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

Kha The Nguyen¹,⁴, Shinya Ju²,⁴, Sang-Yoon Kim¹, Chang-Seok Lee¹, Cheolju Lee²,³,*, and Cheol-Sang Hwang¹,*

¹Department of Life Sciences, Pohang University of Science and Technology, Pohang 37673, Korea, ²Center for Theragnosis, Korea Institute of Science and Technology, Seoul 02792, Korea, ³Division of Bio-Medical Science & Technology, KIST School, Korea University of Science and Technology, Seoul 02792, Korea, ⁴These 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-amino-peptidases (MetAPs), in cases where the residue at position 2 is Gly, Ala, Ser, Thr, Pro, or Val. Intriguingly, Nt-Met excision...
(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-terminal. The MQ-starting Ub (MQ-Ub3-76) is hereafter denoted simply as MQUb. Thus, we hypothesized that MQUb might be processed by these NME-provoked cascade reactions of Nt-deamination and Nt-arginylation to yield Nt-arginylated RE-Ub3-76 (denoted here as REUb) in yeast and mammalian cells (Figs. 1A and 2A). Indeed, we here identify the previously unknown REUb 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 Nt-acetylated residues. (D) The yeast Pro/N-degron pathway that targets for proteolysis Nt-Pro of proteins. (E) The yeast 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.
N-Terminal Modifications Expand the Ubiquitin Code  
Kha The Nguyen et al.

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 MQUbAA-GST, QUbAA-GST, EUbAA-GST, or REUbAA-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.

Fig. 2. Detection of Nt-arginylated REUb and its regulation. (A) A proposed model for the production of REUb in yeast cells in vivo. MQUb, wild-type Ub; ΔUb, Nt-Met-excised MQUb; 4Ub, Nt-deaminated 4Ub; 4Ub, Nt-arginylated 4Ub; NME, Nt-Met excision; MetAPs, Met-aminopeptidases: Nta1, Nt-amidase; Ate1, arginyltransferase. (B) Characterization of anti-REUb that was raised against an REUb Nt-peptide, REIFVKTLTGKC. Dot immunoblotting with affinity-purified anti-REUb, using indicated amounts of the spotted REIFVKTLTGK-GST fusion and its Nt-varied counterparts and MQUb. (C) Same as in (B) but with ΔUb, MQUb-GST, ΔUb, QUb-GST, ΔUb, EUb-GST, and ΔUb, REUb-GST. In ΔUb, C-terminal diglycine of Ub is mutated into dialanine. (D) A scheme for the production of MQUb and REUb. His6Ub-MQUb and His6Ub-REUb were purified from the extracts of E. coli. The Ni-NTA agarose-bound His6Ub-MQUb or His6Ub-REUb was deubiquitylated in vitro with the Usp2-cc deubiquitylating enzyme for the production of MQUb and REUb, respectively. (E) Immunoblotting of purified MQUb and REUb using anti-REUb or anti-Ub (P4D1) antibodies. CBB, Coomassie Brilliant Blue staining. (F) Immunoblotting of the extracts from the stationary-phase naa20ΔΔ cells, ate1Δnaa20ΔΔ cells, and nta1Δnaa20ΔΔ cells with anti-Ub (P4D1) or anti-REUb. (G) Extracts from the exponential-phase naa20ΔΔ cells, ate1Δnaa20ΔΔ cells, and nta1Δnaa20ΔΔ cells were immunoprecipitated (IP) with anti-REUb, followed by immunoblotting (IB) with anti-Ub (P4D1). (H) Same as in (G) but with the stationary-phase naa20ΔΔ cells, ate1Δnaa20ΔΔ cells, and ubr1Δnaa20ΔΔ cells. (I) Same as in (G) but wild-type (WT) cells and ate1Δ cells that were, respectively, cultured to the exponential phase (A600 ≈ 1) either in the absence or presence of 2 mM H2O2. In (G) and (I), * indicates non-specific bands.
N-Terminal Modifications Expand the Ubiquitin Code

Kha The Nguyen et al.

pairs: OCH7003/OCH8110, OCH8089/OCH8110, OCH8088/OCH8110, or OCH6968/OCH8110. The resulting PCR products were digested with SacII/EcoRI and triply ligated into SacII/HindIII-cut pHUE with EcoRI/HindIII-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 SacII/HindIII and ligated into SacII/HindIII-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 SacII/HindIII and ligated into SacII/HindIII-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 $^{65}_{65}$Ub-$^{65}_{65}$Ub$_{65}$-GST), pCH6914 (expressing $^{65}_{65}$Ub-$^{65}_{65}$Ub$_{65}$-GST), pCH6915 (expressing $^{65}_{65}$Ub-$^{65}_{65}$Ub$_{65}$-GST), pCH6916 (expressing $^{65}_{65}$Ub-$^{65}_{65}$Ub$_{65}$-GST), pCH7047 (expressing $^{65}_{65}$Ub-MQIFVKTLTGK-GST), pCH7048 (expressing $^{65}_{65}$Ub-QIFVKTLTGK-GST), pCH7049 (expressing $^{65}_{65}$Ub-EIFVKTLTGK-GST), and pCH7050 (expressing $^{65}_{65}$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, USA).

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

| Material | Description or sequence | Source |
|----------|-------------------------|--------|
| Yeasts   |                         |        |
| JD53     | MATα, his3Δ200, leu2-3, 112 lys2Δ801, trp1Δ63, ura3-52 | Dohmen et al., 1995 |
| CHY367   | naa20::NatMX4 in JD53  | Lab collection |
| CHY2009  | ubr1Δ::KanMX6, naa20Δ::NatMX4 in JD53 | Lab collection |
| CHY2014  | atel1Δ::HphNT2 in JD53 | Lab collection |
| CHY2015  | atel1Δ::HphNT2, naa20Δ::NatMX4 in JD53 | Lab collection |
| CHY2017  | ntel1Δ::HphNT2, naa20Δ::NatMX4 in JD53 | Lab collection |
| CHY3188  | arg4Δ::KnaMX6 in JD53  | Lab collection |
| CHY6056  | arg4Δ::KnaMX6, atel1Δ::HphNT2 in JD53 | This study |
| Plasmids |                         |        |
| pCH20    | pHUE                    | Catanzariti et al., 2004 |
| pCH33    | pGEX4T-3                | Lab collection |
| pCH5097  | $^{65}_{65}$Ub in pRK5  | Lim et al., 2005 |
| pCH6913  | $^{65}_{65}$Ub$_{65}$-GST in pHUE | This study |
| pCH6914  | $^{65}_{65}$Ub$_{65}$-GST in pHUE | This study |
| pCH6915  | $^{65}_{65}$Ub$_{65}$-GST in pHUE | This study |
| pCH6916  | $^{65}_{65}$Ub$_{65}$-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′-GGTAAGCTTCAAAGCGGAACCATCGGATTT-3′ |        |
| OCH6968  | 5′-GATCCCGCGGTGGAAGAACGATCTTGGAAGACCTCG-3′ |        |
| OCH6991  | 5′-GCTGAATTCATGTCCCCTATACTAGGTTATTGGAAAA-3′ |        |
| OCH7003  | 5′-GATCCCGCGGTGGAATGCAGATCTTCGTGAAGACCCTG-3′ |        |
| OCH8085  | 5′-GATCCCGCGGTGGAATGCAGATCTTCGTGAAGACCCTG-3′ |        |
| OCH8088  | 5′-GCTGAATTCATGTCCCCTATACTAGGTTATTGGAAAA-3′ |        |
| OCH8110  | 5′-GCTGAATTCATGTCCCCTATACTAGGTTATTGGAAAA-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 14N2-lysine, MQIFVKTLLGK-GST, QIFVKTLLGK-GST, EIFVKTLLGK-GST, and EIFVKTLLGK-GST, 15N2-lysine, or 15N2-lysine-GST. To remove the free 14N2-lysine molecules, the Usp2-cc-treated samples were incubated with Ni-NTA resin. Subsequently, the Ni-NTA bead-bound 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-15N2-lysine

The indicated yeast cells were cultured in 100 ml of YPD medium to A600 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], 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-15N2-lysine 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).

Immunoprecipitation of 15N2-lysine and its conjugated proteins

The indicated yeast cells were cultured in 100 ml of YPD medium to A600 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-15N2-lysine 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 15N2-lysine-conjugated proteins were identified by anti-lysine (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 A600 of ~1 in 300 ml of SC medium with 20 μg/ml of either light isotope-labeled 13C6-15N4-L-arginine (Cat. No. ULM-8347; Cambridge Isotope Laboratories, USA) and 13C6-15N4-L-lysine (Cat. No. ULM-8766; Cambridge Isotope Laboratories) or 20 μg/ml heavy isotope-labeled 13C6-15N4-L-arginine (Cat. No. CNLM-539-H; Cambridge Isotope Laboratories) and 13C6-15N4-L-lysine (Cat. No. CNLM-291-H; Cambridge Isotope Laboratories). The yeast strains were treated with 2 mM H2O2 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-lysine 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-Exact 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 Q-Exactive 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), deamidation of asparagine/serine (+0.984 Da), deamidation of glutamine (+0.984 Da), and phosphorylation of serine/threonine (+79.549 Da); and static modifications, acetylation of protein-N-term (+42.010565 Da), oxidation of methionine (+15.9949 Da), carbamidomethylation of cysteine (+57.021464 Da), acetylation of protein-N-term (+114.042927 Da), and static modification, carboxymethylation 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 using Percolator module, which was set to use a concatenated target/decoy strategy with a target cut-off q-value of 0.01.

Identification of mass spectrum matched to the Nt-peptide of Ub

The generated mass spectrum files were directly matched to the peptide sequence in the customized Ub protein fasta file using the MS-GF+ search algorithm (v20190418). (Kim and Pevzner, 2014). The 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: 13C6N2 at lysine (+8.014199 Da), label: 13C6N2 at arginine (+10.008269 Da), Gly-Gly at Lys (K-GG) or protein-N-term (+114.042927 Da); and static modification, carboxymethylation of cysteine (+57.021464 Da). Next, the identified spectra were validated using Percorolator 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.

In silico structure modeling of Ub, QUb, EUb, and REUb

The amino acid sequence of S. cerevisiae Ub (UniProt accession: P0CG63) was seeded as a template for the structure modeling of Ub, QUb, EUb, and REUb. In contrast, the REUb template was deduced from Ub by modifying its Nt-MQ into Nt-RE. The three-dimensional structure models of these Ub variants were predicted by CbFold, which combines a protein homolog search MMseq2 (Steinbegger and Soding, 2017) with AlphaFold2 (Jumper et al., 2021). In this setting, the model of Ub was used as a template for accurate comparison with those of QUb, EUb, and REUb. The predicted models with the highest plDDT (predicted Local Distance Difference Test) value were selected for the structure comparison. In the case of Ub, the second-rank model was chosen because the top-ranked one exhibited a covalent bond of Gly-Arg and Ser19, which is most unlikely to form in vitro. 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 Ub

To detect the hypothetical Ub (Fig. 2A), we produced rabbit anti-Ub polyclonal antibody using a synthetic peptide, REIFVKTLTGK-GST, the 11-residue Nt-sequences of Ub and Cys for the KLH (keyhole limpet hemocyanin) conjugation. Dot immunoblotting analysis with the affinity-purified anti-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-sequences of Ub), QIFVKTLTGK-GST (bearing the 10-residue Nt-sequences of Ub), or EIFVKTLTGK-GST (bearing the 10-residue Nt-sequences of 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, Glu; 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-Ub antibody specifically recognized UbAA-GST (UbAA, a ubiquitylation-protective Ub variant with the mutation of C-terminal GG to AA), but not to UbAA-GST, UbAA-GST, or UbAA-GST (Fig. 2C). In agreement with these dot blotting results, immunoblotting with anti-Ub further confirmed the specific recognition of purified Ub but not of Ub; the Ub and Ub were produced by the in vitro ubiquitylation of purified N-terminally 6 His-tagged Ub and Ub, respectively, using the Ub fusion technique (Catanzariti et al., 2004) (Figs. 2D and 2E).

Ub undergoes NME, Nt-deamination, and Nt-arginylation to produce Ub

We next attempted to determine the levels of endogenous Ub and Ub-linked protein species by immunoblotting with the purified anti-Ub. Notably, however, Ub and 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 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 Ub. To rule out this possibility, we performed anti-Ub-based immunoblotting in the extracts from naa20Δ cells that lacked

Mol. Cells 2022; 45(3): 158-167 163
a catalytic subunit of the NatB Nt-acetylase (Ree et al., 2018). Indeed, the high-molecular-weight $^{64}$Ub–protein adducts were detected in overnight-grown (stationary-phase) naa20Δ yeast cells (Fig. 2F, lane 4). As expected, the $^{64}$Ub and $^{64}$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 $^{64}$Ub through the consecutive reactions of NME, Nt-deamination, and subsequent Nt-arginylation (Fig. 2A).

**Upregulation of $^{64}$Ub–protein adducts in the stationary-growth phase or by oxidative stress**

To detect the $^{64}$Ub–protein adducts more effectively and sensitively, we performed anti-$^{64}$Ub-based immunoprecipitation, followed by immunoblotting with a highly Ub-specific monoclonal anti-Ub (P4D1) antibody, in extracts from naa20Δ cells. The hypersensitive immunoprecipitation-immunoblotting assays revealed that those $^{64}$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 $^{64}$Ub–protein adducts were not detected in ate1Δ cells regardless of their growth phase (Fig. 2G; lanes 7 and 8).

Ub1 is 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 $^{64}$Ub and $^{64}$Ub–protein adducts were discernably decreased in ubr1Δ naa20Δ cells, compared with those in naa20Δ cells, indicating the involvement of Ub1 in processing $^{64}$Ub (Fig. 2H; cf., lane 5 vs lane 6).

We also observed that $^{64}$Ub–protein adducts were strongly increased in wild-type cells in the presence of 2 mM H$_2$O$_2$, an oxidative stressor (Fig. 2I; lane 7). In agreement with results with the stationary-phase yeast, however, the $^{64}$Ub-conjugated proteins were hardly detected in ubr1Δ cells, regardless of H$_2$O$_2$ treatment (Fig. 2I; lanes 6 and 8). Taken together, these results suggest that the $^{64}$Ub and $^{64}$Ub–protein adducts are upregulated in yeast cells in the stationary-growth phase or by oxidative stress in an Ate1-dependent manner.

**Identification of $^{64}$Ub and $^{64}$Ub-interacting proteins**

Given that oxidative stress increases the levels of $^{64}$Ub–protein conjugates, we next sought to identify the $^{64}$Ub-conjugated proteins by employing SILAC in the H$_2$O$_2$-treated wild-type cells and ate1Δ cells, followed by anti-$^{64}$Ub-based immunoprecipitation and LC-MS/MS (Fig. 3A).

Out of 541 distinct proteins detected by LC-MS/MS, 83 proteins were identified as direct or indirect $^{64}$Ub-binding candidates in the H$_2$O$_2$-treated wild-type cells (Fig. 3B). STRING enrichment analysis (for protein–protein interaction networks) indicated the functional associations of these putative $^{64}$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-$\varepsilon$-GG ($K_g$G) Ub isopeptide, which allows mapping of the ubiquitylation sites of substrates by LC-MS/MS (Kirkpatrick et al., 2005). Using this Ub-remnant $K_g$ signature-based approach, in the H$_2$O$_2$-treated yeast cells we identified the (putative) $^{MQ}$Ub-conjugated proteins and determined their ubiquitylation sites: Ub ($^{64}$LIFAGK$_{G6}$QLGEDGR$^{57}$); Fdc1 (a ferulic acid decarboxylase, $^{64}$YILHESEEK$_{G6}$); Ino80 (a chromatin remodeler ATPase, $^{64}$NQVSELGDQK$_{G6}$IEIDVLCDLQR$^{90}$); Rpb3 (an RNA polymerase subunit, 586LTCAVK$_{G6}$K$^{161}$); Usol (a vesicle transport protein, $^{64}$ITEIK$_{G6}$AINENLEEMK$^{206}$); Sla1 (a cytoskeleton-binding protein, $^{64}$GIVQYDFMASEQDELTK$^{175}$); Ufd4 (a Ub-fusion decay E3 ligase, $^{54}$ASKDNLQK$_{G6}$); Rad57 (a DNA repair protein, $^{49}$SFK$_{G6}$ASTIQR$^{178}$); Hda1 (a histone deacetylase, $^{56}$SK$_{G6}$LNDLER$^{106}$); and Rpl36a (a ribosome 60S subunit, 85AK$_{G6}$VEEMNNIIAASR$^{99}$) (Fig. 3D). In this setting, however, it should be noted that $^{64}$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 $^{64}$Ub-derived REIFVK Nt-peptide, verifying the actual existence of $^{64}$Ub in vivo. To the best of our knowledge, $^{64}$Ub and its conjugated proteins have not been documented until this study (Fig. 3E).

**Nt-Arg of $^{64}$Ub is flexibly extended to the surface**

To gain insight into the mechanistic role of $^{64}$Ub, we constructed its neural network-based 3D structure (for details, see Materials and Methods section). The overall surface structure of $^{64}$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 $^{64}$Ub are located at almost the same surface positions as those of $^{MQ}$Ub. Tellingly, however, Nt-Arg of $^{64}$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-Glu-Glu of $^{MQ}$Ub, $^{64}$Ub, $^{MQ}$Ub, or $^{64}$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 Glu at position 2 forms two H bonds with Glu16 (2.54 Å, relative to 2.66 Å) and with Glu64 (2.56 Å) (Fig. 4C). In contrast, both Nt-Gln of $^{MQ}$Ub and Nt-Glu of $^{64}$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 $^{64}$Ub, however, Nt-Arg forms two H bonds with Val17 (3.44 Å in length), and Glu at position 2 forms two H bonds with Glu16 (2.54 Å) and with Glu64 (2.56 Å) (Fig. 4F). These results suggest that Nt-Glu of $^{MQ}$Ub interacts more weakly with its surrounding residues Glu16, Val17, and Glu64 than Nt-Met of $^{MQ}$Ub. Consequently, the Nt-Arg flexibly protruding from the Ub globule would increase the accessibility of $^{64}$Ub to the factors that recognize it.
**DISCUSSION**

In this study, we produced for the first time anti-REUb-specific antibody and detected the previously unrecognized REUb or REUb-conjugated proteins in *S. cerevisiae* (Figs. 2B-2E). The scarce REUb and its conjugated proteins were strongly induced in the stationary-growth phase or by oxidative stress. Owing to the scarcity of REUb 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-REUb-based immunoprecipitation–tandem mass analysis, this study provides direct evidence of the existence of REUb in vivo and identifies the REUb-conjugated proteins, suggesting the possible involvement of REUb in epigenetic regulation, DNA repair, and protein degradation (Fig. 3). However, further detailed analyses are needed to define the functions of REUb in these cellular processes.

We found that REUb can be produced in vivo via NME of MQUb, Nt-deamination of QUb, and subsequent Nt-arginylation of EUb (Figs. 2A and 2F). These Nt-modified Ub variants QUb, EUb, 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 **Ub** (PDB:4XOL, resolution 2.91 Å) reveals that the B factors of Met-1, Gln-2 are ~88.1 Å², ~78.1 Å², and ~90.5 Å², respectively, reflecting the rigid property of Nt-Met-Gln region in **Ub**. Accordingly, the relatively flexible exposure of the Nt-Arg moiety of REUb to its surface in comparison to Nt-Met of **Ub** (Figs. 4A-4C and 4F) may provide a binding site for the factors that recognize it, such as Ubr1 that targets basic de-stabilizing 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). REUb is highly upregulated by stationary-growth phase...
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 MQUb under stress conditions and vice versa? As one of parsimonious functions, Ubr1 might mediate the degradation of the apparently faulty MQUb or its derivatives which can perturb the operation of normal Ub molecule. However, such possibility seems to be low, because MQUb and MQUb–protein adducts were rather downregulated in the absence of Ubr1 (Fig. 2H). Conversely, the Nt-Arg-bearing MQUb 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 MQUb regulation and the mechanism behind them remain to be examined.

While no specific enzymes for MQUb NME have been identified thus far, Met-aminopeptidases (Map1 or Map2) in yeast most likely remove Nt-Met of MQUb, 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 MQUb under specific conditions including oxidative stress cannot be ruled out.

The identification of MQUb in this study would enormously increase the repertoire and complexity of the Ub code. In addition, MQUb 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-2020R1A2C2003685 (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
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.

Catanzari, 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, e21154301118.

Chen, S.J., Wu, X., Wadas, B., Oh, J.H., and Varshavsky, A. (2017). An N-end rule pathway that recognizes proline and destroys gluconegenic enzymes. Science 355, eaa63655.

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.

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., Pevzner, P.A., et al. (2018). 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., Tungyasunuvakool, 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 aminocar-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 Pevrzner, P.A. (2014). MS-GF+ makes progress towards a universal database search tool for proteomics. Nat. Commun. 5, 2577.

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-triggered ac/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., Engleander, 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.

Mattirol, F. and Penengo, L. (2021). Histone ubiquitination: an integrative signaling platform in genome stability. Trends Genet. 37, 556-581.

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.

Rhee, 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 alignment 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 of the Arg/N-end rule pathway. J. Biol. Chem. 293, 13226-13236.

Takagi, H., Su, Y., and Komander, D. (2020). UbQ-ONE: an open-source database of ubiquitin-conjugated proteins. Nucleic Acids Res. 48(D1), D478-D484.

Wagner, E. (2018). Remote homology detection in the era of big data. Nat. Rev. Genet. 19, 554-566.

Wang, J., Wang, W., Zhou, Y., Wang, P., and Zhao, X. (2018). Interactive Peptide Spectral Annotator: a versatile web-based tool for proteomic applications. Mol. Cell. Proteomics 18(8 suppl 1), S193-S201.