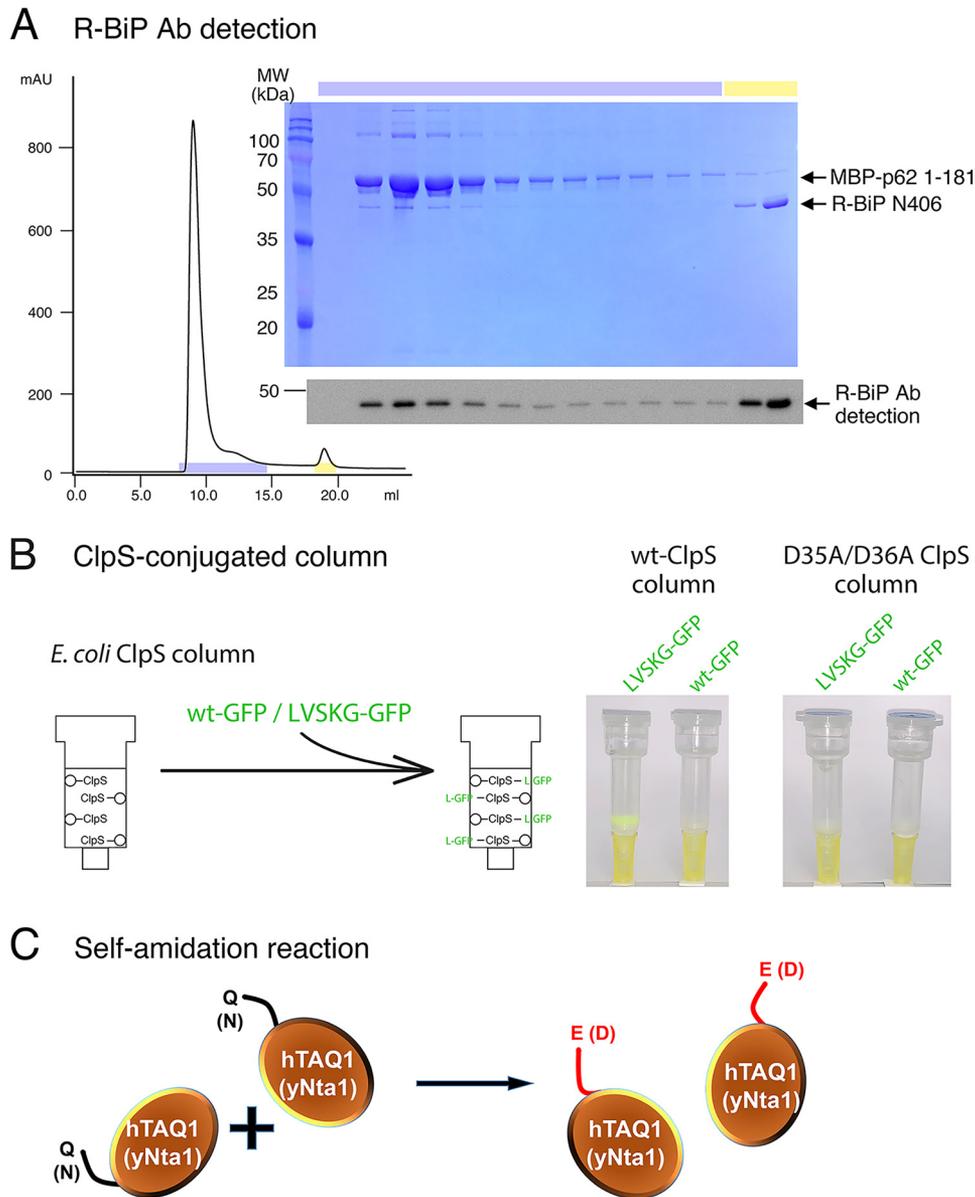


Figure 5. Examples of applications of the LC3B-fusion technique. *A*, generation of the N-degron substrate, R-BiP, using the current technique, and R-BiP detection using an antibody. The mixture of the MBP-PB1-ZZ domain of p62 and R-BiP was applied to the Superose 6 gel filtration column, and the complex formation was monitored. As shown in the Coomassie Blue staining results, the R-BiP fraction that co-migrated with p62 was weakly detected; however, Western blotting using an R-BiP specific antibody showed clear complex formation. *B*, for identifying new type 2 substrates, a ClpS-conjugated affinity column was generated. The LAPPY sequence fused to GFP generated by this method was used as the positive control for this column, which showed a fluorescent green color even after extensive washing. As a negative control, WT GFP was applied to the same column, which did not show any green color. For further validation, the column with D35A/D36A double mutant of ClpS was also tested with wt-GFP and LAPPY-GFP. *C*, enzymatically competent WT human NTAQ1 or yeast Nta1 tagged with N-degron sequences were purified, and the N-terminal sequences were determined by N-terminal sequencing (Table S1). The tertiary destabilizing residues (Gln or Asn) were completely converted to secondary destabilizing residues (Glu or Asp).

times the addition of binding ligands destroys the crystalline lattice (50). Making recombinant proteins with recognition sequences at the N terminus is thus much simpler than either the co-crystallization or soaking techniques. Many other studies have performed successful crystallization using fusion partners (51). Previously, LC3B was fused with its own binding motif, LIR, at its N and C termini to successfully study LC3B-LIR complex structures (22, 23). The current method used in this study with N-degron fused to the N termini of target proteins has now been successfully applied to crystallize proteins in the N-degron pathway, not only to increase crystal quality but

also to yield complex structures. Intriguingly, several structures of other proteins in the N-degron pathway were determined in complexed states unintentionally, although the binding of the N-terminal segments of the neighboring molecules depends on whether such a segment represents a real substrate sequence or not (18, 20, 28, 52) (Table 1). Furthermore, depending on the exact N-degron sequences, it is possible to obtain many different types of crystal packing with the same protein (Table 1), which is also beneficial for structural characterization.

With this technology, we were able to test many biochemical characteristics of proteins in N-degron pathway, because *in*

in vivo substrates generated by enzymatic modifications were successfully produced. Arginylated BiP is a substrate of p62, and leucylated GFP is a model substrate for various applications (Fig. 5, A and B). There are no limitations to the potential of N-degron substrates made using simple mutagenesis at any desired position, except for the addition of proline to the N terminus. Usually, the N-terminal residues do not affect the activity and binding of proteins in the N-degron pathway; therefore, the enzymes we produced are enzymatically competent (Fig. 5C and Table S1), and the proteins could be successfully crystallized and their structures determined (Fig. 3). Using this method, large quantities of N-degron pathway proteins can be obtained, allowing for theoretically any types of *in vitro* study. Therefore, this method can be a standard tool for studying the N-degron pathway.

Experimental procedures

Construct design

Sequences comprising the fusion of LC3B (1–120 amino acids) and proteins involved in the N-degron pathway were amplified from gBlock gene fragments (Integrated DNA Technologies) by PCR and then cloned into BamHI and XhoI restriction sites of a modified pET vector (containing tobacco etch virus and thrombin cleavage sites) to construct the N-terminal hexahistidine-tagged proteins. The resulting plasmids were transformed into *E. coli* BL21 (DE3) cells. We also generated expression vectors for LC3B-fusion proteins using multiple cloning sites (Fig. S1) to ensure easier cloning of target proteins following the LC3B sequence.

Protein expression and purification

The expression of His₆-tagged fusion proteins (LC3B–RLGS–AtPRT6 UBR box, LC3B–M1Q–hNTAQ1, and LC3B–GEEED–ZZ for crystallization, and *E. coli* ClpS, LC3B–LAPPS–GFP for ClpS column) was induced with a final concentration of 0.5 mM isopropyl- β -D-thiogalactoside at 18 °C for 20 h. The cells were harvested by centrifugation and resuspended in buffer A (50 mM Tris-HCl, pH 8.0, and 0.5 mM tris[2-carboxyethyl]phosphine (TCEP)) containing 200 mM NaCl. After sonication, the insoluble material was removed by centrifugation at $27,216 \times g$ for 1 h. The supernatant was loaded onto a HisTrapTM column (GE Healthcare, catalog no. 17-5255-01) and then eluted by gradient purification with buffer A containing 100 mM NaCl and 500 mM imidazole. Eluted samples were further purified using a HiTrapTM Q HP column (GE Healthcare, catalog no. 17-1154-01). The His-LC3B tag was removed using human ATB4B protease at 25 °C overnight. The eluted proteins were concentrated by ultrafiltration (Amicon Ultra 30 K NMWL, Millipore) and loaded onto a HiLoadTM 16/600 SuperdexTM 75pg (GE Healthcare, catalog no. 28-9893-33) column equilibrated with buffer B (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 1 mM TCEP).

Expression of GST-tagged fusion proteins (LC3B–RLGS–yUbr1 UBR box for crystallization and LC3B–R–BiP for antibody recognition) was induced by addition of isopropyl- β -D-thiogalactoside to a final concentration of 0.5 mM at 25 °C for 20 h. The cells were harvested by centrifugation and resuspended in $1 \times$ PBS followed by sonication, after which the insol-

uble material was removed by centrifugation at $27,216 \times g$ for 1 h. The supernatant was then loaded onto a GST column (GE Healthcare) and then eluted with GST elution buffer containing 50 mM Tris-HCl, pH 8.0, 1 mM TCEP, 100 mM NaCl, and 25 mM GSH. The eluted samples were further purified using a HiTrapTM Q HP column (GE Healthcare, catalog no. 17-1154-01). The GST–LC3B tag was removed by cleavage with human ATB4B protease at 25 °C overnight. The eluted proteins were concentrated by ultrafiltration (Amicon Ultra 3 K NMWL, Millipore) and loaded onto a HiLoadTM 16/600 SuperdexTM 75pg column (GE Healthcare, catalog no. 28-9893-33) that was equilibrated with 20 mM Tris-HCl, pH 8.0, 200 mM NaCl, and 1 mM TCEP. The ATG4B protease was prepared as previously reported (53).

Crystallization and data collection

Purified RLGS–AtPRT6 UBR box was concentrated to 9 mg/ml and crystallized at 22 °C using the sitting-drop vapor-diffusion method and by mixing an equal volume of protein with a reservoir solution containing 0.1 M MES, pH 6.5 (Hampton Research, catalog no. HR2-587), and 1.6 M ammonium sulfate (Hampton Research, catalog no. HR2-541). Purified RLGS–yUbr1 UBR box was concentrated to 9 mg/ml and crystallized at 22 °C using the sitting-drop vapor-diffusion method and by mixing an equal volume of protein and reservoir solution containing 0.1 M BIS-Tris, pH 6.5, and 2.7–3.0 M NaCl. The purified M1Q–hNTAQ1 C28S was concentrated to 16 mg/ml and crystallized at 22 °C using the sitting-drop vapor-diffusion method and mixing an equal volume of protein and the reservoir solution containing 0.4 M ammonium sulfate (Hampton Research, catalog no. HR2-541), 0.1 M sodium citrate tribasic dihydrate, pH 5.6 (Hampton Research, catalog no. HR2-735), and 1.0 M lithium sulfate monohydrate (Hampton Research, catalog no. HR2-545). The purified GEEED–p62 ZZ fusion protein was crystallized at 22 °C using the sitting-drop vapor-diffusion method and by mixing an equal volume of protein and a reservoir solution containing 0.17 M ammonium sulfate, 0.085 M sodium cacodylate trihydrate, pH 6.5, and 22–30% (w/v) PEG 8000. The protein crystals were flash-frozen in liquid nitrogen with 20–30% (v/v) glycerol as a cryo-protectant in the original mother liquor. The native protein crystal data were collected using beamlines at Photon Factory and Spring-8 in Japan and Pohang Accelerator Laboratory in South Korea. The diffraction data were indexed, integrated, and scaled using HKL2000 software (54). The data collection statistics are shown in Table 2.

Structure determination and refinement

All structures except that of AtPRT6 UBR box were determined by MR using the p62 ZZ domain (PDB code 5YP7 (17)), yUbr1 UBR box (PDB code 3NIT (18)), and Ntaq1 structures (PDB code 4W79 (20)) as search models. The models obtained by MR calculation were rebuilt and refined in iterative cycles using Coot (55). The initial phases of the AtPRT6 UBR box were determined with a 2.7 Å resolution SAD data set collected at the absorption edge of the zinc atom ($\lambda = 1.282282$ Å) at Beamline 5C at the Pohang Accelerator Laboratory facility (Table S2). Zn-site determination, phasing, and automatic model building

A tool for *in vitro* study of the N-degron pathway

were performed using the SAD phasing module, as implemented in the Phenix software package (56). The initial model was rebuilt and refined in iterative cycles using Coot (55). Ramachandran values were calculated using MolProbity (57). The statistics for collected data and refinements are summarized in Table 2. All structural figures were drawn using PyMOL.

Size-exclusion chromatography coupled with multiangle light scattering

SEC-MALS experiments were performed using a FPLC system (GE Healthcare) connected to a Wyatt MiniDAWN TREOS MALS instrument and Wyatt Optilab rEX differential refractometer. The buffers used were the same as those used in the final protein purification step. The purified protein was loaded onto a Superdex™ 75 Increase 10/300 GL gel filtration column (GE Healthcare, catalog no. 29-1487-21). Ovalbumin was used as the isotropic scatterer for detector normalization. The light scattering was measured and analyzed using ASTRA V software (Wyatt).

R-BiP antibody detection

Purified R-Bip protein was mixed with 2× SDS-PAGE loading buffer (100 mM Tris-HCl, pH 6.8, 4% (w/v) SDS, 0.2% (w/v) bromophenol blue, 20% (v/v) glycerol, and 200 mM DTT) and then was subjected to Western blotting. A rabbit polyclonal anti-R-BiP antibody (Abfrontier, catalog no. AR05-PA0001, 1:1000) was used to detect arginylated BiP.

ClpS-conjugated column

For buffer exchange, purified *E. coli* ClpS protein was loaded onto a Superdex 75 16/600-pg column (GE Healthcare) equilibrated with 20 mM HEPES, pH 7.4, 50 mM NaCl, 1 mM DTT, and 10% (v/v) glycerol. ClpS protein was coupled with Affi-Gel 15 (Bio-Rad) at 4 °C for 4 h. To validate coupling of ClpS to the Affi-Gel 15 column, purified wt-GFP, and LAPPS-GFP proteins were loaded onto the affinity column, which then turned green (Fig. 3B). The same experiment was also performed with the ClpS D35A/D36A double mutant column, and the substrates were not able to bind with it. After several round of washing, the bound GFP was eluted with 1 mg/ml FR dipeptide, removing all green color from the column.

Time-course cleavage assay

Purified proteins (His-LC3B/Ub/SUMO-RLGS-AtPRT6 UBR box) were assayed against proteases over a 120-min time course at a 1:100 molar ratio. Human ATG4B (lab-made) for LC3B protease, hUSP2 (BPS Bioscience, catalog no. 80392) for ubiquitin protease, and yUlp1 (Kerafast, EB 1008-FP) for SUMO protease were used. All reactions were terminated by adding SDS-PAGE sample buffer and analyzed using SDS-PAGE and Coomassie staining.

Accession numbers

The atomic coordinates and structural factors have been deposited in the Protein Data Bank under the PDB ID codes 6LHN (RLGSSS-AtPRT6 UBR box), 6KGI (RLGS-yUbr1 UBR box), 6KGJ (M1Q-hNTAQ1 C28S), and 6KHZ (GEEED-p62 ZZ).

Author contributions—L. K., D. H. K., J. H., and M. R. P. data curation; L. K., J. H., and M. R. P. formal analysis; L. K., D. H. K., and H. K. S. investigation; L. K., J. H., and M. R. P. visualization; L. K., D. H. K., J. H., and M. R. P. methodology; L. K., D. H. K., and H. K. S. writing-original draft; D. H. K. and H. K. S. conceptualization; D. H. K., J. H., M. R. P., and H. K. S. validation; H. K. S. resources; H. K. S. supervision; H. K. S. funding acquisition; H. K. S. project administration; H. K. S. writing-review and editing.

Acknowledgments—We thank the staff at Beamlines 5C and 11C, Pohang Accelerator Laboratory, Republic of Korea and Beamline BL-17A, Photon Factory, Japan, for help with X-ray data collection. This work was performed in part under the International Collaborative Research Program of the Institute for Protein Research, Osaka University Grant ICR-17-05. Diffraction data were collected at the Osaka University Beamline BL44XU at SPring-8 (Harima, Japan) under Proposal 2019A6973.

References

1. Varshavsky, A. (2019) N-degron and C-degron pathways of protein degradation. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 358–366 [CrossRef Medline](#)
2. Bachmair, A., Finley, D., and Varshavsky, A. (1986) *In vivo* half-life of a protein is a function of its amino-terminal residue. *Science* **234**, 179–186 [CrossRef Medline](#)
3. Varshavsky, A. (1992) The N-end rule. *Cell* **69**, 725–735 [CrossRef Medline](#)
4. Varshavsky, A. (1996) The N-end rule: functions, mysteries, uses. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12142–12149 [CrossRef Medline](#)
5. Tobias, J. W., Shrader, T. E., Rocap, G., and Varshavsky, A. (1991) The N-end rule in bacteria. *Science* **254**, 1374–1377 [CrossRef Medline](#)
6. Potuschak, T., Stary, S., Schögelhofer, P., Becker, F., Nejkiskaia, V., and Bachmair, A. (1998) PRT1 of *Arabidopsis thaliana* encodes a component of the plant N-end rule pathway. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7904–7908 [CrossRef Medline](#)
7. Varshavsky, A. (2011) The N-end rule pathway and regulation by proteolysis. *Protein Sci.* **20**, 1298–1345 [CrossRef Medline](#)
8. Hwang, C. S., Shemorry, A., and Varshavsky, A. (2010) N-terminal acetylation of cellular proteins creates specific degradation signals. *Science* **327**, 973–977 [CrossRef Medline](#)
9. Starheim, K. K., Gevaert, K., and Arnesen, T. (2012) Protein N-terminal acetyltransferases: when the start matters. *Trends Biochem. Sci.* **37**, 152–161 [CrossRef Medline](#)
10. Chen, S. J., Wu, X., Wadas, B., Oh, J. H., and Varshavsky, A. (2017) An N-end rule pathway that recognizes proline and destroys gluconeogenic enzymes. *Science* **355**, eaal3655 [CrossRef Medline](#)
11. Kim, J. M., Seok, O. H., Ju, S., Heo, J. E., Yeom, J., Kim, D. S., Yoo, J. Y., Varshavsky, A., Lee, C., and Hwang, C. S. (2018) Formyl-methionine as an N-degron of a eukaryotic N-end rule pathway. *Science* **362**, eaal3655 [CrossRef Medline](#)
12. Zenker, M., Mayerle, J., Lerch, M. M., Tagariello, A., Zerres, K., Durie, P. R., Beier, M., Hülskamp, G., Guzman, C., Rehder, H., Beemer, F. A., Hamel, B., Vanlieferinghen, P., Gershoni-Baruch, R., Vieira, M. W., et al. (2005) Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome). *Nat. Genet.* **37**, 1345–1350 [CrossRef Medline](#)
13. Ichimura, Y., Waguri, S., Sou, Y. S., Kageyama, S., Hasegawa, J., Ishimura, R., Saito, T., Yang, Y., Kouno, T., Fukutomi, T., Hoshii, T., Hirao, A., Takagi, K., Mizushima, T., Motohashi, H., et al. (2013) Phosphorylation of p62 activates the Keap1-Nrf2 pathway during selective autophagy. *Mol. Cell* **51**, 618–631 [CrossRef Medline](#)
14. Mizushima, N., and Komatsu, M. (2011) Autophagy: renovation of cells and tissues. *Cell* **147**, 728–741 [CrossRef Medline](#)
15. Cha-Molstad, H., Yu, J. E., Feng, Z., Lee, S. H., Kim, J. G., Yang, P., Han, B., Sung, K. W., Yoo, Y. D., Hwang, J., McGuire, T., Shim, S. M., Song, H. D., Ganipiseti, S., Wang, N., et al. (2017) p62/SQSTM1/Sequestosome-1 is

- an N-recognin of the N-end rule pathway which modulates autophagosome biogenesis. *Nat. Commun.* **8**, 102 [CrossRef Medline](#)
16. Kwon, D. H., Park, O. H., Kim, L., Jung, Y. O., Park, Y., Jeong, H., Hyun, J., Kim, Y. K., and Song, H. K. (2018) Insights into degradation mechanism of N-end rule substrates by p62/SQSTM1 autophagy adapter. *Nat. Commun.* **9**, 3291 [CrossRef Medline](#)
 17. Kwon, D. H., Kim, L., and Song, H. K. (2019) pH-dependent regulation of SQSTM1/p62 during autophagy. *Autophagy* **15**, 180–181 [CrossRef Medline](#)
 18. Choi, W. S., Jeong, B. C., Joo, Y. J., Lee, M. R., Kim, J., Eck, M. J., and Song, H. K. (2010) Structural basis for the recognition of N-end rule substrates by the UBR box of ubiquitin ligases. *Nat. Struct. Mol. Biol.* **17**, 1175–1181 [CrossRef Medline](#)
 19. Kim, M. K., Oh, S. J., Lee, B. G., and Song, H. K. (2016) Structural basis for dual specificity of yeast N-terminal amidase in the N-end rule pathway. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12438–12443 [CrossRef Medline](#)
 20. Park, M. S., Bitto, E., Kim, K. R., Bingman, C. A., Miller, M. D., Kim, H. J., Han, B. W., and Phillips, G. N., Jr. (2014) Crystal structure of human protein N-terminal glutamine amidohydrolase, an initial component of the N-end rule pathway. *PLoS One* **9**, e111142 [CrossRef Medline](#)
 21. Dong, C., Zhang, H., Li, L., Tempel, W., Loppnau, P., and Min, J. (2018) Molecular basis of GID4-mediated recognition of degrons for the Pro/N-end rule pathway. *Nat. Chem. Biol.* **14**, 466–473 [CrossRef Medline](#)
 22. Kwon, D. H., Kim, L., Kim, B. W., Kim, J. H., Roh, K. H., Choi, E. J., and Song, H. K. (2017) A novel conformation of the LC3-interacting region motif revealed by the structure of a complex between LC3B and RavZ. *Biochem. Biophys. Res. Commun.* **490**, 1093–1099 [CrossRef Medline](#)
 23. Rogov, V. V., Suzuki, H., Fiskin, E., Wild, P., Kniss, A., Rozenknop, A., Kato, R., Kawasaki, M., McEwan, D. G., Löhr, F., Güntert, P., Dikic, I., Wakatsuki, S., and Dötsch, V. (2013) Structural basis for phosphorylation-triggered autophagic clearance of *Salmonella*. *Biochem. J.* **454**, 459–466 [CrossRef Medline](#)
 24. Bartel, B., Wüning, L., and Varshavsky, A. (1990) The recognition component of the N-end rule pathway. *EMBO J.* **9**, 3179–3189 [CrossRef Medline](#)
 25. Garzón, M., Eifler, K., Faust, A., Scheel, H., Hofmann, K., Koncz, C., Yephremov, A., and Bachmair, A. (2007) PRT6/At5g02310 encodes an *Arabidopsis* ubiquitin ligase of the N-end rule pathway with arginine specificity and is not the CER3 locus. *FEBS Lett.* **581**, 3189–3196 [CrossRef Medline](#)
 26. Baker, R. T. (1996) Protein expression using ubiquitin fusion and cleavage. *Curr. Opin. Biotechnol.* **7**, 541–546 [CrossRef Medline](#)
 27. Satakarni, M., and Curtis, R. (2011) Production of recombinant peptides as fusions with SUMO. *Protein Expr. Purif.* **78**, 113–119 [CrossRef Medline](#)
 28. Matta-Camacho, E., Kozlov, G., Li, F. F., and Gehring, K. (2010) Structural basis of substrate recognition and specificity in the N-end rule pathway. *Nat. Struct. Mol. Biol.* **17**, 1182–1187 [CrossRef Medline](#)
 29. Muñoz-Escobar, J., Matta-Camacho, E., Cho, C., Kozlov, G., and Gehring, K. (2017) Bound waters mediate binding of diverse substrates to a ubiquitin ligase. *Structure* **25**, 719–729. [e713 CrossRef Medline](#)
 30. Bachmair, A., and Varshavsky, A. (1989) The degradation signal in a short-lived protein. *Cell* **56**, 1019–1032 [CrossRef Medline](#)
 31. Baker, R. T., and Varshavsky, A. (1995) Yeast N-terminal amidase: a new enzyme and component of the N-end rule pathway. *J. Biol. Chem.* **270**, 12065–12074 [CrossRef Medline](#)
 32. Wang, H., Piatkov, K. I., Brower, C. S., and Varshavsky, A. (2009) Glutamine-specific N-terminal amidase, a component of the N-end rule pathway. *Mol. Cell* **34**, 686–695 [CrossRef Medline](#)
 33. Grigoryev, S., Stewart, A. E., Kwon, Y. T., Arfin, S. M., Bradshaw, R. A., Jenkins, N. A., Copeland, N. G., and Varshavsky, A. (1996) A mouse amidase specific for N-terminal asparagine: the gene, the enzyme, and their function in the N-end rule pathway. *J. Biol. Chem.* **271**, 28521–28532 [CrossRef Medline](#)
 34. Cha-Molstad, H., Sung, K. S., Hwang, J., Kim, K. A., Yu, J. E., Yoo, Y. D., Jang, J. M., Han, D. H., Molstad, M., Kim, J. G., Lee, Y. J., Zakrzewska, A., Kim, S. H., Kim, S. T., Kim, S. Y., *et al.* (2015) Amino-terminal arginylation targets endoplasmic reticulum chaperone BiP for autophagy through p62 binding. *Nat. Cell Biol.* **17**, 917–929 [CrossRef Medline](#)
 35. Erbsse, A., Schmidt, R., Bornemann, T., Schneider-Mergener, J., Mogk, A., Zahn, R., Dougan, D. A., and Bukau, B. (2006) ClpS is an essential component of the N-end rule pathway in *Escherichia coli*. *Nature* **439**, 753–756 [CrossRef Medline](#)
 36. Lee, B. G., Kim, M. K., and Song, H. K. (2011) Structural insights into the conformational diversity of ClpP from *Bacillus subtilis*. *Mol. Cells* **32**, 589–595 [CrossRef Medline](#)
 37. Schuenemann, V. J., Kralik, S. M., Albrecht, R., Spall, S. K., Truscott, K. N., Dougan, D. A., and Zeth, K. (2009) Structural basis of N-end rule substrate recognition in *Escherichia coli* by the ClpAP adaptor protein ClpS. *EMBO Rep.* **10**, 508–514 [CrossRef Medline](#)
 38. Nishimura, K., Asakura, Y., Friso, G., Kim, J., Oh, S. H., Rutschow, H., Ponnala, L., and van Wijk, K. J. (2013) ClpS1 is a conserved substrate selector for the chloroplast Clp protease system in *Arabidopsis*. *Plant Cell* **25**, 2276–2301 [CrossRef Medline](#)
 39. Uhlén, M., and Moks, T. (1990) Gene fusions for purpose of expression: an introduction. *Methods Enzymol.* **185**, 129–143 [CrossRef Medline](#)
 40. Varshavsky, A. (2000) Ubiquitin fusion technique and its descendants. *Methods Enzymol.* **327**, 578–593 [CrossRef Medline](#)
 41. Hondred, D., Walker, J. M., Mathews, D. E., and Vierstra, R. D. (1999) Use of ubiquitin fusions to augment protein expression in transgenic plants. *Plant Physiol.* **119**, 713–724 [CrossRef Medline](#)
 42. Catanzariti, A. M., Soboleva, T. A., Jans, D. A., Board, P. G., and Baker, R. T. (2004) An efficient system for high-level expression and easy purification of authentic recombinant proteins. *Protein Sci.* **13**, 1331–1339 [CrossRef Medline](#)
 43. Panavas, T., Sanders, C., and Butt, T. R. (2009) SUMO fusion technology for enhanced protein production in prokaryotic and eukaryotic expression systems. *Methods Mol. Biol.* **497**, 303–317 [CrossRef Medline](#)
 44. Butt, T. R., Edavettal, S. C., Hall, J. P., and Mattern, M. R. (2005) SUMO fusion technology for difficult-to-express proteins. *Protein Expr. Purif.* **43**, 1–9 [CrossRef Medline](#)
 45. Malakhov, M. P., Mattern, M. R., Malakhova, O. A., Drinker, M., Weeks, S. D., and Butt, T. R. (2004) SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J. Struct. Funct. Genomics* **5**, 75–86 [CrossRef Medline](#)
 46. Wang, X., Liu, H., Liu, Y., Li, Y., Yan, L., Yuan, X., Zhang, Y., Wu, Y., Liu, J., Zhang, C., and Chu, Y. (2016) A novel strategy for the preparation of codon-optimized truncated Ulp1 and its simplified application to cleavage the SUMO fusion protein. *Protein J.* **35**, 115–123 [CrossRef Medline](#)
 47. Noda, N. N., Kumeta, H., Nakatogawa, H., Satoo, K., Adachi, W., Ishii, J., Fujioka, Y., Ohsumi, Y., and Inagaki, F. (2008) Structural basis of target recognition by Atg8/LC3 during selective autophagy. *Genes Cells* **13**, 1211–1218 [CrossRef Medline](#)
 48. Kabeya, Y., Mizushima, N., Ueno, T., Yamamoto, A., Kirisako, T., Noda, T., Kominami, E., Ohsumi, Y., and Yoshimori, T. (2000) LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. *EMBO J.* **19**, 5720–5728 [CrossRef Medline](#)
 49. Hwang, C. S., Sukalo, M., Batygin, O., Addor, M. C., Brunner, H., Aytes, A. P., Mayerle, J., Song, H. K., Varshavsky, A., and Zenker, M. (2011) Ubiquitin ligases of the N-end rule pathway: assessment of mutations in UBR1 that cause the Johanson-Blizzard syndrome. *PLoS One* **6**, e24925 [CrossRef Medline](#)
 50. Hassell, A. M., An, G., Bledsoe, R. K., Bynum, J. M., Carter, H. L., 3rd, Deng, S. J., Gampe, R. T., Grisard, T. E., Madauss, K. P., Nolte, R. T., Rocque, W. J., Wang, L., Weaver, K. L., Williams, S. P., Wisely, G. B., *et al.* (2007) Crystallization of protein–ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 72–79 [CrossRef Medline](#)
 51. Kobe, B., Ve, T., and Williams, S. J. (2015) Fusion-protein-assisted protein crystallization. *Acta Crystallogr. F Struct. Biol. Commun.* **71**, 861–869 [CrossRef Medline](#)
 52. Román-Hernández, G., Hou, J. Y., Grant, R. A., Sauer, R. T., and Baker, T. A. (2011) The ClpS adaptor mediates staged delivery of N-end rule substrates to the AAA+ ClpAP protease. *Mol. Cell* **43**, 217–228 [CrossRef Medline](#)
 53. Kwon, D. H., Kim, S., Jung, Y. O., Roh, K. H., Kim, L., Kim, B. W., Hong, S. B., Lee, I. Y., Song, J. H., Lee, W. C., Choi, E. J., Hwang, K. Y., and Song, H. K. (2017) The 1:2 complex between RavZ and LC3 reveals a mechanism

A tool for *in vitro* study of the N-degron pathway

- for deconjugation of LC3 on the phagophore membrane. *Autophagy* **13**, 70–81 [CrossRef Medline](#)
54. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 [CrossRef Medline](#)
55. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 [CrossRef Medline](#)
56. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 [CrossRef Medline](#)
57. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 [CrossRef Medline](#)

Use of the LC3B-fusion technique for biochemical and structural studies of proteins involved in the N-degron pathway

Leehyeon Kim, Do Hoon Kwon, Jiwon Heo, Mi Rae Park and Hyun Kyu Song

J. Biol. Chem. 2020, 295:2590-2600.

doi: 10.1074/jbc.RA119.010912 originally published online January 9, 2020

Access the most updated version of this article at doi: [10.1074/jbc.RA119.010912](https://doi.org/10.1074/jbc.RA119.010912)

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

This article cites 57 references, 15 of which can be accessed free at <http://www.jbc.org/content/295/9/2590.full.html#ref-list-1>