Molecules and Cells

N-Terminal Modifications of Ubiquitin via Methionine Excision, Deamination, and Arginylation Expand the Ubiquitin Code

Kha The Nguyen[1,4], Shinyeong Ju[2,4], Sang-Yoon Kim[1], Chang-Seok Lee[1], Cheolju Lee[2,3,*], and Cheol-Sang Hwang[1,*]

[1]Department of Life Sciences, Pohang University of Science and Technology, Pohang 37673, Korea, [2]Center for Theragnosis, Korea Institute of Science and Technology, Seoul 02792, Korea, [3]Division of Bio-Medical Science & Technology, KIST School, Korea University of Science and Technology, Seoul 02792, Korea, [4]These authors contributed equally to this work.
*Correspondence: cshwang@postech.ac.kr (C-SH); clee270@kist.re.kr (CL)
https://doi.org/10.14348/molcells.2022.2027
www.molcells.org

Ubiquitin (Ub) is post-translationally modified by Ub itself or Ub-like proteins, phosphorylation, and acetylation, among others, which elicits a variety of Ub topologies and cellular functions. However, N-terminal (Nt) modifications of Ub remain unknown, except the linear head-to-tail ubiquitylation via Nt-Met. Here, using the yeast *Saccharomyces cerevisiae* and an Nt-arginylated Ub-specific antibody, we found that the detectable level of Ub undergoes Nt-Met excision, Nt-deamination, and Nt-arginylation. The resulting Nt-arginylated Ub and its conjugated proteins are upregulated in the stationary-growth phase or by oxidative stress. We further proved the existence of Nt-arginylated Ub *in vivo* and identified Nt-arginylated Ub-protein conjugates using stable isotope labeling by amino acids in cell culture (SILAC)-based tandem mass spectrometry. *In silico* structural modeling of Nt-arginylated Ub predicted that Nt-Arg flexibly protrudes from the surface of the Ub, thereby most likely providing a docking site for the factors that recognize it. Collectively, these results reveal unprecedented Nt-arginylated Ub and the pathway by which it is produced, which greatly expands the known complexity of the Ub code.

Keywords: arginylation, deamination, methionine excision, N-degron, proteolysis, ubiquitin code

INTRODUCTION

The covalent attachment of ubiquitin (Ub) to target proteins recruits the cascade reactions of the E1 activating enzymes, E2 conjugation enzymes, and E3 ligases. Ub can also be post-translationally tagged by Ub itself or Ub-like proteins, phosphorylation, acetylation, ribosylation, and deamination (Mattiroli and Penengo, 2021; Swatek and Komander, 2016) (for details, see the references therein) (Fig. 1A). These post-translational modifications induce distinct topologies of the Ub molecule, leading to a vast range of cellular outcomes (cell division, death, development, etc.) via the creation of distinct words for the Ub code (Dittmar and Selbach, 2017; Komander and Rape, 2012; Kwon and Ciechanover, 2017). The identification of new types of Ub modification can reveal unrecognized Ub codes, which deepens our understanding of a vast range of Ub-associated biological processes. However, the post-translational modifications of Ub *per se* and their regulatory mechanisms remain incompletely understood.

Virtually all newly made polypeptides from ribosomes contain methionine (Met) due to the design of the AUG initiation codon. The initiator Met is frequently excised by Met-aminopeptidases (MetAPs), in cases where the residue at position 2 is Gly, Ala, Ser, Thr, Pro, or Val. Intriguingly, Nt-Met excision

Received 19 November, 2021; revised 16 December, 2021; accepted 12 January, 2022; published online 2 March, 2022

eISSN: 0219-1032

(NME) is an evolutionarily conserved and essential process that is applied to two-thirds of proteins (Giglione et al., 2004). NME also operates even in a set of proteins with a residue larger than Val at position 2, although the mechanism underlying this is still unclear (Chen and Kashina, 2021; Redman and Rubenstein, 1981; Sadis et al., 1995). Specifically in the yeast *Saccharomyces cerevisiae*, the Nt-Asn and Nt-Gln exposed after NME are deaminated to Nt-Asp and Nt-Glu, respectively, by Nta1 Nt-amidases and subsequently arginylated by Ate1 arginyltransferase (Nguyen et al., 2019). The resulting Nt-arginylated proteins are targeted for degradation by the Arg/N-degron pathway (previously known as the N-end rule pathway), the prototype of N-degron pathways (Fig. 1B) (Bachmair et al., 1986; Tasaki et al., 2012; Varshavsky, 2019). The hitherto established N-degron pathways of *S. cerevisiae* encompass the Arg/N-degron pathway (that targets unmodified Nt-Arg, His, Lys, Leu, Ile, Trp, Phe, Tyr, and Met-Φ [Met followed by a hydrophobic residue]) (Kim and Hwang, 2014; Kim et al., 2014), the Ac/N-degron pathway (that targets acetylated Nt-residues) (Hwang et al., 2010; Lee et al., 2016; Nguyen et al., 2018; Shemorry et al., 2013), the Pro/N-degron pathway (that targets Nt-Pro) (Chen et al., 2017; 2021), and the fMet/N-degron pathway (that targets Nt-formylated Met) (Kim et al., 2018) (Figs. 1B-1E).

The 76-residue Ub starts with the Met-Gln (MQ) N-terminus. The MQ-starting Ub (MQ-Ub$^{3\text{-}76}$) is hereafter denoted simply as ^{MQ}Ub. Thus, we hypothesized that ^{MQ}Ub might be processed by these NME-provoked cascade reactions of Nt-deamination and Nt-arginylation to yield Nt-arginylated RE-Ub$^{3\text{-}76}$ (denoted here as ^{RE}Ub) in yeast and mammalian cells (Figs. 1A and 2A). Indeed, we here identify the previously unknown ^{RE}Ub in yeast and consider the ramifications of this discovery, which greatly expands the known complexity and biological significance of the Ub code.

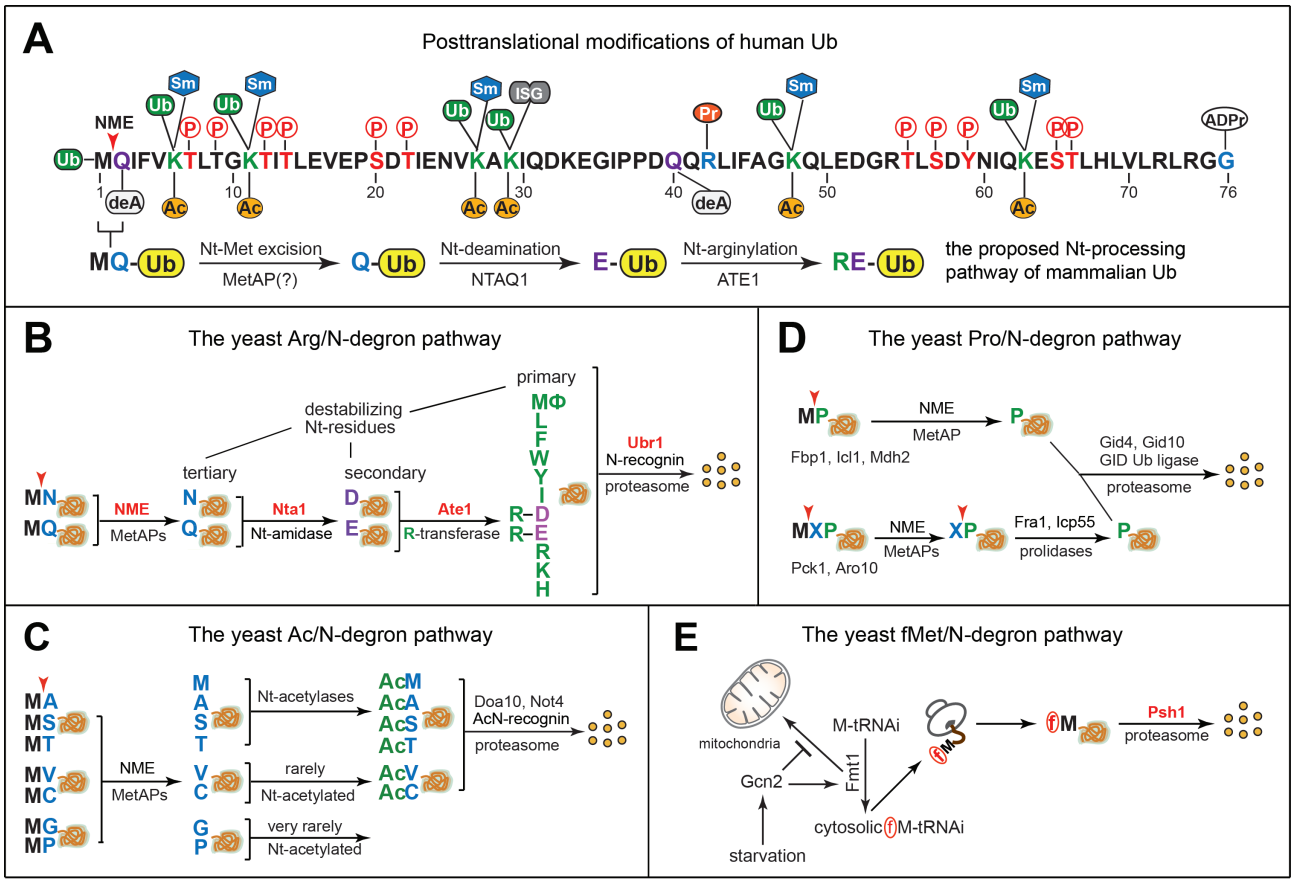


Fig. 1. Post-translational modification sites on Ub and the yeast N-degron pathways. (A) The post-translational modification sites of human Ub are depicted. Ub, ubiquitylation (green); Sm, sumonylation (blue); ISG, ISGylation (grey); P, phosphorylation (open red circle); Ac, acetylation (orange); deA, deamination (light grey); Pr, phosphoribosylation (red); ADPr, ADP-ribosylation (white). The proposed Nt-modifications of mammalian Ub based on this study involve NME by (unknown) MetAP Met-aminopeptidases, Nt-deamination by NTAQ1 Nt-Gln amidase, and Nt-arginylation by ATE1 arginyltransferase. See previous reports (Dittmar and Selbach, 2017; Mattiroli and Penengo, 2021; Swatek and Komander, 2016) and references therein for details. (B) The yeast Arg/N-degron pathway that targets for proteolysis destabilizing Nt-residues. In Met-Φ, Φ denotes a bulky hydrophobic residue. (C) The yeast Ac/N-degron pathway that targets for proteolysis the Nt-acetyl moiety of Nt-acetylated proteins. (D) The Pro/N-degron pathway that targets for proteolysis Nt-Pro of proteins. (E) The fMet/N-degron pathway that targets for proteolysis Nt-formylated Met of proteins. See previous reports (Chen et al., 2021; Kim et al., 2018; Lee et al., 2016; Varshavsky, 2019) and references therein for details.

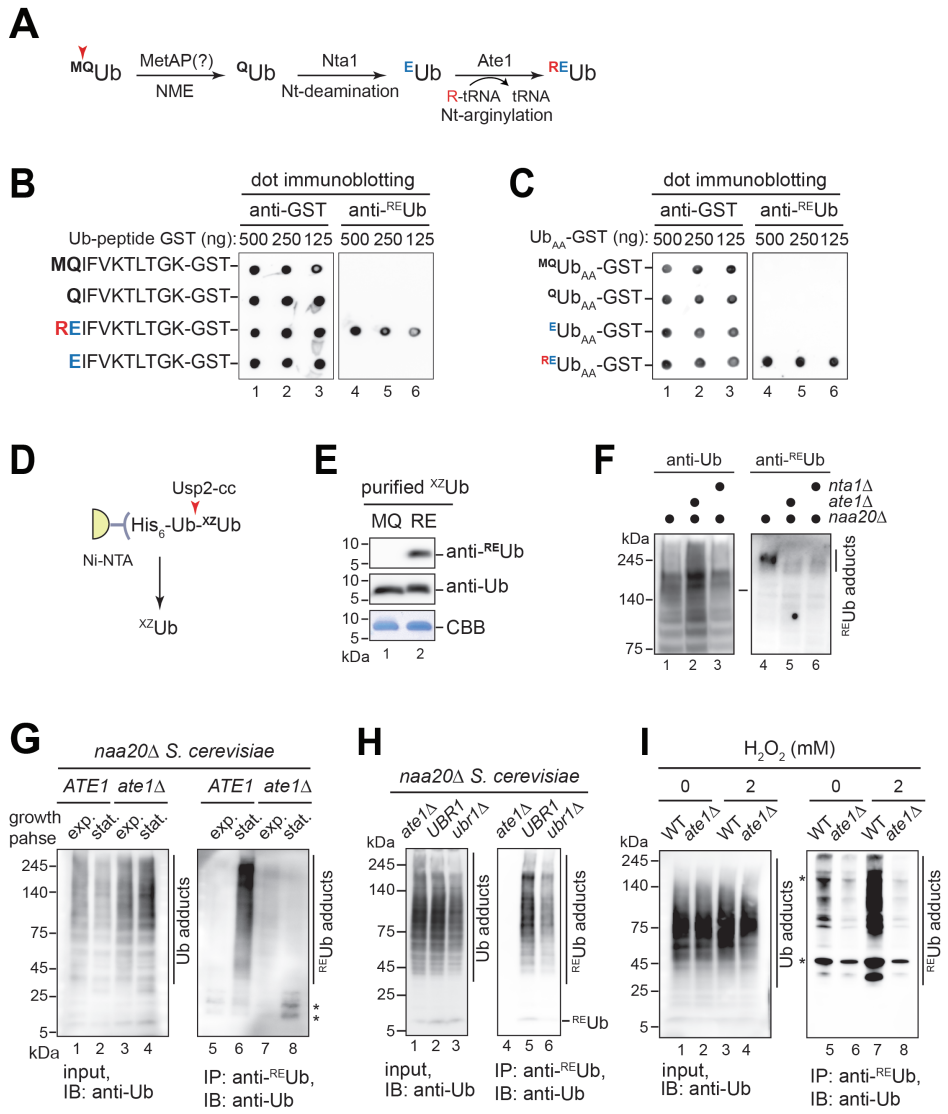


Fig. 2. Detection of Nt-arginylated ^{RE}Ub and its regulation. (A) A proposed model for the production of ^{RE}Ub in yeast cells *in vivo*. ^{MQ}Ub, wild-type Ub; ^{Q}Ub, Nt-Met-excised ^{MQ}Ub; ^{E}Ub, Nt-deaminated ^{Q}Ub; ^{RE}Ub, Nt-arginylated ^{E}Ub; NME, Nt-Met excision; MetAPs, Met-aminopeptidases; Nta1, Nt-amidase; Ate1, arginyltransferase. (B) Characterization of anti-^{RE}Ub that was raised against an ^{RE}Ub Nt-peptide, REIFVKTLTGKC. Dot immunoblotting with affinity-purified anti-^{RE}Ub, using indicated amounts of the spotted REIFVKTLTGK-GST fusion and its Nt-varied counterparts and ^{MQ}Ub. (C) Same as in (B) but with ^{MQ}Ub$_{AA}$-GST, ^{Q}Ub$_{AA}$-GST, ^{E}Ub$_{AA}$-GST, and ^{RE}Ub$_{AA}$-GST. In Ub$_{AA}$, C-terminal diglycine of Ub is mutated into dialanine. (D) A scheme for the production of ^{MQ}Ub and ^{RE}Ub. ^{His6}Ub-^{MQ}Ub and ^{His6}Ub-^{RE}Ub were purified from the extracts of *E. coli*. The Ni-NTA agarose-bound ^{His6}Ub-^{MQ}Ub or ^{His6}Ub-^{RE}Ub was deubiquitylated *in vitro* with the Usp2-cc deubiquitylating enzyme for the production of ^{MQ}Ub and ^{RE}Ub, respectively. (E) Immunoblotting of purified ^{MQ}Ub and ^{RE}Ub using anti-^{RE}Ub or anti-Ub (P4D1) antibodies. CBB, Coomassie Brilliant Blue staining. (F) Immunoblotting of the extracts from the stationary-phase *naa20Δ* cells, *ate1Δ naa20Δ* cells, and *nta1Δ naa20Δ* cells with anti-Ub (P4D1) or anti-^{RE}Ub. (G) Extracts from the exponential-phase ($A_{600} \approx 1$) or stationary-phase ($A_{600} \approx 3$) *naa20Δ* cells and *ate1Δ naa20Δ* cells were immunoprecipitated (IP) with anti-^{RE}Ub, followed by immunoblotting (IB) with anti-Ub (P4D1). (H) Same as in (G) but with the stationary-phase *naa20Δ* cells, *ate1Δ naa20Δ* cells, and *ubr1Δ naa20Δ* cells. (I) Same as in (G) but wild-type (WT) cells and *ate1Δ* cells that were, respectively, cultured to the exponential phase ($A_{600} \approx 1$) either in the absence or presence of 2 mM H_2O_2. In (G) and (I), * indicates non-specific bands.

MATERIALS AND METHODS

Yeast strains, plasmids, and culture media

The yeast strains, plasmids, and oligomers used in this study are listed in Table 1. To construct pCH6913, pCH6914, pCH6915, or pCH6916 expressing ^{MQ}Ub$_{AA}$-GST, ^{Q}Ub$_{AA}$-GST, ^{E}Ub$_{AA}$-GST, or ^{RE}Ub$_{AA}$-GST in the pHUE vector (Catanzariti et al., 2004), respectively, Ub variant-encoding open reading frames (ORFs) were polymerase chain reaction (PCR)-amplified from pCH5097 as a template using a set of primer

pairs: OCH7003/OCH8110, OCH8089/OCH8110, OCH8088/OCH8110, or OCH6968/OCH8110. The resulting PCR products were digested with *Sac*II/*Eco*RI and triply ligated into *Sac*II/*Hin*dIII-cut pHUE with *Eco*RI/*Hin*dIII-cut GST-encoding DNA (that was PCR-amplified from pCH33 using OCH6991/OCH6512 primers). To construct pCH7047 expressing MQIFVKTLTGK-GST, the PCR-amplified DNA fragment using OCH8085/OCH6512 primers was digested with *Sac*II/*Hin*dIII and ligated into *Sac*II/*Hin*dIII-cut pHUE. To construct pCH7048, pCH7049, and pCH7050, the DNA fragments encoding QIFVKTLTGK-GST, EIFVKTLTGK-GST, or REIFVKTLTGK-GST were PCR-amplified from pCH7047 using the set of primers OCH8089/OCH6512, OCH8088/OCH6512 or OCH6968/OCH6512 and pCH7047, respectively. The resulting PCR products were digested with *Sac*II/*Hin*dIII and ligated into *Sac*II/*Hin*dIII-cut pHUE, yielding pCH7048, pCH7049, and pCH7050. Further details of the cloning procedures for the plasmids are available upon request. All final plasmids were verified by DNA sequencing.

Yeast cells were cultured in YPD (1% yeast extract, 2% peptone, 2% glucose) and synthetic complete (0.67% yeast nitrogen base with ammonium sulfate, 2% glucose, specific compounds essential for the growth of an auxotrophic strain) media. The construction of yeast strains and DNA transformation were performed following standard techniques (Gietz and Schiestl, 2007; Janke et al., 2004; Sherman, 2002).

Purification of Ub-GST fusions and their derivatives

pCH6913 (expressing ^{6His}Ub-^{MQ}Ub$_{AA}$-GST), pCH6914 (expressing ^{6His}Ub-^{Q}Ub$_{AA}$-GST), pCH6915 (expressing ^{6His}Ub-^{E}Ub$_{AA}$-GST), pCH6916 (expressing ^{6His}Ub-^{RE}Ub$_{AA}$-GST), pCH7047 (expressing ^{6His}Ub-MQIFVKTLTGK-GST), pCH7048 (expressing ^{6His}Ub-QIFVKTLTGK-GST), pCH7049 (expressing ^{6His}Ub-EIFVKTLTGK-GST), and pCH7050 (expressing ^{6His}Ub-REIFVKTLTGK-GST) were transformed into BL21 (DE3) *Escherichia coli*. Five milliliters of overnight culture of these transformants was inoculated into 500 ml of Lysogeny Broth (LB) medium with ampicillin (final concentration of 100 μg/ml), followed by growth at 37°C to A_{600} of ~0.6. Expression of the Ub fusions was induced by incubating *E. coli* cells with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Cat. No. I2481C; GoldBio, USA) at 18°C overnight.

Cell pellets were thawed and resuspended in lysis buffer (50 mM HEPES, 150 mM NaCl, 10% glycerol, 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride [PMSF, Cat. No. 10837091001; Sigma-Aldrich, USA], 1 mM DTT, 0.1% Triton X-100) with lysozyme (final concentration of 1 mg/ml). Resuspended cells were incubated on ice for 20 min and disrupted by sonication (VCX-130; SONICS, USA), five times for 30 s each, at 1 min intervals with 40% amplitude. Cell lysates were centrifuged at 25,000 × *g* for 20 min at 4°C and the supernatants were incubated with 0.5 ml of pre-washed Glutathione Sepharose 4B (Cat. No. GE17075605; GE Healthcare,

Table 1. Yeast strains, plasmids, and primers used in this study

Material	Description or sequence	Source
Yeasts		
JD53	*MATα, his3-Δ200, leu2-3, 112 lys2-801, trp1-Δ63, ura3-52*	Dohmen et al., 1995
CHY367	*naa20::NatMX4* in JD53	Lab collection
CHY2009	*ubr1Δ::KanMX6, naa20Δ::NatMX4* in JD53	Lab collection
CHY2014	*ate1Δ::HphNT2* in JD53	Lab collection
CHY2015	*ate1Δ::HphNT2, naa20Δ::NatMX4* in JD53	Lab collection
CHY2017	*nta1Δ::HphNT2, naa20Δ::NatMX4* in JD53	Lab collection
CHY3188	*arg4Δ::KnaMX6* in JD53	Lab collection
CHY6056	*arg4Δ::KanMX6, ate1Δ::HphNT2* in JD53	This study
Plasmids		
pCH20	pHUE	Catanzariti et al., 2004
pCH33	pGEX4T-3	Lab collection
pCH5097	^{ha}Ub in pRK5	Lim et al., 2005
pCH6913	^{MQ}Ub$_{AA}$-GST in pHUE	This study
pCH6914	^{Q}Ub$_{AA}$ -GST in pHUE	This study
pCH6915	^{E}Ub$_{AA}$-GST in pHUE	This study
pCH6916	^{RE}Ub$_{AA}$-GST in pHUE	This study
pCH7047	MQIFVKTLTGK-GST in pHUE	This study
pCH7048	QIFVKTLTGK-GST in pHUE	This study
pCH7049	EIFVKTLTGK-GST in pHUE	This study
pCH7050	REIFVKTLTGK-GST in pHUE	This study
Oligomers		
OCH6512	5′-GGTAAGCTTTCAACGCGGAACCAGATCCGATTT-3′	
OCH6968	5′-GATCCCGCGGTGGAAGAGAGATCTTCGTGAAGACCCTG-3′	
OCH6991	5′-GCTGAATTCATGTCCCCTATACTAGGTTATTGGAAAA-3′	
OCH7003	5′-GATCCCGCGGTGGAATGCAGATCTTCGTGAAGACCCTG-3′	
OCH8085	5′-GATCCCGCGGTGGAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCGGTGCTGGAGCAGGTGCTATGTCCCCT-3′	
OCH8088	5′-GATCCCGCGGTGGAGAGATCTTCGTGAAGACCCTG-3′	
OCH8110	5′-GCTGAATTCAGCAGCTCTGAGACGGAGGAC-3′	

USA) at 4°C for 4 h. After washing the beads three times in 10 ml of washing buffer (50 mM HEPES, 150 mM NaCl, 10% glycerol, 0.1% Triton X-100, 1 mM DTT), the bound proteins were eluted with 6 ml of elution buffer (50 mM HEPES, 150 mM NaCl, 10% glycerol, 0.1% Triton X-100, 1 mM DTT, 30 mM reduced glutathione, pH 8.0), followed by 4 h of dialysis against storage buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 10% glycerol, 1 mM DTT) and overnight incubation with the catalytic core Usp2 (Usp2-cc) at 4°C. The Usp2-cc treatment yielded free ^{6His}Ub and MQIFVKTLTGK-GST, QIFVKTLTGK-GST, EIFVKTLTGK-GST, REIFVKTLTGK-GST, $^{MQ}Ub_{AA}$-GST, $^{Q}Ub_{AA}$-GST, $^{E}Ub_{AA}$-GST, or $^{RE}Ub_{AA}$-GST. To remove the free ^{6His}Ub molecules, the Usp2-cc-treated samples were incubated with Ni-NTA resins. Subsequently, the Ni-NTA bead-unbound proteins were re-dialyzed against the above storage buffer and further purified by gel filtration chromatography. The active fractions were collected, concentrated, and stored at −80°C.

Production of anti-^{RE}Ub

The synthetic peptide REIFVKTLTGKC was used for the production of anti-^{RE}Ub as the antigen. Rabbit polyclonal antisera to the peptide were raised by AbClon (Korea). Antibodies specific for REIFVKTLTGKC were "negatively" selected from antiserum by sequentially incubating 5 ml of clarified antiserum with 1 ml of Glutathione Sepharose beads that had been pre-incubated individually with 2 mg of the purified MQIFVKTLTGK-GST, QIFVKTLTGK-GST, and EIFVKTLTGK-GST. The flow-through of antisera was further "positively" selected by overnight incubation with REIFVKTLTGK-GST-conjugated SulfoLink Coupling Resin (Cat. No. 20401; Thermo Fisher Scientific, USA) at 4°C. Bound antibodies were eluted by adding 9 ml of 0.1 M glycine HCl (pH 3.5). The eluted antibodies were further incubated with Affi-Gel 10/15 (Cat. No. 1536098; Bio-Rad, USA) that were pre-conjugated with 30 mg of proteins in the extracts of CHY2014 (*ate1Δ*) cells. The unbound anti-^{RE}Ub was concentrated by Protein A-Sepharose affinity chromatography (Amicogen, Korea).

Immunoprecipitation of ^{RE}Ub and its conjugated proteins

The indicated yeast cells were cultured in 100 ml of YPD medium to A_{600} of ~1. Cells were harvested and washed twice with ice-cold phosphate-buffered saline (PBS). Cell pellets were resuspended in 2 ml of lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 10 mM N-ethylmaleimide [NEM, Cat. No. E3876; Sigma-Aldrich], 10% glycerol, 1% nonidet P40 substitute [NP-40, Cat. No. 11754599001; Sigma-Aldrich], and 1× protease inhibitor cocktail [Cat. No. P8215, Sigma-Aldrich]) and disrupted by bead beating using Mini Beadbeater-24 (BioSpec Products, USA), 10 times for 10 s each at 1 min intervals on ice. Cell extracts were precipitated by centrifugation at 13,000 × *g* for 15 min at 4°C. The supernatants were incubated with 1 μg of anti-^{RE}Ub for 6 h at 4°C and then with 5 μl of Dynabeads Protein A (Cat. No. 10001D; Thermo Fisher Scientific) for 1 h. After washing the beads three times in the lysis buffer, the bound proteins were eluted by boiling in 1× SDS sample buffer and fractionated using 4%-20% Mini-PROTEAN TGX precast protein gels (Cat. No. 4561094; Bio-Rad). Fractionated proteins were transferred into PVDF membranes (polyvinylidene difluoride, Cat. No. 88518; Thermo Fisher Scientific) and ^{RE}Ub-conjugated proteins were identified by anti-Ub (P4D1) (Cat. No. Sc-8017; Santa Cruz Biotechnology, USA).

Preparation of SILAC (stable isotope labeling by amino acids in cell culture) samples

S. cerevisiae CHY3188 (*arg4Δ lys2Δ*) and CHY6056 (*ate1Δ arg4Δ lys2Δ*) were grown at 30°C to A_{600} of ~1 in 300 ml of SC medium with 20 μg/ml of either light isotope-labeled $^{12}C_6{}^{14}N_4$-L-arginine (Cat. No. ULM-8347; Cambridge Isotope Laboratories, USA) and $^{12}C_6{}^{14}N_2$-L-lysine (Cat. No. ULM-8766; Cambridge Isotope Laboratories) or 20 μg/ml heavy isotope-labeled $^{13}C_6{}^{15}N_4$-L-arginine (Cat. No. CNLM-539-H; Cambridge Isotope Laboratories) and $^{13}C_6{}^{15}N_2$-L-lysine (Cat. No. CNLM-291-H; Cambridge Isotope Laboratories). The yeast strains were treated with 2 mM H_2O_2 for 1 h, followed by mixing an equal amount of each culture and subsequent centrifugation at 2,000 × *g* and 4°C for 10 min. Cell pellets were washed twice with ice-cold PBS, resuspended in 5 ml of lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 10 mM NEM, 10% glycerol, 1% NP-40, 1× protease inhibitor cocktail), and disrupted by bead beating using Mini Beadbeater-24, 10 times for 10 s each at 1 min intervals on ice. Cell lysates were centrifuged at 13,000 × *g* and 4°C for 15 min. The resulting supernatants were incubated with 2 μg of anti-^{RE}Ub at 4°C for 6 h before adding 10 μl of Dynabeads-Protein A for an additional 1 h of incubation. After washing three times with 1 ml of the yeast lysis buffer, the binding proteins were eluted by boiling in 30 μl of 1× SDS-polyacrylamide gel electrophoresis (PAGE)-sample buffer. The eluted proteins were fractionated by 4%-20% Mini-PROTEAN TGX precast protein gels (Cat. No. 4561096; Bio-Rad) and stained with Coomassie Brilliant Blue (CBB).

Liquid chromatography and tandem mass spectrometry (LC-MS/MS)

The CBB-stained protein bands of interest were excised, destained in 50 mM ammonium bicarbonate with 50% acetonitrile (ACN), and dehydrated with 100% ACN, followed by drying in a vacuum evaporator. The destained proteins were proteolyzed with trypsin by the in-gel digestion method (Kim et al., 2013). The tryptic peptides were reconstituted in 7 μl of 0.1% formic acid, and a 5 μl aliquot was injected into a reverse-phase EASY-Spray PepMap RSLC C18 LC column (0.075 mm inner diameter × 500 mm length) on the Eksigent NanoLC Ultra system with an integrated column heater at 40°C; the column was pre-equilibrated with 96% buffer A (0.1% formic acid in water) and 4% buffer B (0.1% formic acid in ACN). The peptides were eluted with a 4%-35% gradient of buffer B over 150 min and a 32%-80% gradient of buffer B over 40 min. The total ion acquisition run time was set to 190 min at a flow rate of 250 nl/min. A Q-Exactive mass spectrometer (Thermo Fisher Scientific) was operated in data-dependent acquisition mode for the entire analysis. Full scans (m/z 300-1600) were acquired at a resolution of 70,000 using an automatic gain control (AGC) target value of 1e6 and a maximum ion injection time of 30 ms. Tandem mass spectra were generated for up to 12 precursors by

high-energy collision dissociation using a normalized collision energy of 35%. The dynamic exclusion was set to 60 s. Fragment ions were detected in normal scan mode using an AGC target value of 5e4 and a maximum ion injection time of 120 ms. Source ionization parameters were as follows: spray voltage, 1.9 kV; capillary temperature, 275°C; and S-Lens RF Level, 50.

Proteomic analysis of SILAC-based LC-MS/MS data

The RAW files from Q-Exactive were directly conveyed to the proteomic data analysis program suite, Proteome Discoverer v2.2 (Thermo Fisher Scientific). We employed the adjusted exemplar workflow for the SILAC experiment and the basic Sequest HT search module. Parameters for the Sequest search module were as follows: enzyme, specifically trypsin; protein database, UniProt reference proteome database of yeast (UP000002311, released in 04/2019) plus modified cRAP contaminant database with accessions in the dust/contact category (https://www.thegpm.org/crap/); fragment mass tolerance, 0.05 Da; precursor mass tolerance, 20 ppm; dynamic modifications, acetylation of protein-N-term (+42.010565 Da), oxidation of methionine (+15.9949 Da), label: $^{13}C_6{}^{15}N_2$ at lysine (+8.014199 Da), label: $^{13}C_6{}^{15}N_4$ at arginine (+10.008269 Da), Gly–Gly at Lys (Kε-GG) or protein-N-term (+114.042927 Da); and static modification, carbamidomethylation of cysteine (+57.021464 Da). Next, the identified spectra were validated using Percolator module, which was set to use a concatenated target/decoy strategy with a target cut-off q-value of 0.01 at PSM (peptide spectrum match) level. To integrate chromatographic features of identified SILAC pairs, we used Minora Feature Detector module with maximal trace retention time window of 5 min. Peptide abundance from the feature detector module was normalized with total sum of the abundance values over all identified peptides. Each result of the replicates was treated separately.

Identification of mass spectrum matched to the Nt-peptide of ^{RE}Ub

The generated mass spectrum files were directly matched to the peptide sequence in the customized ^{RE}Ub protein fasta file using the MS-GF+ search algorithm (v20190418) (Kim and Pevzner, 2014). The ^{RE}Ub sequence is a virtual one that replaces the two N-terminal residues of Ub MQ with RE. The search engine settings were as follows: trypsin enzyme specificity with 1 as the number of tolerable termini; 20 ppm for MS1 tolerance; variable modifications: oxidation of methionine (+15.9949 Da), label: $^{13}C_6{}^{15}N_2$ at lysine (+8.014199 Da), label: $^{13}C_6{}^{15}N_4$ at arginine (+10.008269 Da); and fixed modification: carbamidomethylation of cysteine (+57.021464 Da). All spectra matched to the Nt-peptide of the ^{RE}Ub were manually inspected. For annotating peptide spectra, we used a freely accessible version of the Interactive Peptide Spectrum Annotator (IPSA) (Brademan et al., 2019).

In silico structure modeling of ^{MQ}Ub, ^{Q}Ub, ^{E}Ub, and ^{RE}Ub

The amino acid sequence of *S. cerevisiae* Ub (UniProt accession: P0CG63) was seeded as a template for the structure modeling of ^{MQ}Ub, ^{Q}Ub, and ^{E}Ub. In contrast, the ^{RE}Ub template was deduced from ^{MQ}Ub by modifying its Nt-MQ into Nt-RE. The three-dimensional structure models of these Nt-varied Ubs were predicted by ColabFold, which combines a protein homolog search MMseqs2 (Steinegger and Soding, 2017) with AlphaFold2 (Jumper et al., 2021). In this setting, the model of ^{MQ}Ub was used as a template for accurate comparison with those of ^{RE}Ub, ^{Q}Ub, and ^{E}Ub. The predicted models with the highest pLDDT (predicted Local Distance Difference Test) value were selected for the structure comparison. In the case of ^{RE}Ub, the second-ranked model was chosen because the top-ranked one exhibited a covalent bond of Nt-Arg and Ser19, which is most unlikely to form *in vivo*. Surface structure modeling of the Ub variants was performed by Chimera (Pettersen et al., 2004) and the hydrogen bonds (H bonds) were displayed by PyMOL (http://www.pymol.org).

RESULTS

Antibody to recognize Nt-arginylated ^{RE}Ub

To detect the hypothetical ^{RE}Ub (Fig. 2A), we produced rabbit anti-^{RE}Ub polyclonal antibody using a synthetic peptide, REIFVKTLTGKC, the 11-residue Nt-sequences of ^{RE}Ub and Cys for the KLH (keyhole limpet hemocyanin) conjugation. Dot immunoblotting analysis with the affinity-purified anti-^{RE}Ub revealed the specific binding of the antibody to the REIFVKTLTGK-GST (glutathione *S*-transferase) fusion, but not to MQIFVKTLTGK-GST (with the 11-residue Nt-sequence of ^{MQ}Ub), QIFVKTLTGK-GST (bearing the 10-residue Nt-sequence of ^{Q}Ub), or EIFVKTLTGK-GST (bearing the 10-residue Nt-sequence of ^{E}Ub) (Fig. 2B). The following single-letter abbreviations denote amino acid residues: M, Met; A, Ala; C, Cys; G, Gly; S, Ser; T, Thr; V, Val; N, Asn; Q, Gln; D, Asp; E, Glu; L, Leu; I, Ile; F, Phe; Y, Tyr; W, Trp; H, His; K, Lys; R, Arg; and P, Pro.

Moreover, the anti-^{RE}Ub antibody specifically recognized ^{RE}Ub$_{AA}$-GST (Ub$_{AA}$, a deubiquitylation-protective Ub variant with the mutation of C-terminal GG to AA), but not ^{MQ}Ub$_{AA}$-GST, ^{Q}Ub$_{AA}$-GST, or ^{E}Ub$_{AA}$-GST (Fig. 2C). In agreement with these dot blotting results, immunoblotting with anti-^{RE}Ub further confirmed the specific recognition of purified ^{RE}Ub but not of ^{MQ}Ub; the ^{RE}Ub and ^{MQ}Ub were produced by the *in vitro* deubiquitylation of purified N-terminally 6 His-tagged ^{His6}Ub-^{MQ}Ub and ^{His6}Ub-^{RE}Ub, respectively, using the Ub fusion technique (Catanzariti et al., 2004) (Figs. 2D and 2E).

^{MQ}Ub undergoes NME, Nt-deamination, and Nt-arginylation to produce ^{RE}Ub

We next attempted to determine the levels of endogenous ^{RE}Ub and ^{RE}Ub-linked protein species by immunoblotting with the purified anti-^{RE}Ub. Notably, however, ^{RE}Ub and ^{RE}Ub–protein conjugates in wild-type *S. cerevisiae* were barely detectable under normal growth conditions (data not shown), suggesting that their endogenous levels are most likely near or below the detection limit of the antibody.

Since ^{MQ}Ub is a putative substrate of NatB Nt-acetylase that targets MQ, MN, ME, and MD at N-termini (Nguyen et al., 2018; Ree et al., 2018), we conjectured that Nt-acetylation might inhibit the NME-triggered Nt-modifications of ^{MQ}Ub. To rule out this possibility, we performed anti-^{RE}Ub-based immunoblotting in the extracts from *naa20Δ* cells that lacked

a catalytic subunit of the NatB Nt-acetylase (Ree et al., 2018). Indeed, the high-molecular-weight ^{RE}Ub–protein adducts were detected in overnight-grown (stationary-phase) *naa20Δ* yeast cells (Fig. 2F, lane 4). As expected, the ^{RE}Ub and ^{RE}Ub–protein adducts were abolished in *nta1Δ naa20Δ* cells (lacking Nta1 Nt-amidase) and *ate1Δ naa20Δ* cells (lacking Ate1 arginyltransferase) (Nguyen et al., 2018), whereas total Ub and Ub-conjugated proteins were still detected (Fig. 2F). In sum, we conclude that, *in vivo*, ^{MQ}Ub can be converted into ^{RE}Ub through the consecutive reactions of NME, Nt-deamination, and subsequent Nt-arginylation (Fig. 2A).

Upregulation of ^{RE}Ub-protein adducts in the stationary-growth phase or by oxidative stress

To detect the ^{RE}Ub–protein adducts more effectively and sensitively, we performed anti-^{RE}Ub-based immunoprecipitation, followed by immunoblotting with a highly Ub-specific monoclonal anti-Ub (P4D1) antibody, in extracts from *naa20Δ* cells. The hypersensitive coimmunoprecipitation–immunoblotting assays revealed that those ^{RE}Ub–protein adducts were markedly induced in wild-type cells in the stationary-growth phase, but hardly at all in the exponential-growth phase (Fig. 2G; cf, lane 6 vs lane 5). In contrast, the ^{RE}Ub–protein adducts were not detected in *ate1Δ* cells regardless of their growth phase (Fig. 2G; lanes 7 and 8).

Ubr1 is as a key E3 Ub ligase of the Arg/N-degron pathway and directly detects the Nt-Arg of Nt-arginylated proteins via its UBR box (Choi et al., 2010). Interestingly, the levels of ^{RE}Ub and ^{RE}Ub-protein adducts were discernably decreased in *ubr1Δ naa20Δ* cells, compared with those in *naa20Δ* cells, indicating the involvement of Ubr1 in processing ^{RE}Ub (Fig. 2H; cf, lane 5 vs lane 6).

We also observed that ^{RE}Ub–protein adducts were strongly increased in wild-type cells in the presence of 2 mM H_2O_2, an oxidative stressor (Fig. 2I; lane 7). In agreement with results with the stationary-phase yeast, however, the ^{RE}Ub-conjugated proteins were hardly detected in *ate1Δ* cells, regardless of H_2O_2 treatment (Fig. 2I; lanes 6 and 8). Taken together, these results suggest that the ^{RE}Ub and ^{RE}Ub–protein adducts are upregulated in yeast cells in the stationary-growth phase or by oxidative stress in an Ate1-dependent manner.

Identification of ^{RE}Ub and ^{RE}Ub-interacting proteins

Given that oxidative stress increases the levels of ^{RE}Ub–protein conjugates, we next sought to identify the ^{RE}Ub-attached proteins by employing SILAC in the H_2O_2-treated wild-type cells and *ate1Δ* cells, followed by anti-^{RE}Ub-based coimmunoprecipitation and LC-MS/MS (Fig. 3A).

Out of 541 distinct proteins detected by LC-MS/MS, 83 proteins were identified as direct or indirect ^{RE}Ub-binding candidates in the H_2O_2-treated wild-type cells (Fig. 3B). STRING enrichment analysis (for protein–protein interaction networks) indicated the functional associations of these putative ^{RE}Ub-interacting proteins with a set of cellular processes including gene transcription, vesicle trafficking, DNA replication, RNA processing, and mitochondrial regulation (Fig. 3C).

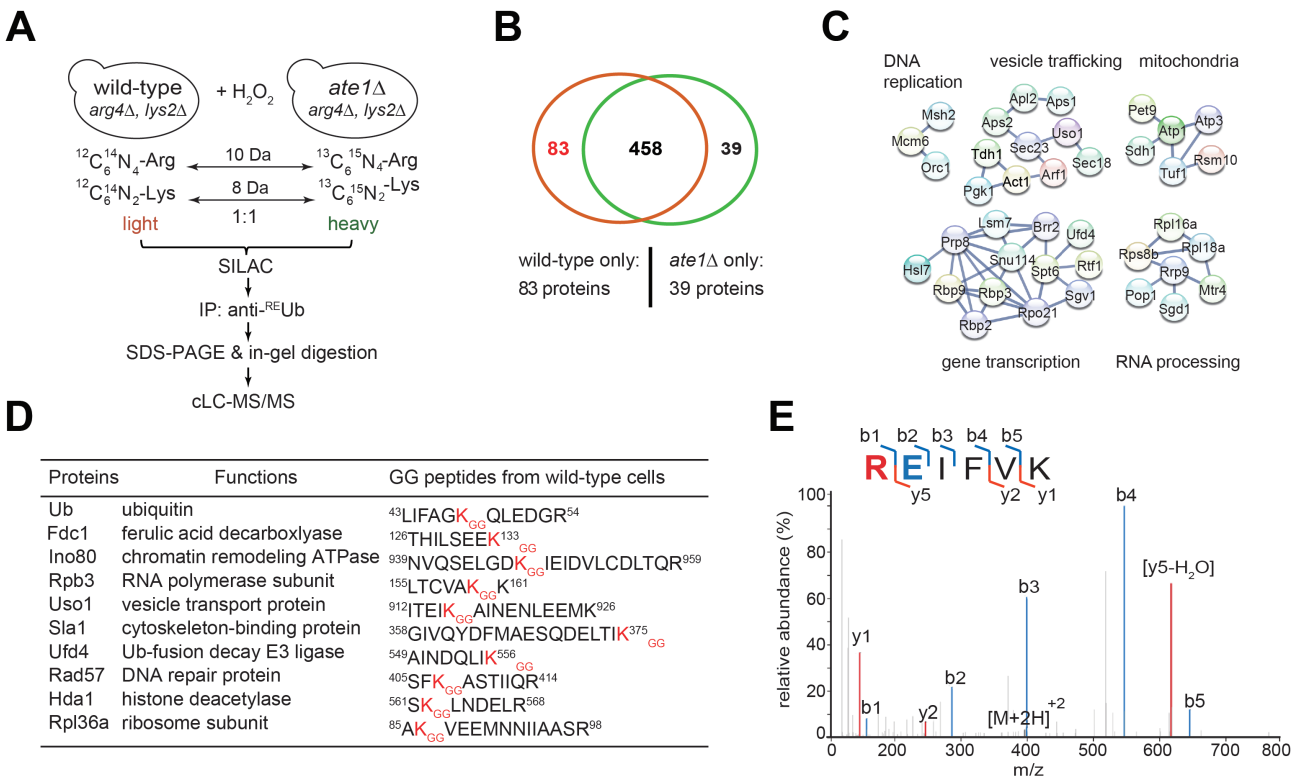
The tryptic digestion of Ub-conjugated proteins produces a K-ε-GG (K_{GG}) Ub isopeptide, which allows mapping of the ubiquitylation sites of substrates by LC-MS/MS (Kirkpatrick et al., 2005). Using this Ub-remnant K_{GG} signature-based approach, in the H_2O_2-treated yeast cells we identified the (putative) ^{RE}Ub-conjugated proteins and determined their ubiquitylation sites: Ub (^{43}LIFAGK$_{GG}$QLEDGR^{54}); Fdc1 (a ferulic acid decarboxylase, ^{126}THILSEEK$^{133}_{GG}$); Ino80 (a chromatin remodeling ATPase, ^{939}NVQSELGDK$_{GG}$IEIDVLCDLTQR^{959}); Rpb3 (an RNA polymerase subunit, ^{155}LTCVAK$_{GG}$K^{161}); Uso1 (a vesicle transport protein, ^{912}ITEIK$_{GG}$AINENLEEMK^{926}); Sla1 (a cytoskeleton-binding protein, ^{358}GIVQYDFMAESQDELTIK$^{375}_{GG}$); Ufd4 (a Ub-fusion decay E3 ligase, ^{549}AINDQLIK$^{556}_{GG}$); Rad57 (a DNA repair protein, ^{405}SFK$_{GG}$ASTIIQR^{414}); Hda1 (a histone deacetylase, ^{561}SK$_{GG}$LNDELR^{568}); and Rpl36a (a ribosome 60S subunit, ^{85}AK$_{GG}$VEEMNNIIAASR^{98}) (Fig. 3D). In this setting, however, it should be noted that ^{RE}Ub might be attached not only directly to the identified proteins but also indirectly to them by way of the Ub moiety of the already Ub-tagged proteins.

Strikingly, the exhaust manual survey of the tandem MS/MS database revealed the ^{RE}Ub-derived REIFVK Nt-peptide, verifying the actual existence of ^{RE}Ub *in vivo*. To the best of our knowledge, ^{RE}Ub and its conjugated proteins have not been documented until this study (Fig. 3E).

Nt-Arg of ^{RE}Ub is flexibly extended to the surface

To gain insight into the mechanistic role of ^{RE}Ub, we constructed its neural network-based 3D structure (for details, see Materials and Methods section). The overall surface structure of ^{RE}Ub was almost the same as that of ^{MQ}Ub, except its Nt-region (Figs. 4A and 4B). Specifically, the top and side views of the predicted models revealed that seven ubiquitylatable Lys residues and C-terminal Gly of ^{RE}Ub are located at almost the same surface positions as those of ^{MQ}Ub. Tellingly, however, Nt-Arg of ^{RE}Ub is stretched outward from the globular Ub protein, whereas ^{MQ}Ub retains its Nt-Met buried inward (Figs. 4A and 4B).

Given these findings, using the three-dimensional structure viewer programs we further estimated the H-bond acceptor–donor lengths of Nt-Met-Gln, Nt-Gln, Nt-Glu, or Nt-Arg-Glu of ^{MQ}Ub, ^{Q}Ub, ^{E}Ub, or ^{RE}Ub with their surrounding residues, respectively (Figs. 4C-4F) (see Materials and Methods section for details). In ^{MQ}Ub, Nt-Met forms two H bonds with Val17 (each 2.69 Å in length), and Gln at position 2 forms two H bonds with Glu16 (2.54 Å) and with Glu64 (2.56 Å) (Fig. 4C). In contrast, both Nt-Gln of ^{Q}Ub and Nt-Glu of ^{E}Ub form only one H bond with Glu64, resulting in H-bond acceptor–donor lengths of 2.77 Å and 2.67 Å, respectively (Figs. 4D and 4E). In the case of ^{RE}Ub, however, Nt-Arg forms two H bonds with Val17 (3.44 Å and 2.63 Å) and Glu at position 2 with Glu64 (2.66 Å). Notably, the H-bond acceptor–donor length between the amino group of Nt-Arg and the oxygen atom of Val17 in ^{RE}Ub was increased by 3.44 Å, relative to that in ^{MQ}Ub with 2.69 Å (Fig. 4F). These results suggest that Nt-Arg-Glu of ^{RE}Ub interacts more weakly with its surrounding residues Glu16, Val17, and Glu64 than Nt-Met-Gln of ^{MQ}Ub. Consequently, the Nt-Arg flexibly protruding from the Ub globule would increase the accessibility of ^{RE}Ub to the factors that recognize it.



Proteins	Functions	GG peptides from wild-type cells
Ub	ubiquitin	^{43}LIFAGK$_{GG}$QLEDGR^{54}
Fdc1	ferulic acid decarboxlyase	^{126}THILSEEK$^{133}_{GG}$
Ino80	chromatin remodeling ATPase	^{939}NVQSELGDK$_{GG}$IEIDVLCDLTQR^{959}
Rpb3	RNA polymerase subunit	^{155}LTCVAK$_{GG}$K^{161}
Uso1	vesicle transport protein	^{912}ITEIK$_{GG}$AINENLEEMK^{926}
Sla1	cytoskeleton-binding protein	^{358}GIVQYDFMAESQDELTIK$^{375}_{GG}$
Ufd4	Ub-fusion decay E3 ligase	^{549}AINDQLIK$^{556}_{GG}$
Rad57	DNA repair protein	^{405}SFK$_{GG}$ASTIIQR^{414}
Hda1	histone deacetylase	^{561}SK$_{GG}$LNDELR^{568}
Rpl36a	ribosome subunit	^{85}AK$_{GG}$VEEMNNIIAASR^{98}

Fig. 3. Proteomic analyses of ^{RE}Ub and ^{RE}Ub-protein conjugates. (A) Scheme for quantitative proteomic analysis employing SILAC, enrichment for ^{RE}Ub, ^{RE}Ub-protein conjugates, and their interacting proteins by coimmunoprecipations with anti-^{RE}Ub and LC-MS/MS. Wild-type *S. cerevisiae* (with *arg4Δ lys2Δ* genetic background) was cultured in a "light" medium ($^{12}C_6{}^{14}N_4$Arg, $^{12}C_6{}^{14}N_2$Lys), while the control *ate1Δ S. cerevisiae* was cultured in the otherwise identical "heavy" medium ($^{13}C_6{}^{15}N_4$Arg, $^{13}C_6{}^{15}N_2$Lys). Equal amounts of cells from light and heavy media were mixed prior to the preparation of cell extracts, followed by the enrichment for ^{RE}Ub or ^{RE}Ub-conjugated proteins, protease digestion, and LC-MS/MS. (B) Venn diagram indicating the numbers of specific ^{RE}Ub-associated proteins identified by LC-MS/MS analyses in the wild-type versus *ate1Δ S. cerevisiae*. (C) STRING enrichment analysis of the identified ^{RE}Ub-associated proteins. (D) The identified ^{RE}Ub and ^{RE}Ub-attached proteins in *S. cerevisiae*. (E) MS/MS spectrum of ^{RE}Ub identified from wild-type *S. cerevisiae*.

DISCUSSION

In this study, we produced for the first time anti-^{RE}Ub-specific antibody and detected the previously unrecognized ^{RE}Ub or ^{RE}Ub-conjugated proteins in *S. cerevisiae* (Figs. 2B-2E). The scarce ^{RE}Ub and its conjugated proteins were strongly induced in the stationary-growth phase or by oxidative stress. Owing to the scarcity of ^{RE}Ub *in vivo*, its physiological roles and regulation can be confined to the localized pools of Ub that might suddenly arise in response to specific stresses, such as nutrient starvation or reactive oxygen species (Figs. 2G and 2I).

Using anti-^{RE}Ub-based immunoprecipitation–tandem mass analysis, this study provides direct evidence of the existence of ^{RE}Ub *in vivo* and identifies the ^{RE}Ub-conjugated proteins, suggesting the possible involvement of ^{RE}Ub in epigenetic regulation, DNA repair, and protein degradation (Fig. 3). However, further detailed analyses are needed to define the functions of ^{RE}Ub in these cellular processes.

We found that ^{RE}Ub can be produced *in vivo* via NME of ^{MQ}Ub, Nt-deamination of ^{Q}Ub, and subsequent Nt-arginylation of ^{E}Ub (Figs. 2A and 2F). These Nt-modified Ub variants ^{Q}Ub, ^{E}Ub, and ^{RE}Ub have different physicochemical properties (charge, hydrophobicity, topology, etc.) of their Nt-region (Fig. 4), possibly impeding the ubiquitylation processes of the E1/E2/E3 cascade reactions, the Ub-chain assembly and disassembly, and their interplay with other Ub modifications for versatile biological processes and outcomes. Specifically, Nt-modifications of Ub can directly block the Met1-mediated linear polyubiquitylation, which is important for immune response, cell death, and proteostasis (Iwai et al., 2014).

The actual X-ray crystal structure of ^{MQ}Ub (PDB:4XOL, resolution 2.91 Å) reveals that the B factors of ^{MQ}Ub, Met-1, and Gln-2 are ~88.1 Å^2, ~78.1 Å^2, and ~90.5 Å^2, respectively, reflecting the rigid property of Nt-Met-Gln region in ^{MQ}Ub. Accordingly, the relatively flexible exposure of the Nt-Arg moiety of ^{RE}Ub to its surface in comparison to Nt-Met of ^{MQ}Ub (Figs. 4A-4C and 4F) may provide a binding site for the factors that recognize it, such as Ubr1 that targets basic destabilizing Nt-residues including Nt-Arg for the Arg/N-degron pathway (Nguyen et al., 2019; Varshavsky, 2019).

Ubr1 contains at least three substrate-binding sites for basic (type-1) Nt-residues, bulky hydrophobic (type-2) Nt-residues, and internal degrons (Nguyen et al., 2019; Varshavsky, 2019). ^{RE}Ub is highly upregulated by stationary-growth phase

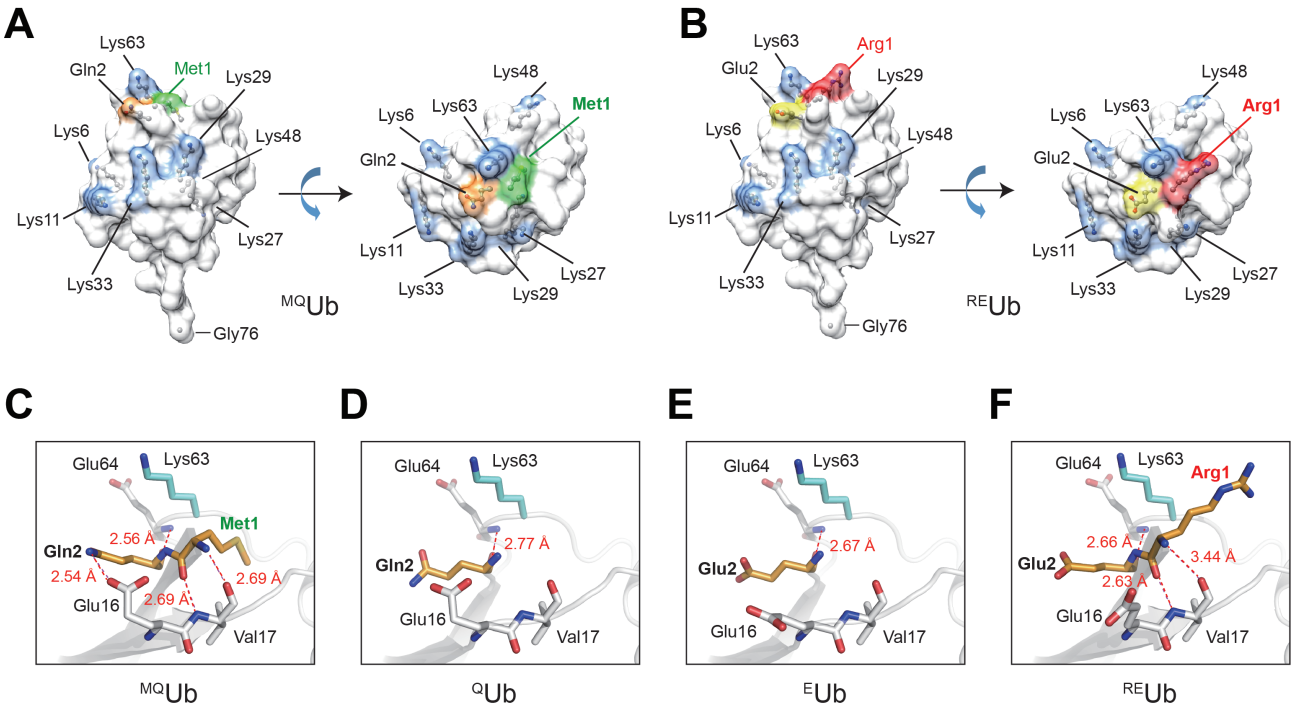


Fig. 4. Three-dimensional structure modeling of ^{MQ}Ub, ^{Q}Ub, ^{E}Ub, and ^{RE}Ub. (A) Side and top views of ^{MQ}Ub surface model. (B) Same as in (A) but with ^{RE}Ub. Nt-Met in ^{MQ}Ub (green), Nt-Arg in ^{RE}Ub (red), Gln2 in ^{MQ}Ub (orange), Glu2 in ^{RE}Ub (yellow), and seven Lys residues (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, and Lys63; blue). (C-F) Predicted locations of Nt-residues of ^{MQ}Ub, ^{Q}Ub, ^{E}Ub, and ^{RE}Ub. Nt-Met-Gln of ^{MQ}Ub (C), Nt-Gln of ^{Q}Ub (D), Nt-Glu of ^{E}Ub (E), and Nt-Arg-Glu of ^{RE}Ub (F) are indicated as sticks with bright orange on the transparent white-colored Ub surface cartoon. The residues that form H bonds with those Nt-residues are indicated as white-colored sticks (Glu16, Val17, and Glu64) and the H bonds are dotted in red. Rear-sighted Lys63 is also indicated as cyan sticks. The amino group and oxygen atom are in blue and red, respectively.

and oxidative stress that most likely induces Ubr1 activation for protein quality control (Szoradi et al., 2018). If so, is Ubr1 mechanistically or functionally linked to the metabolic regulation of ^{RE}Ub under stress conditions and vice versa? As one of parsimonious functions, Ubr1 might mediate the degradation of the apparently faulty ^{RE}Ub or its derivatives which can perturb the operation of normal Ub molecule. However, such possibility seems to be low, because ^{RE}Ub and ^{RE}Ub–protein adducts were rather downregulated in the absence of Ubr1 (Fig. 2H). Conversely, the Nt-Arg-bearing ^{RE}Ub might occupy the type-1 substrate-binding site of Ubr1, thereby modulating the degradation of a wide range of its substrates, similarly to the hydrophilin protein Roq1 (Szoradi et al., 2018). The precise roles of Ubr1 in ^{RE}Ub regulation and the mechanism behind them remain to be examined.

While no specific enzymes for ^{MQ}Ub NME have been identified thus far, Met-aminopeptidases (Map1 or Map2) in yeast most likely remove Nt-Met of ^{MQ}Ub, as in the case of other MQ- or MN-starting proteins (Nguyen et al., 2019). However, the possibility that other dedicated proteases might mediate the NME of ^{MQ}Ub under specific conditions including oxidative stress cannot be ruled out.

The identification of ^{RE}Ub in this study would enormously increase the repertoire and complexity of the Ub code. In addition, ^{RE}Ub would be present and actively operate in multicellular organisms as well because NME, Nt-deamination, and Nt-arginylation are conserved in nearly all eukaryotes (Sriram et al., 2011; Varshavsky, 2019).

ACKNOWLEDGMENTS

We are grateful to the current and former members of the Hwang and Lee laboratories for their assistance and advice. We also thank Edanz (www.edanz.com/ac) for editing a draft of this manuscript. This work was supported by grants from the Korean Government (MSIP) NRF-2020R1A3B2078127 and NRF2017R1A5A1015366 (C.-S.H.) and NRF-2020 R1A2C2003685 (C.L.), and the BK21 Plus program (C.-S.H.).

AUTHOR CONTRIBUTIONS

C.-S.H., C.L., K.T.N., and S.J. designed the research. K.T.N., S.J., S.-Y.K., and C.-S.L. performed the research and all coauthors analyzed the data. C.-S.H., C.L., K.T.N., S.J., and S.-Y.K. wrote the paper.

CONFLICT OF INTEREST

The authors have no potential conflicts of interest to disclose.

ORCID

Kha The Nguyen https://orcid.org/0000-0001-9191-8754
Shinyeong Ju https://orcid.org/0000-0001-5483-4690

Sang-Yoon Kim https://orcid.org/0000-0002-6195-3687
Chang-Seok Lee https://orcid.org/0000-0001-5249-3466
Cheolju Lee https://orcid.org/0000-0001-8482-4696
Cheol-Sang Hwang https://orcid.org/0000-0002-0105-5957

REFERENCES

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.

Brademan, D.R., Riley, N.M., Kwiecien, N.W., and Coon, J.J. (2019). Interactive Peptide Spectral Annotator: a versatile web-based tool for proteomic applications. Mol. Cell. Proteomics *18*(8 suppl 1), S193-S201.

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.

Chen, L. and Kashina, A. (2021). Post-translational modifications of the protein termini. Front. Cell Dev. Biol. *9*, 719590.

Chen, S.J., Kim, L., Song, H.K., and Varshavsky, A. (2021). Aminopeptidases trim Xaa-Pro proteins, initiating their degradation by the Pro/N-degron pathway. Proc. Natl. Acad. Sci. U. S. A. *118*, e2115430118.

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.

Choi, W.S., Jeong, B.C., Joo, Y.J., Lee, M.R., Kim, J., Eck, M.J., and Song, H.K. (2010). Structure basis for the recognition of N-end rule substrates by the UBR box of ubiquitin ligases. Nat. Struct. Mol. Biol. *17*, 1175-1181.

Dittmar, G. and Selbach, M. (2017). Deciphering the ubiquitin code. Mol. Cell *65*, 779-780.

Dohmen, R.J., Stappen, R., McGrath, J.P., Forrova, H., Kolarov, J., Goffeau, A., and Varshavsky, A. (1995). An essential yeast gene encoding a homolog of ubiquitin-activating enzyme. J. Biol. Chem. *270*, 18099-18109.

Gietz, R.D. and Schiestl, R.H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protoc. *2*, 31-34.

Giglione, C., Boularot, A., and Meinnel, T. (2004). Protein N-terminal methionine excision. Cell. Mol. Life Sci. *61*, 1455-1474.

Hwang, C.S., Shemorry, A., and Varshavsky, A. (2010). N-terminal acetylation of cellular proteins creates specific degradation signals. Science *327*, 973-977.

Iwai, K., Fujita, H., and Sasaki, Y. (2014). Linear ubiquitin chains: NF-kappaB signalling, cell death and beyond. Nat. Rev. Mol. Cell Biol. *15*, 503-508.

Janke, C., Magiera, M.M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E., et al. (2004). A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast *21*, 947-962.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583-589.

Kim, H.K., Kim, R.R., Oh, J.H., Cho, H., Varshavsky, A., and Hwang, C.S. (2014). The N-terminal methionine of cellular proteins as a degradation signal. Cell *156*, 158-169.

Kim, J.M. and Hwang, C.S. (2014). Crosstalk between the Arg/N-end and Ac/N-end rule. Cell Cycle *13*, 1366-1367.

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*, eaat0174.

Kim, K., Park, S.J., Na, S., Kim, J.S., Choi, H., Kim, Y.K., Paek, E., and Lee, C. (2013). Reinvestigation of aminoacyl-tRNA synthetase core complex by affinity purification-mass spectrometry reveals TARSL2 as a potential member of the complex. PLoS One *8*, e81734.

Kim, S. and Pevzner, P.A. (2014). MS-GF+ makes progress towards a universal database search tool for proteomics. Nat. Commun. *5*, 5277.

Kirkpatrick, D.S., Denison, C., and Gygi, S.P. (2005). Weighing in on ubiquitin: the expanding role of mass-spectrometry-based proteomics. Nat. Cell Biol. *7*, 750-757.

Komander, D. and Rape, M. (2012). The ubiquitin code. Annu. Rev. Biochem. *81*, 203-229.

Kwon, Y.T. and Ciechanover, A. (2017). The ubiquitin code in the ubiquitin-proteasome system and autophagy. Trends Biochem. Sci. *42*, 873-886.

Lee, K.E., Heo, J.E., Kim, J.M., and Hwang, C.S. (2016). N-terminal acetylation-targeted N-end rule proteolytic system: the Ac/N-end rule pathway. Mol. Cells *39*, 169-178.

Lim, K.L., Chew, K.C., Tan, J.M., Wang, C., Chung, K.K., Zhang, Y., Tanaka, Y., Smith, W., Engelender, S., Ross, C.A., et al. (2005). Parkin mediates nonclassical, proteasomal-independent ubiquitination of synphilin-1: implications for Lewy body formation. J. Neurosci. *25*, 2002-2009.

Mattiroli, F. and Penengo, L. (2021). Histone ubiquitination: an integrative signaling platform in genome stability. Trends Genet. *37*, 566-581.

Nguyen, K.T., Kim, J.M., Park, S.E., and Hwang, C.S. (2019). N-terminal methionine excision of proteins creates tertiary destabilizing N-degrons of the Arg/N-end rule pathway. J. Biol. Chem. *294*, 4464-4476.

Nguyen, K.T., Mun, S.H., Lee, C.S., and Hwang, C.S. (2018). Control of protein degradation by N-terminal acetylation and the N-end rule pathway. Exp. Mol. Med. *50*, 91.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605-1612.

Redman, K. and Rubenstein, P.A. (1981). NH2-terminal processing of Dictyostelium discoideum actin in vitro. J. Biol. Chem. *256*, 13226-13229.

Ree, R., Varland, S., and Arnesen, T. (2018). Spotlight on protein N-terminal acetylation. Exp. Mol. Med. *50*, 90.

Sadis, S., Atienza, C., Jr., and Finley, D. (1995). Synthetic signals for ubiquitin-dependent proteolysis. Mol. Cell. Biol. *15*, 4086-4094.

Shemorry, A., Hwang, C.S., and Varshavsky, A. (2013). Control of protein quality and stoichiometries by N-terminal acetylation and the N-end rule pathway. Mol. Cell *50*, 540-551.

Sherman, F. (2002). Getting started with yeast. Methods Enzymol. *350*, 3-41.

Sriram, S.M., Kim, B.Y., and Kwon, Y.T. (2011). The N-end rule pathway: emerging functions and molecular principles of substrate recognition. Nat. Rev. Mol. Cell Biol. *12*, 735-747.

Steinegger, M. and Soding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat. Biotechnol. *35*, 1026-1028.

Swatek, K.N. and Komander, D. (2016). Ubiquitin modifications. Cell Res. *26*, 399-422.

Szoradi, T., Schaeff, K., Garcia-Rivera, E.M., Itzhak, D.N., Schmidt, R.M., Bircham, P.W., Leiss, K., Diaz-Miyar, J., Chen, V.K., Muzzey, D., et al. (2018). SHRED is a regulatory cascade that reprograms Ubr1 substrate specificity for enhanced protein quality control during stress. Mol. Cell *70*, 1025-1037.e5.

Tasaki, T., Sriram, S.M., Park, K.S., and Kwon, Y.T. (2012). The N-end rule pathway. Annu. Rev. Biochem. *81*, 261-289.

Varshavsky, A. (2019). N-degron and C-degron pathways of protein degradation. Proc. Natl. Acad. Sci. U. S. A. *116*, 358-366.