PERSPECTIVE

N-degron and C-degron pathways of protein degradation

Alexander Varshavsky[a,1]

Edited by F. Ulrich Hartl, Max Planck Institute of Biochemistry, Martinsried, Germany, and approved December 3, 2018 (received for review November 5, 2018)

This perspective is partly review and partly proposal. N-degrons and C-degrons are degradation signals whose main determinants are, respectively, the N-terminal and C-terminal residues of cellular proteins. N-degrons and C-degrons include, to varying extents, adjoining sequence motifs, and also internal lysine residues that function as polyubiquitylation sites. Discovered in 1986, N-degrons were the first degradation signals in short-lived proteins. A particularly large set of C-degrons was discovered in 2018. We describe multifunctional proteolytic systems that target N-degrons and C-degrons. We also propose to denote these systems as "N-degron pathways" and "C-degron pathways." The former notation replaces the earlier name "N-end rule pathways." The term "N-end rule" was introduced 33 years ago, when only some N-terminal residues were thought to be destabilizing. However, studies over the last three decades have shown that all 20 amino acids of the genetic code can act, in cognate sequence contexts, as destabilizing N-terminal residues. Advantages of the proposed terms include their brevity and semantic uniformity for N-degrons and C-degrons. In addition to being topologically analogous, N-degrons and C-degrons are related functionally. A proteolytic cleavage of a subunit in a multisubunit complex can create, at the same time, an N-degron (in a C-terminal fragment) and a spatially adjacent C-degron (in an N-terminal fragment). Consequently, both fragments of a subunit can be selectively destroyed through attacks by the N-degron and C-degron pathways.

degron | proteolysis | ubiquitin | proteasome | N-end rule

The lifespans of protein molecules in a cell range from less than a minute to many days. Regulated protein degradation protects cells from misfolded, aggregated, or otherwise abnormal proteins, and also controls the levels of proteins that evolved to be short-lived in vivo. Some proteolytic pathways can selectively destroy a specific subunit of a protein complex. Such pathways can act as protein-remodeling devices (1). They can either activate or inactivate a protein machine, change its enzymatic specificity, alter its subunit composition, or repair an oligomeric complex, for example, by destroying fragments of a cleaved subunit that are still embedded in the complex. This would allow a replacement of the cleaved subunit by its intact counterpart. Many biological transitions involve remodeling of protein complexes through subunit-selective degradation, in settings that range from cell-division cycles and circadian circuits to cell differentiation and responses to stresses.

One function of protein degradation is the quality control of nascent and newly formed proteins. Selective proteolysis eliminates those proteins (including mutant ones) that fold too slowly, misfold, or do not satisfy other requirements of quality control. Most proteins function as multisubunit complexes, which often assemble cotranslationally. Quality-control systems destroy subunits that are either overproduced relative to other subunits of a complex or do not become incorporated into a complex rapidly enough. The intracellular protein degradation is mediated largely by the ubiquitin (Ub)-proteasome system (UPS) and by autophagy-lysosome pathways, with molecular chaperones being a part of both systems (1–14).

The UPS comprises a set of pathways that have in common two classes of enzymes: E3-E2 Ub ligases and deubiquitylases (DUBs). A Ub ligase recognizes a substrate protein through its degradation signal (degron) (15) and conjugates Ub, a 9-kDa protein (usually in the form of a poly-Ub chain), to an amino acid residue (usually an internal lysine) of the targeted substrate (*SI Appendix*, Fig. S1*A*). DUBs deubiquitylate Ub-conjugated proteins and

[a]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125

Author contributions: A.V. designed research and wrote the paper.

The author declares no conflict of interest.

This article is a PNAS Direct Submission.

[1]Email: avarsh@caltech.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1816596116/-/DCSupplemental.

Published online January 7, 2019.

edit poly-Ub chains. In addition, DUBs produce free Ub through cleavages of Ub precursors encoded by Ub genes (1–9).

Most UPS pathways also involve a multisubunit, ATP-dependent protease called the 26S proteasome (*SI Appendix*, Fig. S1*A*). This protease binds to a ubiquitylated protein substrate through a substrate-linked poly-Ub chain, unfolds the protein using proteasome's ATPases (often with involvement of the Cdc48/p97 unfoldase), and processively destroys the protein to ∼10-residue peptides (16–22).

Some UPS pathways have nonproteolytic functions as well. A mammalian genome encodes more than 800 E3 Ub ligases, which target, in general, different degrons. The multitude and diversity of Ub ligases underlie the immense functional reach of UPS. Its pathways participate in just about every physiological process in all eukaryotes, play major roles in aging, and are involved in causation of many diseases, from impairments of immunity to cancer and neurodegeneration.

Terminology for Proteolytic Pathways That Target N-Termini and C-Termini

Degradation signals are features of proteins that make them short-lived in vivo. Such signals determine, in particular, the specificity of UPS. The problem of degradation signals preceded the onset of Ub studies in the 1980s, and remained a mystery until 1986, when the first degradation signals, later termed degrons (15), were discovered at the N-termini of short-lived proteins, an advance made possible by the invention of the Ub fusion technique (*SI Appendix*, Fig. S1*B*) (23–26). A set of these N-terminal (Nt) signals, later called N-degrons (15), was referred to by the term "N-end rule," which related the in vivo half-life of a protein to the identity of its Nt-residue (23–27). Studies over the next three decades identified proteolytic systems that recognize distinct classes of N-degrons and destroy, often conditionally, specific proteins or their natural fragments that bear N-degrons (Figs. 1 and 2 and *SI Appendix*, Fig. S2*A*) (22–41). Other studies, in the 1990s and afterward, have also identified many internal degrons, defined as degradation signals whose functionally essential elements do not include either Nt-residues or C-terminal (Ct) residues.

The first example of physiologically relevant Ct-degradation signals, called C-degrons below, was identified in 1996. A specific RNA (SsrA) can terminate, *in trans*, a stalled translation of a bacterial protein while tagging the released protein with the Ct-sequence ANDENYALAA. This segment acts as a C-degron, targeting a protein for degradation by the proteasome-like bacterial protease ClpXP (42). In 2018, the laboratories of Elledge and Yen discovered a remarkably large set of diverse natural C-degrons in human proteins (*SI Appendix*, Fig. S3) (43–45). While differing from N-degrons mechanistically and location-wise, C-degrons are topologically analogous to N-degrons. Specific C-degrons and N-degrons can also be associated functionally through their coformation upon a proteolytic cut, as described below.

In 1986, only some Nt-residues were thought to be destabilizing (23). However, later studies by our laboratory showed that every one of the 20 amino acids in the genetic code can act, in cognate sequence contexts, as a destabilizing Nt-residue of an N-degron (Figs. 1 and 2 and *SI Appendix*, Fig. S2*A*) (10, 23–26, 35, 38, 39, 46–48). The term N-end rule and its definition, cited above, are not commensurate with involvement of the entire gamut of Nt-residues in protein degradation. This understanding, as well as benefits of accurate notations, is the reason for renaming N-end rule pathways as "N-degron pathways." Advantages of this terminology include a recall of both the N-terminus ("N") and degradation ("degron"), and the ease of extending this notation from N-degrons to C-degrons.

In sum, proteolytic systems that target N-degrons are proposed to be called the Arg/N-degron pathway, the Pro/N-degron pathway, and the Ac/N-degron pathway in eukaryotes, the fMet/N-degron pathway in eukaryotes and bacteria, and the Leu/N-degron pathway in bacteria (Fig. 1 and *SI Appendix*, Fig. S2*A*). The prefixes Arg, Pro, Ac, fMet, and Leu specify each pathway by highlighting their unique features, for example, the step of Nt-arginylation as a part of the Arg/N-degron pathway (Fig. 1*G* and *SI Appendix*, Fig. S2*A*).

A Ub ligase of an N-degron pathway can contain several degron-recognizing sites. Such a ligase can bind not only to N-degrons but also to internal degradation signals in proteins that lack an N-degron (26). In the proposed terminology, a substrate of, for example, the Arg/N-degron pathway (Fig. 1*G* and *SI Appendix*, Fig. S2*A*) can be called an Arg/N-degron substrate or an Arg/N-d substrate. Another protein, recognized by the same Ub ligase through a protein's internal degron, can be denoted as an Arg/N-id substrate: that is, a substrate bearing an internal degron of the Arg/N-degron pathway.

N-Degron Pathways of Protein Degradation

The N-degron pathways (formerly "N-end rule pathways") comprise a set of proteolytic systems whose unifying feature is their ability to recognize proteins containing N-degrons, thereby causing the degradation of these proteins by the 26S proteasome or autophagy in eukaryotes and by the proteasome-like ClpAP protease in bacteria (Fig. 1 and *SI Appendix*, Figs. S2*A* and S4) (13, 23, 24, 26, 28, 30–38, 40, 46, 49–54). The main determinant of an N-degron is a destabilizing Nt-residue of a protein. In eukaryotes, an N-degron includes an internal lysine (or lysines) of a substrate protein that acts as the site of polyubiquitylation.

Initially, most N-degrons are pro–N-degrons. They are converted to N-degrons either constitutively (e.g., during the emergence of a protein from a ribosome) or conditionally, via regulated steps. Among the routes to N-degrons are cleavages of proteins by proteases that can expose a destabilizing Nt-residue (29, 55–57). An exopeptidase, for example the mammalian Dpp9 aminopeptidase (it removes dipeptides from N-termini), can convert a pro–N-degron at the N-terminus of a specific protein, such as the Syk kinase, to an N-degron (58). The Dpp9 aminopeptidase, Met-aminopeptidases (they remove Nt-Met from some nascent proteins) (*SI Appendix*, Fig. S1*D*), and endoproteases that include caspases, separases, calpains, and cathepsins, have all been shown to generate N-degrons in vivo through their cleavages of intracellular proteins (26, 29, 55–59). Operationally, these proteases are components of N-degron pathways.

A different and mutually nonexclusive route to N-degrons is through enzymatic Nt-modifications of proteins, including Nt-acetylation, Nt-deamidation, Nt-arginylation, Nt-leucylation, and Nt-formylation of the α-amino groups of Nt-residues (Fig. 1 and *SI Appendix*, Fig. S2*A*). Recognition components of N-degron pathways are called N-recognins. They are either specific E3 Ub ligases or other proteins that can target N-degrons (Figs. 1 and 2 and *SI Appendix*, Figs. S2*A* and S4). All 20 amino acids of the genetic code can act, in cognate sequence contexts, as destabilizing Nt-residues (Fig. 1*A*). Consequently, many proteins in a cell are conditionally short-lived N-degron substrates, either as full-length proteins or as protease-generated Ct-fragments (29, 35, 47, 55–57).

Selective degradation of proteins or their natural fragments by N-degron pathways has been shown to regulate a multitude of processes, including: the sensing of oxygen, nitric oxide (NO), heme, and short peptides; the control of subunit stoichiometries in protein complexes; the elimination of misfolded or otherwise abnormal proteins; the degradation of proteins that are retrotranslocated to the cytosol from other compartments, such as mitochondria; the regulation of apoptosis and repression of neurodegeneration; the regulation of DNA repair, transcription, replication, and chromosome cohesion/segregation; the regulation of G proteins, cytoskeletal proteins, autophagy, gluconeogenesis, peptide transport, meiosis, immunity, circadian rhythms, fat metabolism, cell migration, cardiovascular development, spermatogenesis, and neurogenesis; and the regulation of leaf and shoot development, oxygen/NO sensing, and many other

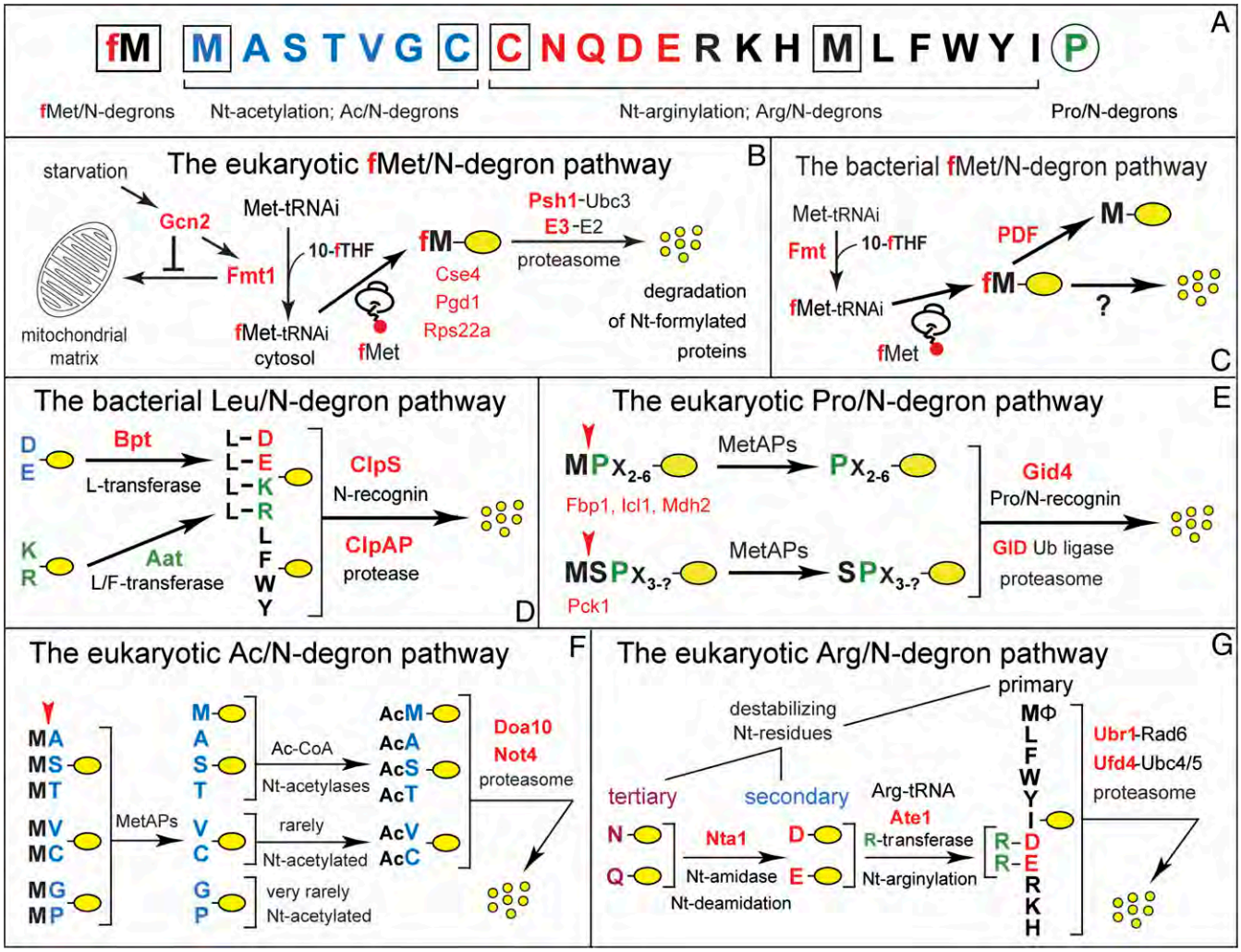

Fig. 1. N-degron pathways. Nt-residues are indicated by single-letter abbreviations. A yellow oval denotes the rest of a protein substrate. (*A*) Twenty amino acids of the genetic code are arranged to delineate specific N-degrons. Nt-Met is cited three times because it can be recognized by the Ac/N-degron pathway (as Nt-acetylated Ac-Met), by the Arg/N-degron pathway (as unacetylated Nt-Met), and by the fMet/N-degron pathway (as Nt-formylated fMet). Nt-Cys is cited twice, because it can be recognized by the Ac/N-degron pathway (as Nt-acetylated Cys) and by the Arg/N-degron pathway (as an oxidized, arginylatable Nt-Cys sulfinate or sulfonate, formed in multicellular eukaryotes but apparently not in unstressed *S. cerevisiae*). (*B*) The eukaryotic (*S. cerevisiae*) fMet/N-degron pathway (39); 10-fTHF, 10-formyltetrahydrofolate. (*C*) The bacterial (*E. coli*) fMet/N-degron pathway (38). (*D*) The bacterial (*V. vulnificus*) Leu/N-end rule pathway (51). (*E*) The eukaryotic (*S. cerevisiae*) Pro/N-degron pathway (35–37). (*F*) The eukaryotic (*S. cerevisiae*) Ac/N-degron pathway (10, 46–48). (*G*) The eukaryotic (*S. cerevisiae*) Arg/N-degron pathway (26, 31). Modified with permission from ref. 38.

processes in plants (see refs. 26, 30–35, 48, 56, 57, and 59 and references therein).

The field of N-degrons and C-degrons is too large for a comprehensive review in a Perspective-size article. Instead of describing all pathways equally briefly, the Arg/N-degron pathway is discussed below in relative detail, followed by much shorter accounts of other pathways.

The Arg/N-Degron Pathway

This eukaryotic pathway targets unacetylated Nt-residues (Figs. 1*G* and 2*B* and *SI Appendix*, Figs. S2*A* and S4) (23, 55, 59). Nt-Arg, -Lys, -His, -Leu, -Phe, -Tyr, -Trp, -Ile, and -Met (if Nt-Met is followed by a bulky hydrophobic residue) are directly recognized by Arg/N-recognins. Examples of Arg/N-recognins include: the *Saccharomyces cerevisiae* Ubr1 E3; the mammalian Ubr1, Ubr2, Ubr4, and Ubr5 E3s; the Prt1 and Prt6 E3s of plants; and the mammalian non-E3 autophagy regulator p62/Sqstm1 (Fig. 1*G* and *SI Appendix*, Figs. S2*A* and S3) (13, 26, 32, 33, 47, 49, 50). The Nt-Asn, -Gln, -Glu, and -Asp residues (as well as Nt-Cys, under some conditions) are destabilizing because of enzymatic deamidation of Nt-Asn and -Gln, and Nt-arginylation of Nt-Asp, -Glu, and (oxidized) -Cys (Fig. 1*G* and *SI Appendix*, Fig. S2*A*) (40, 60–62).

Double-E3 Design of the Arg/N-Degron Pathway. Ubr1 is the sole Arg/N-recognin in *S. cerevisiae*, but the pathway's targeting complex contains two E3s: the 225-kDa RING-type Ubr1 and the 168-kDa HECT-type Ufd4, in association with their respective E2 enzymes Rad6 and Ubc4/Ubc5 (63) (Fig. 1*G*). The Ubr1-bound Ufd4 increases the processivity of polyubiquitylation (63). In contrast to Ubr1, Ufd4 is not an Arg/N-recognin. Specifically, Ufd4 does not, by itself, recognize Arg/N-degrons. However, Ufd4 can bind to substrate proteins such as Mgt1, Cup9, and Chk1, through their internal degrons that are also recognized by Ubr1 (63, 64). Exactly how the recognition of an internal degron by both Ubr1 and Ufd4 is achieved within the Ubr1–Ufd4 complex (do Ubr1 and Ufd4 compete for the same elements of a degron, or did

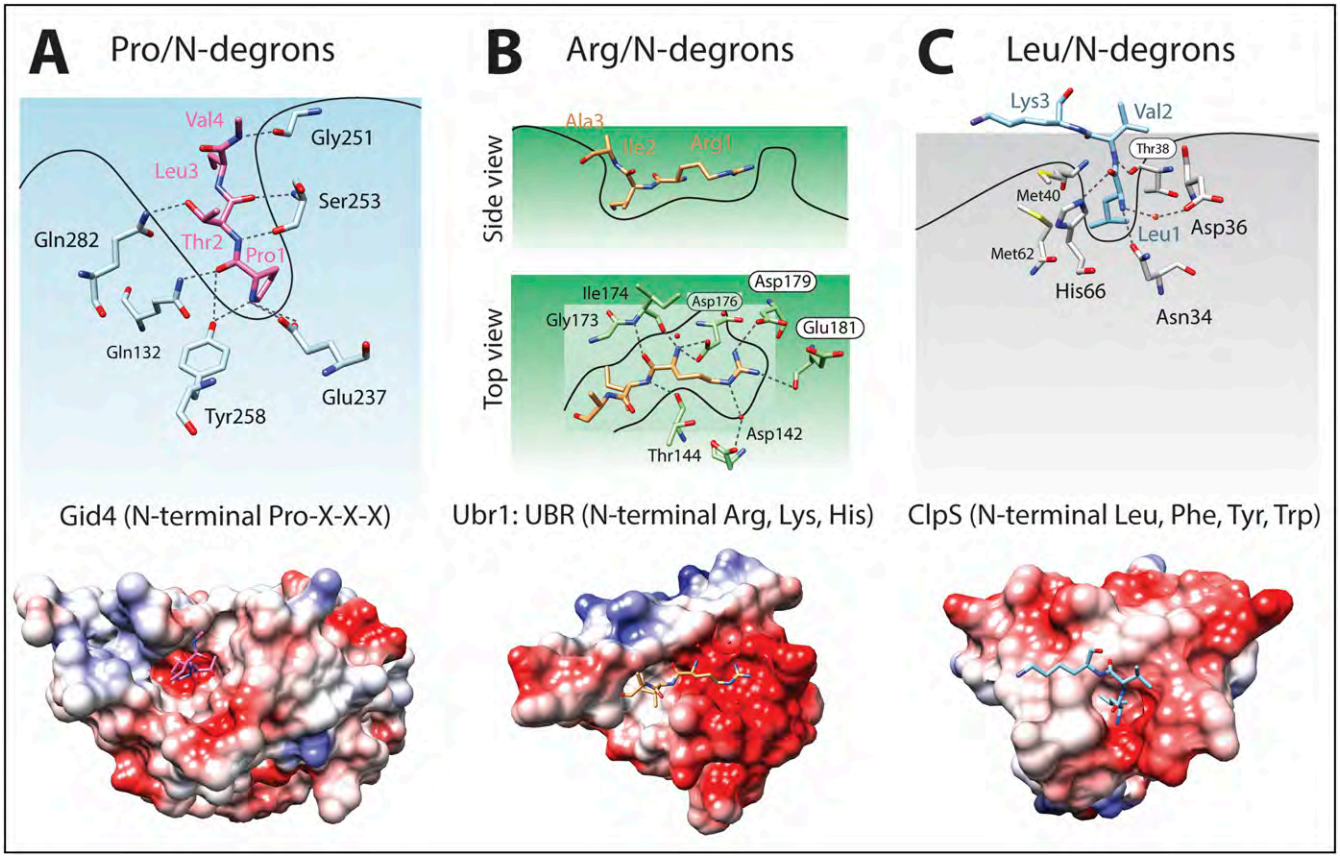

Fig. 2. Structural basis of N-degron recognition. The upper diagrams schematically depict the substrate-binding sites of different N-recognins, with corresponding space-filling images indicating electrostatic potential (red, negative; blue, positive) below the diagrams. (*A*) The substrate-binding site of the human Gid4 Pro/N-recognin (35–37). (*B*) One of substrate-binding sites (the UBR box, which recognizes basic Nt-residues) of the *S. cerevisiae* Ubr1 Arg/N-recognin. (*C*) The substrate-binding site of the *E. coli* ClpS Leu/N-recognin, which recognizes bulky hydrophobic Nt-residues (30, 52, 53). Modified with permission from ref. 35.

these E3s evolve to recognize a cognate degron simultaneously?) remains to be understood.

Substrates of the Arg/N-Degron Pathway and Their Protection by Chaperones. A molecule of a protein, including a newly formed protein, would be longer-lived if that molecule succeeds, rapidly enough, to become a subunit of a "protective" complex, often a cognate complex in which that subunit normally functions. Stabilization of the subunit would be caused by steric shielding of its degrons within the complex (10). A protection can also be attained through the binding of a vulnerable protein to a molecular chaperone, particularly the Hsp90 system, which comprises Hsp90 and more than 10 of its cochaperones. Hsp90 reversibly binds to at least 20% of cellular proteins, called Hsp90 clients, including most kinases and transcriptional regulators. Hsp90 assists its clients, often repeatedly, in maintaining their active conformations (14).

Oh et al. (64) showed that a weakening of the *S. cerevisiae* Hsp90 system (i.e., an increase in the fraction of Hsp90 clients that are not bound to Hsp90), causes many otherwise long-lived proteins to become short-lived, because of their rapid degradation by the Arg/N-degron pathway. Diverse Hsp90 clients, including Chk1, Kar4, Tup1, Gpd1, Ste11, and also, remarkably, Hsp82 [i.e., the Hsp90 chaperone itself (suggesting that Hsp90 is its own client)], become short-lived substrates of the Arg/N-degron pathway under conditions of hypoactive Hsp90 (64). The cited proteins are targeted by Ubr1/Ufd4 through their internal degrons (64). The Arg/N-degron pathway has also been shown to destroy a variety of misfolded proteins (reviewed in refs. 26, 31, and 34).

Mammalian Ubr1, Ubr2, Ubr4, and Ubr5. In contrast to *S. cerevisiae*, in which the Ubr1 E3 is the sole Arg/N-recognin (Figs. 1*G* and 2*B*), a mammalian genome encodes at least four E3s that can recognize Arg/N-degrons: the 200-kDa Ubr1 and Ubr2, the 570-kDa Ubr4 (p600; Big), and the 300-kDa Ubr5 (Edd1; Hyd) (31) (*SI Appendix*, Fig. S2*A*). Ubr1 and Ubr2 are highly sequelogous (similar in sequence) (65) to each other and to *S. cerevisiae* Ubr1 (26, 31). In contrast, the sequelogy (sequence similarity) (65) between, for example, Ubr1 and either Ubr4 or Ubr5, is largely confined to their ∼80-residue UBR domains. ["Sequelog" denotes a sequence that is similar, to a specified extent, to another sequence (65). Derivatives of sequelog include "sequelogous" (similar in sequence) and "sequelogy" (sequence similarity). The usefulness and appeal of sequelog and derivative terms stem from the rigor of their evolutionary neutrality. In contrast, the terms "homolog," "ortholog," and "paralog," which invoke, respectively, common descent and functional similarity or dissimilarity, are interpretation-laden and often less than precise notations. Homolog, ortholog, and paralog are compatible with the sequelog terminology, and can be used to convey understanding about common descent and biological functions, if this additional information (it is distinct from sequence similarities per se) is actually present (65).]

Ubr4, a huge (570 kDa) Arg/N-recognin, functions, in particular: in neurogenesis; in cell migration; in the biogenesis of endosomes; in cardiovascular development and autophagy; in the degradation of

podocin, a protein that maintains the renal filtration barrier; and in auxin transport in plants (ref. 66 and references therein). The functions of the Ubr5 Arg/N-recognin include: regulation of Wnt/β-catenin; the degradation of huntingtin, hPXR, Gkn1, and many other proteins; and also specific roles as either an oncoprotein or a tumor suppressor (ref. 67 and references therein). Connections between the functions of Ubr4/Ubr5 and their ability to recognize Arg/N-degrons remain to be understood.

Johanson-Blizzard Syndrome. Johanson-Blizzard Syndrome (JBS) patients lack Ubr1 but retain other Arg/N-recognins, including Ubr2, a sequelog (65), and functional analog of Ubr1 (*SI Appendix*, Fig. S2*A*). Symptoms of JBS include an exocrine pancreatic insufficiency and inflammation, multiple malformations (e.g., a near-absence of nasal wings), as well as mental retardation and deafness (refs. 26 and 68 and references therein). *Ubr1*$^{-/-}$ mice exhibit JBS symptoms in a milder form. Mice lacking Ubr2 have other defects, including infertility in males (because of apoptosis of spermatocytes) and genomic instability (31). In contrast to viability of *Ubr1*$^{-/-}$ and *Ubr2*$^{-/-}$ mouse strains, double-mutant mice, lacking both Ubr1 and Ubr2, die as midgestation embryos, with defects in neurogenesis and cardiovascular development (refs. 26 and 31 and references therein).

Structure and Targeting of Arg/N-Degrons. The main determinant of an Arg/N-degron is a substrate's specific Nt-residue (Fig. 1*G* and *SI Appendix*, Figs. S2*A* and S5). Once an Arg/N-recognin (as a part of a targeting complex) binds to a destabilizing Nt-residue of a substrate, a race against time begins, given the transience of the bound state and the necessity, for a successful targeting, to produce a substrate-linked poly-Ub chain. The synthesis of a proteasome-binding poly-Ub chain is initiated at an internal lysine of a substrate. This lysines is the second determinant of an Arg/N-degron (*SI Appendix*, Fig. S5) (1, 22, 24, 26, 31). The third determinant of an Arg/N-degron is an unstructured segment that the substrate-bound proteasome uses to initiate proteolysis (21, 22, 24, 26).

Subunit Selectivity of Protein Degradation. The Arg/N-degron pathway can destroy a subunit of a complex while sparing the rest of the complex (1, 26). Subunit selectivity, discovered in 1990, involved in addition the discovery of *trans*-targeting (*SI Appendix*, Fig. S5) (1). In this process, an Arg/N-recognin binds to a destabilizing Nt-residue of a subunit that lacks an efficacious second-determinant lysine. It was found that the subunit-bound Arg/N-recognin could polyubiquitylate *in trans* another subunit of the same complex (if it contained a "suitable" second-determinant lysine), and thereby would target for degradation specifically that subunit rather than the initially bound one. In sum, the multideterminant organization of an Arg/N-degron allows it to be "split" between subunits of a complex, leading to a targeting *in trans* (*SI Appendix*, Fig. S5) (1, 26).

Physiological Arg/N-Degron Substrates. The list of physiological Arg/N-degron substrates is already large and continues to grow (*SI Appendix*, Figs. S6–S8). An example of Arg/N-degron substrates that are not cited in *SI Appendix*, Figs. S6–S8 is Phe-Pink1, a Ct-fragment of the Pink1 kinase. Pink1 is imported into mitochondria and is conditionally cleaved there. The Youle laboratory (see ref. 69 for review) showed that the Phe-Pink1 Ct-fragment is retrotranslocated to the cytosol and is destroyed by the Arg/N-degron pathway. Pink1 phosphorylates, in particular, the E3 Ub ligase parkin and Ub itself. Null mutations in both copies of human *PINK1* result in early-onset Parkinson disease, a neurodegeneration syndrome. The uncleaved Pink1 accumulates in the outer mitochondrial membrane (OMM) and recruits parkin to OMM, a step that can lead to an autophagosome-mediated engulfment of damaged mitochondria and their destruction in lysosomes. Generation and degradation of the Phe-Pink1 Ct-fragment are a part of circuits that regulate the levels of OMM-bound uncleaved Pink1 and mitochondrial quality control (ref. 69 and references therein).

Roq1 as a Substrate and Regulator of the Arg/N-Degron Pathway. A natural Ct-fragment of *S. cerevisiae* Roq1 acts as both a substrate and regulator of the Arg/N-degron pathway (41). Tunicamycin, a drug that causes protein misfolding in the endoplasmic reticulum (ER), increases the level of *ROQ1* mRNA. An artificial increase of *ROQ1* mRNA can accelerate the degradation, by the Arg/N-degron pathway, of an ER membrane-embedded reporter protein. The Ynm3 endoprotease can cleave Roq1, generating the Arg-Roq1 Ct-fragment. This cleavage of Roq1 is required for the accelerated degradation of the above reporter. Arg-Roq1 is destroyed, in part, by the Arg/N-degron pathway (41). Remarkably, interactions between Arg-Roq1 and Ubr1 can alter the targeting efficacy of Ubr1 toward its other substrates, such as, for example, Cup9, which bears an internal degron. One possibility is that Arg-Roq1 modulates the specificity and efficacy of Ubr1 under conditions of stress (41).

Accelerators of Apoptosis as Arg/N-Degron Substrates. During apoptosis, caspases cleave more than 1,000 different proteins in a mammalian cell. Caspase-mediated cleavages of cellular proteins can generate proapoptotic Ct-fragments, defined as those that increase the probability of apoptosis. Such Ct-fragments often bear destabilizing Nt-residues (*SI Appendix*, Fig. S6). It was found that the natural proapoptotic Ct-fragments Cys-Ripk1, Cys-Traf1, Asp-Brca1, Leu-Likk1, Tyr-Nedd9, Arg-Bid, Asp-Bcl$_{XL}$, Arg-Bim$_{EL}$, Asp-Epha4, and Tyr-Met were short-lived substrates of the Arg/N-degron pathway (*SI Appendix*, Fig. S6) (refs. 34 and 55 and references therein). In agreement with these results, even a partial ablation of the Arg/N-degron pathway sensitizes cells to apoptosis (55). In sum, the Arg/N-degron pathway is a regulator of apoptosis, acting largely (but not necessarily exclusively) (34) as an antiapoptotic circuit (55). By destroying proapoptotic Ct-fragments, the Arg/N-degron pathway contributes to thresholds that prevent a transient or otherwise weak proapoptotic signal from reaching the point of commitment to apoptosis.

Ubr1 Binds to Caspases. Weaver et al. (70) showed that the Ubr1 Arg/N-recognin of the nematode *Caenorhabditis elegans* binds to both the procaspase Ced3 and its proteolytically activated form. One substrate of Ced3 is Lin28, a regulator of cell differentiation. Activated Ced3 cleaves Lin28, generating its Nt-Asn–bearing Ct-fragment that is rapidly destroyed by the Arg/N-degron pathway. In *ubr1Δ* worms the level of Lin28 was increased (as would be expected), but Lin28 was also at most weakly cleaved by Ced3 (70). The latter finding suggested that Ubr1 not only mediates the degradation of the caspase-generated Asn^{31}-Lin28, but may also activate the Ced3 procaspase. If so, the Arg/N-degron pathway might be a previously unknown route for activation of caspases, a most interesting possibility.

Regulation of Peptide Transport by the Arg/N-Degron Pathway. In the absence of extracellular di/tripeptides, the *S. cerevisiae* transcriptional repressor Cup9 shuts off (nearly but not entirely) the *PTR2* gene, which encodes the transmembrane peptide importer. This makes cells nearly (but not entirely) incapable of importing di/tripeptides (*SI Appendix*, Fig. S9). The type 1 and type 2 binding sites of Ubr1 recognize Arg/N-degrons through their binding, respectively, to basic and bulky hydrophobic Nt-residues in either proteins or short peptides (26, 63, 71).

If a cell finds itself in the presence of extracellular di/tripeptides, they are imported inefficiently at first, because of low initial levels of the Ptr2 transporter. However, imported di/tripeptides

that bear destabilizing Nt-residues can bind to the type 1/2 sites of Ubr1. These interactions activate, allosterically, a separate (third) binding site of Ubr1, the one that recognizes an internal degron of the Cup9 repressor (*SI Appendix*, Fig. S9) (26, 71). The resulting "activated" form of Ubr1 targets Cup9 for degradation, reducing its half-life to ∼1 minute and its levels to negligible. As a result, *PTR2* is derepressed and the Prt2 transporter is overproduced, greatly increasing the capacity of cells to import di/tripeptides (*SI Appendix*, Fig. S9) (26, 71). This positive-feedback circuit enables both the budding yeast *S. cerevisiae* and the fission yeast *Schizosaccharomyces pombe* to detect the presence of extracellular di/tripeptides and to react by accelerating their uptake (71, 72).

Deamidation of Nt-Asn and -Gln. In *S. cerevisiae*, Nt-deamidation is mediated by the 52-kDa Nta1 Nt-amidase (Fig. 1*G*) (40, 73). Remarkably, the bulk of Nta1 is located in the inner mitochondrial matrix (https://yeastgfp.yeastgenome.org/). Nevertheless, a low cytosolic (and presumably nuclear) level of Nta1 suffices to mediate the Arg/N-degron pathway (40, 73). Physiological substrates of yeast Nta1 remain to be discovered. Mitochondrial Nta1 might be a component of a distinct (still to be identified) N-degron pathway in the mitochondrial matrix (26).

In animals and plants, Nt-deamidation is mediated by the *Ntan1*-encoded, Nt-Asn–specific Nt^N-amidase and the *Ntaq1*-encoded, Nt-Gln–specific Nt^Q-amidase (*SI Appendix*, Fig. S2*A*) (74). Ntan1 and Ntaq1 are present in the cytosol/nucleus, in contrast to the largely mitochondrial *S. cerevisiae* Nta1 (ref. 74 and references therein). In the fly *Drosophila melanogaster*, the cleavage, by a caspase, of the antiapoptotic Ub ligase Diap1 generates the short-lived Asn^{21}-Diap1 Ct-fragment that is much less efficacious than full-length Diap1 in repressing apoptosis. Degradation of Asn^{21}-Diap1 requires Ntan1 (*SI Appendix*, Fig. S2*A*) (75). A virus would benefit from a delay of apoptosis, as this would facilitate the completion of viral replication in an infected cell. Remarkably, a picorno-like RNA virus induces, through an unknown mechanism, the proteasome-dependent degradation of the Ntan1 Nt^N-amidase in infected insect cells, resulting in a partial stabilization of Asn^{21}-Diap1 (76). This way, a viral infection can down-regulate apoptosis, benefiting the virus (76).

Nt-Arginylation. The 60-kDa Ate1 R-transferase catalyzes the conjugation of Arg (provided by Arg-tRNA) to the α-amino group of a specific Nt-residue of a protein. The resulting Nt-Arg can be bound by Arg/N-recognins (Fig. 1*G* and *SI Appendix*, Figs. S2 and S4). In mammals, there are six isoforms of R-transferase, produced through alternative splicing of the *Ate1* pre-mRNA (*SI Appendix*, Fig. S2 *B* and *C*) (ref. 26 and references therein). A number of natural Ct-fragments, including nearly full-length proteins, are either confirmed or putative substrates of the Ate1 R-transferase and the rest of the Arg/N-degron pathway (*SI Appendix*, Figs. S6–S8).

Arginylation and the Sensing of Oxygen and NO. In 2005, it was discovered that the mammalian Arg/N-degron pathway is a new kind of oxygen (O_2) and NO sensor. The NO/O_2-dependent oxidation of Nt-Cys converts it to Nt-Cys-sulfinate or Nt-Cys-sulfonate, which can be Nt-arginylated, in contrast to unmodified Nt-Cys (*SI Appendix*, Fig. S2*A*) (60, 61). The NO/O_2-dependent proteolysis by the Arg/N-degron pathway controls the levels of a subset of proteins that bear Nt-Cys, including Rgs4, Rgs5, and Rgs16 (60, 61). These conditionally short-lived proteins are regulators of specific G proteins.

The Arg/N-degron pathway is also the main sensor of NO/O_2 in plants, through the NO/O_2-dependent oxidation of Nt-Cys in conditionally short-lived transcription factors that include Rap2.12, Rap2.2, Rap2.3, Hre1, and Hre2 (33, 77, 78). In plants, and possibly in other multicellular eukaryotes as well, the NO/O_2-dependent oxidation of Nt-Cys is catalyzed by Cys-oxidases, in addition to a nonenzymatic oxidation of Nt-Cys. In vivo levels of the above transcription factors and the expression of regulons controlled by them underlie adaptations to a broad range of stresses experienced by plants (refs. 32, 33, 77, and 78 and references therein).

The Arg/N-Degron Pathway as a Sensor of Heme. Both mammalian and yeast Ate1 R-transferases are inhibited by low micromolar levels of hemin (Fe^{3+}-heme) (79). Hemin also accelerates, in vivo, the degradation of mouse Ate1, thereby acting as both a "stoichiometric" and "catalytic" down-regulator of Nt-arginylation. Thus, in addition to being a sensor of NO, O_2, and short peptides, the Arg/N-degron pathway is also a sensor of heme (*SI Appendix*, Fig. S2*A*) (79).

Arginylation, Autophagy, and the Arg/N-Degron Pathway. Kwon and colleagues (13, 49, 50) discovered that p62/Sqstm1 (called p62 below), a component of the autophagy-lysosome system, is also a non-E3 Arg/N-recognin that binds to cytosolic proteins that bear either Nt-Arg or specific hydrophobic Nt-residues. p62 mediates the capture of these proteins by autophagy and their subsequent destruction in lysosomes (refs. 12, 13, 49, and 50 and references therein). Either a proteasome inhibitor or natural stresses can up-regulate the p62/autophagy branch of the Arg/N-degron pathway, termed the Arg/N-$degron^{p62}$ pathway (*SI Appendix*, Fig. S4) (12, 13, 49, 50).

BiP (one of Hsp70 chaperones), calreticulin (another ER chaperone), and protein disulfide isomerase are among ER-resident proteins that bear Nt-arginylatable Nt-residues, such as Nt-Asp or Nt-Glu. Upon stresses, including heat shock and unfolded protein response, a fraction of these ER proteins is transferred to the cytosol, followed by their Nt-arginylation. The resulting Nt-Arg–bearing proteins are captured either by the p62 Arg/N-recognin or by E3 Arg/N-recognins, and are destroyed by the autophagy-lysosome system (the Arg/N-$degron^{p62}$ pathway) or by the 26S proteasome (*SI Appendix*, Fig. S4) (49, 50). In sum, the Arg/N-degron pathway is a major functional link between UPS and autophagy (refs. 12, 13, 49, and 50 and references therein).

The Ac/N-Degron Pathway

About 60% and more than 80% of, respectively, *S. cerevisiae* and human proteins are irreversibly N^{α}-terminally acetylated (Nt-acetylated) by Nt-acetylases (80). The 2010 discovery of Ac/N-degrons (46) identified a major function of Nt-acetylation, a universally present modification whose significance was, until then, largely obscure. The Ac/N-degron pathway targets proteins for degradation by recognizing their Nt-acetylated Nt-residues (Fig. 1*F*) (10, 46–48). The E3 Ub ligases (Ac/N-recognins) of this pathway are the ER membrane-embedded yeast Doa10 and its mammalian counterpart Teb4, and also Not4, the E3 subunit of Ccr4-Not, a multifunctional cytosolic/nuclear complex (10, 46, 48).

Schulman and coworkers (81) showed that the Nt-Ac group of a subunit in a protein complex usually increases thermodynamic stability of the complex. The affinity-enhancing effect of Nt-acetylation provides an explanation for at least intermittently long half-lives of many Nt-acetylated proteins. Specifically, natural Ac/N-degrons tend to be conditional, because of their rapid sequestration within cognate protein complexes (10). The functions of the Ac/N-degron pathway (Fig. 1*F*) include quality control and the regulation of input protein stoichiometries in vivo. For example, *S. cerevisiae* Nt-Ac-Cog1, a short-lived Ac/N-degron substrate, can be made long-lived by coexpressing Cog2 or Cog3, the Cog1-binding subunits of the Golgi-associated COG complex (10). Analogously, *S. pombe* Nt-Ac-Hcn1, a short-lived Ac/N-degron substrate, can be stabilized by coexpressing Cut9, a cognate ligand of Hcn1 in the APC/C Ub ligase (10).

The Pro/N-Degron Pathway

When glucose is low or absent, cells synthesize it through gluconeogenesis. In yeast, the main gluconeogenesis-specific cytosolic enzymes are the Fbp1 fructose-1,6-bisphosphatase, the Icl1 isocitrate lyase, the Mdh2 malate dehydrogenase, and the Pck1 phosphoenolpyruvate carboxykinase. When *S. cerevisiae* grows on a nonfermentable carbon source such as, for example, ethanol, the gluconeogenic enzymes are expressed and long-lived. Transition to a medium containing glucose inhibits the synthesis of these enzymes and induces their degradation, mediated by the multisubunit GID Ub ligase and the proteasome (ref. 35 and references therein).

We discovered that Gid4, a subunit of GID, is the N-recognin of a proteolytic system termed the Pro/N-degron pathway (Figs. 1*E* and 2*A*) (35). Gid4 recognizes a substrate through its Nt-Pro residue or a Pro at position 2, in the presence of distinct (but nonunique) adjoining sequence motifs. The gluconeogenic enzymes Fbp1, Icl1, Mdh2, and Pck1 bear either Nt-Pro or a Pro at position 2, and are conditionally short-lived substrates of the Gid4-dependent Pro/N-degron pathway (Fig. 1*E*) (35–37). The structure of Gid4 comprises an antiparallel β-barrel that contains a deep and narrow substrate-binding cleft (Fig. 2*A*) (36, 37).

The Eukaryotic fMet/N-Degron Pathway

Nascent proteins bear Nt-Met, encoded by the AUG initiation codon. In bacteria and in eukaryotic organelles, mitochondria, and chloroplasts, formyltransferases Nt-formylate the Met moiety of initiator Met-tRNAs. Consequently, nascent bacterial proteins start with Nt-fMet. In contrast, proteins synthesized by the cytosolic ribosomes of eukaryotes bear unformylated Nt-Met, which is often cotranslationally Nt-acetylated, resulting in Ac/N-degrons (Fig. 1*F*) (10, 46–48).

In 2015, it was found that Nt-fMet residues of nascent bacterial proteins can act as bacterial N-degrons, termed fMet/N-degrons (Fig. 1*C*) (38). Remarkably, it was recently discovered that Nt-formylation of proteins, previously thought to be confined to bacteria and bacteria-derived eukaryotic organelles, can also occur at the start of translation by the cytosolic ribosomes of a eukaryote, such as *S. cerevisiae* (Fig. 1*B*) (39). Nt-formylation of yeast cytosolic proteins is mediated by the nuclear DNA-encoded Fmt1 formyltransferase, whose translocation from the cytosol to the inner matrix of mitochondria was found to be not as efficacious, even under normal conditions, as had previously been assumed, and is strongly impaired under conditions of stationary phase and other stresses (39). The cytosolic retention of Fmt1, and the resulting upsurge in the levels of Nt-formylated cytosolic proteins in nutritionally stressed cells, require Gcn2, a protein kinase (39). It was also discovered that Nt-formylated cytosolic proteins are targeted for selective degradation by the Psh1 E3 Ub ligase, which acts as the fMet/N-recognin of the previously unknown eukaryotic fMet/N-degron pathway (Fig. 1*B*) (39).

The Bacterial Leu/N-Degron Pathway

The bacterial Leu/N-degron pathway, which does not involve ubiquitylation, was discovered in 1991 (28) and characterized in Gram-negative bacteria (Fig. 1*D*) (refs. 26, 30, and 51–54 and references therein). This pathway comprises the following components: (*i*) ClpAP, a proteasome-like, ATP-dependent protease; (*ii*) ClpS, the 12-kDa Leu/N-recognin that binds to Nt-Leu, -Phe, -Trp, or -Tyr and delivers bound substrates to the ClpAP protease; (*iii*) Aat, an L/F-transferase that employs Leu-tRNA or Phe-tRNA as a cosubstrate to conjugate largely Leu (and occasionally Phe) to the N-termini of proteins bearing Nt-Lys or Nt-Arg (Fig. 1*D*); and (*iv*) Bpt, an L-transferase that employs Leu-tRNA to conjugate Leu to Nt-Asp, -Glu, and (possibly) oxidized -Cys (Fig. 1*D*). *Vibrio vulnificus*, a human pathogen, contains both Aat and Bpt, while *Escherichia coli* contains only Aat (51). Physiological substrates of the *E. coli* Leu/N-degron pathway include Dps, an 18-kDa DNA-binding protein that compacts the *E. coli* nucleoid in starving cells, and the YgjG putrescine-aminotransferase (PATase) (refs. 30 and 52 and references therein). Although *E. coli* ClpS is nearly 20-fold smaller than yeast or human Ubr1, there is a significant sequelogy between the substrate-binding region of ClpS and a functionally analogous region of Ubr1, suggesting a common descent of bacterial and eukaryotic N-recognins (26, 30).

Studies by Groisman and coworkers indicated that ClpS can target not only N-degrons (Fig. 1*D*), but also N-terminus–proximal internal degrons in bacterial proteins, such as PhoP (ref. 54 and references therein). In a pathway that regulates PhoP, the MgtC protein competes with ClpS for the binding to PhoP, and thereby protects PhoP from degradation. In addition, PhoP, a conditionally short-lived substrate of ClpS, is a transcriptional repressor of ClpS expression. The resulting circuits differentially regulate the rates of degradation of specific ClpS substrates under conditions of low intracellular Mg^{2+} (54). Because bacterial ClpS is a sequelog (65) of much bigger eukaryotic Arg/N-recognins E3s, such as Ubr1 (26, 31), the largely unexplored regulation of the Arg/N-degron pathway in yeast and multicellular eukaryotes may prove to be at least as functionally rich as the already revealed regulation of ClpS and the bacterial Leu/N-degron pathway (ref. 54 and references therein).

Eukaryotic C-Degron Pathways

Because of its free carboxyl group, the Ct-residue of a polypeptide is stereochemically unique, analogously to the Nt-residue and its α-amino group. In 2018, the laboratories of Elledge and Yen discovered a remarkably large set of Ct-degradation signals in human proteins (*SI Appendix*, Fig. S3) (43–45). They also showed that specific E3 Ub ligases of the cullin-RING (CRL) family, and other E3s as well, can recognize these degrons (43–45).

The authors' terms for Ct-degradation signals and pathways that recognize them were, respectively, "C-end degrons" and "DesCEND" (destruction via C-end degron) (43, 44). For reasons discussed at the beginning of this paper, we propose to denote "C-end degrons" as "C-degrons," and "DesCEND pathways" as "C-degron pathways" (*SI Appendix*, Fig. S3). In addition to their succinctness as well as semantic uniformity vis-à-vis N-degrons, it is easy to adapt these terms to specific settings. For example, a pathway mediated by the C-degron–recognizing Kldhc3 subunit of the Crl2 Ub ligase (43, 44) can be called the C-degron$^{Kldhc3}_{Crl2}$ pathway.

Functional Aspects of C-degrons. C-degrons can be present in full-length proteins, in truncated proteins that result from premature termination of translation, and in protein fragments that form upon proteolytic cuts (*SI Appendix*, Fig. S3) (43, 44). All such proteins would be afforded, in vivo, a transient stochastic opportunity to fold or associate in ways that would shield their C-degrons. A C-degron–containing polypeptide that fails to shield its C-degron rapidly enough would face the rising probability of destruction by a cognate C-degron pathway. This temporal pattern is universal among C-degrons and other degradation signals, in that it is relevant to any protein whose degron-based susceptibility to a proteolytic attack changes as a function of time, with the clock beginning to tick at the time of protein's emergence from the ribosomal tunnel.

Cocreation of C-Degrons and N-Degrons upon a Proteolytic Cut. Usp1 is a mammalian DUB (82). Usp1 forms a heterodimer with Uaf1, a non-DUB protein (*SI Appendix*, Fig. S10). Usp1 can autocleave immediately after its internal Gly-Gly sequence (82). The resulting Ct-fragment, Gln-Usp1^{Ct}, bears a deamidation/arginylation-dependent Arg/N-degron (*SI Appendix*, Fig. S10) (56). Nevertheless, the DUB activity of autocleaved Usp1 can be transiently maintained, inasmuch as Usp1^{Nt}, the Nt-fragment of autocleaved Usp1, can remain bound

to the Gln-Usp1^{Ct} Ct-fragment within the cleaved Usp1-Uaf1 heterodimer. The Ct-sequence Gly-Gly of the Usp1^{Nt} Nt-fragment can act as a C-degron, which is recognized by the Kldch2 adaptor subunit of the Clr2 Ub ligase (44). In the resulting mechanism (not yet analyzed in detail), Uaf1 would hold together two Usp1 fragments, allowing them to function, temporarily, as a DUB enzyme, until successful attacks on both fragments by the N-degron and C-degron pathways (refs. 56 and 82 and references therein). Usp1 is the first experimentally addressed setting in which an N-degron and a C-degron can be cocreated upon a cleavage (self-cleavage, in this case) of a full-length protein.

Concluding Remarks

In 1984–1990, studies by our laboratory described the discovery of the first degradation signals (N-degrons) in short-lived proteins; the singular biological significance of UPS; the first physiological functions of ubiquitylation, in the cell cycle, DNA repair, protein synthesis, transcriptional regulation, and stress responses; the Arg/N-degron pathway as the first specific UPS pathway; the subunit selectivity of Ub-dependent proteolysis; the first specific poly-Ub chains and their necessity for protein degradation; the Matα2 repressor as the first physiological substrate of UPS; the first nonproteolytic function of Ub (its role as a cotranslational chaperone); and initiated the molecular genetic understanding of UPS, including the cloning of the first E3 Ub ligase (Ubr1), the first DUBs (Ubp1–Ubp3), and the first precursors of free Ub (Ubi1-Ubi4) (refs. 3 and 4 and references therein). Just how broad and elaborate Ub functions are was understood systematically over the next three decades through studies by many laboratories that entered the field in the 1990s and afterward, an expansion that continues to this day.

Studies of N-degron pathways remained a fount of new genetic and biochemical methods for more than three decades, giving rise to the Ub fusion technique, the Ub reference technique, the Ub translocation technique, the split-Ub technique, the Ub sandwich technique, the heat-inducible N-degron (refs. 3, 4, and 25 and references therein), and other methods by other laboratories.

UPS is of major relevance to medicine. Pharmaceutical companies and academic laboratories are developing compounds that target specific UPS components. The fruits of their labors have already become—or will soon become—clinically useful drugs. Work in this arena is producing not only "conventional" inhibitors or activators of specific enzymes, but also drugs that can direct a Ub ligase to target, destroy, and thereby down-regulate any specific protein. Given the broad functional range of N-degron and C-degron pathways, they will be a part of these advances.

Acknowledgments

Studies in the author's laboratory are supported by the NIH Grants DK039520 and GM031530.

1 Johnson ES, Gonda DK, Varshavsky A (1990) *Cis-trans* recognition and subunit-specific degradation of short-lived proteins. *Nature* 346:287–291.
2 Hershko A, Ciechanover A, Varshavsky A (2000) The ubiquitin system. *Nat Med* 6:1073–1081.
3 Varshavsky A (2008) Discovery of cellular regulation by protein degradation. *J Biol Chem* 283:34469–34489.
4 Varshavsky A (2014) Discovery of the biology of the ubiquitin system. *JAMA* 311:1969–1970.
5 Finley D, Ulrich HD, Sommer T, Kaiser P (2012) The ubiquitin-proteasome system of *Saccharomyces cerevisiae*. *Genetics* 192:319–360.
6 Rape M (2018) Ubiquitylation at the crossroads of development and disease. *Nat Rev Mol Cell Biol* 19:59–70.
7 Joazeiro CAP (2017) Ribosomal stalling during translation: Providing substrates for ribosome-associated protein quality control. *Annu Rev Cell Dev Biol* 33:343–368.
8 Vittal V, Stewart MD, Brzovic PS, Klevit RE (2015) Regulating the regulators: Recent revelations in the control of E3 ubiquitin ligases. *J Biol Chem* 290:21244–21251.
9 Zheng N, Shabek N (2017) Ubiquitin ligases: Structure, function, and regulation. *Annu Rev Biochem* 86:129–157.
10 Shemorry A, Hwang CS, Varshavsky A (2013) Control of protein quality and stoichiometries by N-terminal acetylation and the N-end rule pathway. *Mol Cell* 50:540–551.
11 Shiber A, et al. (2018) Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. *Nature* 561:268–272.
12 Grumati P, Dikic I (2018) Ubiquitin signaling and autophagy. *J Biol Chem* 293:5404–5413.
13 Ji CH, Kwon YT (2017) Crosstalk and interplay between the ubiquitin-proteasome system and autophagy. *Mol Cells* 40:441–449.
14 Balchin D, Hayer-Hartl M, Hartl FU (2016) In vivo aspects of protein folding and quality control. *Science* 353:aac4354.
15 Varshavsky A (1991) Naming a targeting signal. *Cell* 64:13–15.
16 Bard JAM, et al. (2018) Structure and function of the 26S proteasome. *Annu Rev Biochem* 87:697–724.
17 Schweitzer A, et al. (2016) Structure of the human 26S proteasome at a resolution of 3.9 Å. *Proc Natl Acad Sci USA* 113:7816–7821.
18 Finley D, Chen X, Walters KJ (2016) Gates, channels, and switches: Elements of the proteasome machine. *Trends Biochem Sci* 41:77–93.
19 Collins GA, Goldberg AL (2017) The logic of the 26S proteasome. *Cell* 169:792–806.
20 Budenholzer L, Cheng CL, Li Y, Hochstrasser M (2017) Proteasome structure and assembly. *J Mol Biol* 429:3500–3524.
21 Gödderz D, et al. (2015) Cdc48-independent proteasomal degradation coincides with a reduced need for ubiquitylation. *Sci Rep* 5:7615.
22 Inobe T, Fishbain S, Prakash S, Matouschek A (2011) Defining the geometry of the two-component proteasome degron. *Nat Chem Biol* 7:161–167.
23 Bachmair A, Finley D, Varshavsky A (1986) In vivo half-life of a protein is a function of its amino-terminal residue. *Science* 234:179–186.
24 Bachmair A, Varshavsky A (1989) The degradation signal in a short-lived protein. *Cell* 56:1019–1032.
25 Varshavsky A (2005) Ubiquitin fusion technique and related methods. *Methods Enzymol* 399:777–799.
26 Varshavsky A (2011) The N-end rule pathway and regulation by proteolysis. *Protein Sci* 20:1298–1345.
27 Bartel B, Wünning I, Varshavsky A (1990) The recognition component of the N-end rule pathway. *EMBO J* 9:3179–3189.
28 Tobias JW, Shrader TE, Rocap G, Varshavsky A (1991) The N-end rule in bacteria. *Science* 254:1374–1377.
29 Rao H, Uhlmann F, Nasmyth K, Varshavsky A (2001) Degradation of a cohesin subunit by the N-end rule pathway is essential for chromosome stability. *Nature* 410:955–959.
30 Dougan DA, Micevski D, Truscott KN (2012) The N-end rule pathway: From recognition by N-recognins, to destruction by AAA+proteases. *Biochim Biophys Acta* 1823:83–91.
31 Tasaki T, Sriram SM, Park KS, Kwon YT (2012) The N-end rule pathway. *Annu Rev Biochem* 81:261–289.
32 Gibbs DJ, Bacardit J, Bachmair A, Holdsworth MJ (2014) The eukaryotic N-end rule pathway: Conserved mechanisms and diverse functions. *Trends Cell Biol* 24:603–611.
33 Dissmeyer N, Rivas S, Graciet E (2018) Life and death of proteins after protease cleavage: Protein degradation by the N-end rule pathway. *New Phytol* 218:929–935.
34 Eldeeb MA, Leitao LCA, Fahlman RP (2018) Emerging branches of the N-end rule pathways are revealing the sequence complexities of N-termini dependent protein degradation. *Biochem Cell Biol* 96:289–294.
35 Chen SJ, Wu X, Wadas B, Oh J-H, Varshavsky A (2017) An N-end rule pathway that recognizes proline and destroys gluconeogenic enzymes. *Science* 355:366.

36 Dougan DA, Varshavsky A (2018) Understanding the Pro/N-end rule pathway. *Nat Chem Biol* 14:415–416.
37 Dong C, et al. (2018) Molecular basis of GID4-mediated recognition of degrons for the Pro/N-end rule pathway. *Nat Chem Biol* 14:466–473.
38 Piatkov KI, Vu TT, Hwang CS, Varshavsky A (2015) Formyl-methionine as a degradation signal at the N-termini of bacterial proteins. *Microb Cell* 2:376–393.
39 Kim JM, et al. (2018) Formyl-methionine as an N-degron of a eukaryotic N-end rule pathway. *Science* 362:eaat0174.
40 Kim MK, Oh SJ, Lee BG, Song HK (2016) Structural basis for dual specificity of yeast N-terminal amidase in the N-end rule pathway. *Proc Natl Acad Sci USA* 113:12438–12443.
41 Szoradi T, 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.
42 Keiler KC, Waller PR, Sauer RT (1996) Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* 271:990–993.
43 Koren I, et al. (2018) The eukaryotic proteome is shaped by E3 ubiquitin ligases targeting C-terminal degrons. *Cell* 173:1622–1635.e14.
44 Lin HC, et al. (2018) C-terminal end-directed protein elimination by CRL2 ubiquitin ligases. *Mol Cell* 70:602–613.e3.
45 Chatr-Aryamontri A, van der Sloot A, Tyers M (2018) At long last, a C-terminal bookend for the ubiquitin code. *Mol Cell* 70:568–571.
46 Hwang CS, Shemorry A, Varshavsky A (2010) N-terminal acetylation of cellular proteins creates specific degradation signals. *Science* 327:973–977.
47 Kim HK, et al. (2014) The N-terminal methionine of cellular proteins as a degradation signal. *Cell* 156:158–169.
48 Park SE, et al. (2015) Control of mammalian G protein signaling by N-terminal acetylation and the N-end rule pathway. *Science* 347:1249–1252.
49 Shim SM, et al. (2018) The endoplasmic reticulum-residing chaperone BiP is short-lived and metabolized through N-terminal arginylation. *Sci Signal* 11:eaan0630.
50 Yoo YD, et al. (2018) N-terminal arginylation generates a bimodal degron that modulates autophagic proteolysis. *Proc Natl Acad Sci USA* 115:E2716–E2724.
51 Graciet E, et al. (2006) Aminoacyl-transferases and the N-end rule pathway in a human pathogen. *Proc Natl Acad Sci USA* 103:3078–3083.
52 Schmidt R, Zahn R, Bukau B, Mogk A (2009) ClpS is the recognition component for *Escherichia coli* substrates of the N-end rule degradation pathway. *Mol Microbiol* 72:506–517.
53 Rivera-Rivera I, Román-Hernández G, Sauer RT, Baker TA (2014) Remodeling of a delivery complex allows ClpS-mediated degradation of N-degron substrates. *Proc Natl Acad Sci USA* 111:E3853–E3859.
54 Yeom J, Gao X, Groisman EA (2018) Reduction in adaptor amounts establishes degradation hierarchy among protease substrates. *Proc Natl Acad Sci USA* 115:E4483–E4492.
55 Piatkov KI, Brower CS, Varshavsky A (2012) The N-end rule pathway counteracts cell death by destroying proapoptotic protein fragments. *Proc Natl Acad Sci USA* 109:E1839–E1847.
56 Piatkov KI, Colnaghi L, Békés M, Varshavsky A, Huang TT (2012) The auto-generated fragment of the Usp1 deubiquitylase is a physiological substrate of the N-end rule pathway. *Mol Cell* 48:926–933.
57 Piatkov KI, Oh J-H, Liu Y, Varshavsky A (2014) Calpain-generated natural protein fragments as short-lived substrates of the N-end rule pathway. *Proc Natl Acad Sci USA* 111:E817–E826.
58 Justa-Schuch D, et al. (2016) DPP9 is a novel component of the N-end rule pathway targeting the tyrosine kinase Syk. *eLife* 5:e16370.
59 Brower CS, Piatkov KI, Varshavsky A (2013) Neurodegeneration-associated protein fragments as short-lived substrates of the N-end rule pathway. *Mol Cell* 50:161–171.
60 Hu R-G, et al. (2005) The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators. *Nature* 437:981–986.
61 Lee MJ, et al. (2005) RGS4 and RGS5 are in vivo substrates of the N-end rule pathway. *Proc Natl Acad Sci USA* 102:15030–15035.
62 Kwon YT, et al. (2002) An essential role of N-terminal arginylation in cardiovascular development. *Science* 297:96–99.
63 Hwang CS, Shemorry A, Auerbach D, Varshavsky A (2010) The N-end rule pathway is mediated by a complex of the RING-type Ubr1 and HECT-type Ufd4 ubiquitin ligases. *Nat Cell Biol* 12:1177–1185.
64 Oh JH, Hyun JY, Varshavsky A (2017) Control of Hsp90 chaperone and its clients by N-terminal acetylation and the N-end rule pathway. *Proc Natl Acad Sci USA* 114:E4370–E4379.
65 Varshavsky A (2004) 'Spalog' and 'sequelog': Neutral terms for spatial and sequence similarity. *Curr Biol* 14:R181–R183.
66 Kim ST, et al. (2018) The N-recognin UBR4 of the N-end rule pathway is required for neurogenesis and homeostasis of cell surface proteins. *PLoS One* 13:e0202260.
67 Shearer RF, Iconomou M, Watts CK, Saunders DN (2015) Functional roles of the E3 ubiquitin ligase UBR5 in cancer. *Mol Cancer Res* 13:1523–1532.
68 Zenker M, et al. (2005) Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome). *Nat Genet* 37:1345–1350.
69 Sekine S, Youle RJ (2018) PINK1 import regulation; a fine system to convey mitochondrial stress to the cytosol. *BMC Biol* 16:2.
70 Weaver BP, Weaver YM, Mitani S, Han M (2017) Coupled caspase and N-end rule ligase activities allow recognition and degradation of pluripotency factor LIN-28 during non-apoptotic development. *Dev Cell* 41:665–673.e6.
71 Turner GC, Du F, Varshavsky A (2000) Peptides accelerate their uptake by activating a ubiquitin-dependent proteolytic pathway. *Nature* 405:579–583.
72 Kitamura K, Fujiwara H (2013) The type-2 N-end rule peptide recognition activity of Ubr11 ubiquitin ligase is required for the expression of peptide transporters. *FEBS Lett* 587:214–219.
73 Baker RT, Varshavsky A (1995) Yeast N-terminal amidase. A new enzyme and component of the N-end rule pathway. *J Biol Chem* 270:12065–12074.
74 Wang H, Piatkov KI, Brower CS, Varshavsky A (2009) Glutamine-specific N-terminal amidase, a component of the N-end rule pathway. *Mol Cell* 34:686–695.
75 Ditzel M, et al. (2003) Degradation of DIAP1 by the N-end rule pathway is essential for regulating apoptosis. *Nat Cell Biol* 5:467–473.
76 Wang Z, et al. (2017) A picorna-like virus suppresses the N-end rule pathway to inhibit apoptosis. *eLife* 6:e30590.
77 Vicente J, et al. (2017) The Cys-Arg/N-end rule pathway is a general sensor of abiotic stress in flowering plants. *Curr Biol* 27:3183–3190.e4.
78 White MD, et al. (2017) Plant cysteine oxidases are dioxygenases that directly enable arginyl transferase-catalysed arginylation of N-end rule targets. *Nat Commun* 8:14690.
79 Hu R-G, Wang H, Xia Z, Varshavsky A (2008) The N-end rule pathway is a sensor of heme. *Proc Natl Acad Sci USA* 105:76–81.
80 Aksnes H, Drazic A, Marie M, Arnesen T (2016) First things first: Vital protein marks by N-terminal acetyltransferases. *Trends Biochem Sci* 41:746–760.
81 Scott DC, Monda JK, Bennett EJ, Harper JW, Schulman BA (2011) N-terminal acetylation acts as an avidity enhancer within an interconnected multiprotein complex. *Science* 334:674–678.
82 Huang TT, et al. (2006) Regulation of monoubiquitinated PCNA by DUB autocleavage. *Nat Cell Biol* 8:339–347.

PNAS
www.pnas.org

Supplementary Information (SI) for

The N-degron and C-degron pathways of protein degradation

Alexander Varshavsky

Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125

Email: avarsh@caltech.edu

Second SI revision, November 27, 2018

This PDF includes:

Figures S1 to S10.

References for SI reference citations.

www.pnas.org/cgi/doi/10.1073/pnas.1816596115

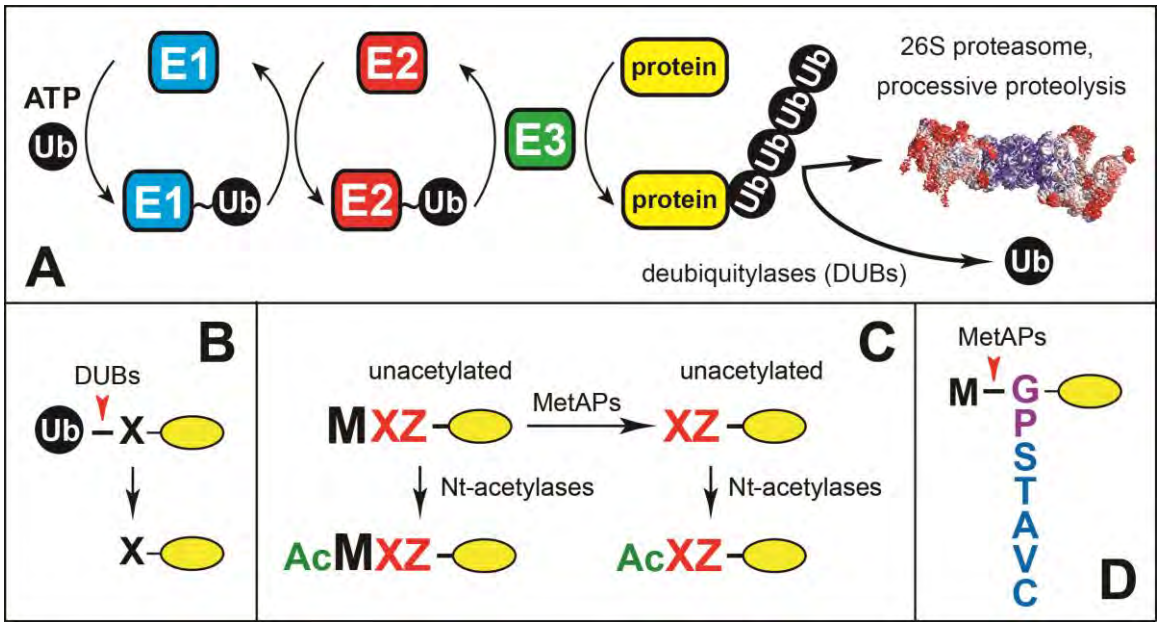

Fig. S1. The ubiquitin (Ub)-proteasome system (UPS), the Ub fusion technique, the N-terminal processing of newly formed proteins, and specificity of Met-aminopeptidases.

(*A*) The UPS. Conjugation of Ub to other proteins involves a preliminary ATP-dependent step in which the last residue of Ub (Gly^{76}) is joined, via a thioester bond, to a Cys residue of the E1 (Ub-activating) enzyme. The "activated" Ub moiety is then conjugated to a Cys residue in one of Ub-conjugating (E2) enzymes, and from there, through an isopeptide bond, to a Lys residue of an ultimate acceptor, denoted as "protein". E2s function as subunits of E2-E3 Ub ligase complexes that can produce substrate-linked poly-Ub chains. These chains have specific Ub-Ub topologies, depending on the identity of a specific Lys residue of Ub that forms an isopeptide bond with C-terminal Gly^{76} of the adjacent Ub moiety in a poly-Ub chain. Specific poly-Ub chains can confer a processive degradation of a substrate by the 26S proteasome, or other metabolic fates. One role of E3 is the recognition of a substrate's degradation signal (degron). An individual mammalian genome encodes more than 800 distinct E3 Ub ligases. See also the main text.

(*B*) The Ub fusion technique. In eukaryotes, linear fusions of Ub to other proteins are cotranslationally cleaved by deubiquitylases (DUBs) at the last residue of Ub, making it possible to produce, in vivo, different residues at the N-termini of otherwise identical proteins (1, 2).

(C) N-terminal processing of nascent eukaryotic proteins by N^{α}-terminal acetylases (Nt-acetylases) and Met-aminopeptidases (MetAPs) (3). "Ac" denotes the N^{α}-terminal acetyl moiety. M, Met. X and Z, single-letter abbreviations for any amino acid residue. Yellow ovals denote the rest of a protein.

(*D*) Met-aminopeptidases (MetAPs) cotranslationally cleave off the N-terminal Met residue of a nascent protein if a residue at position 2 belongs to the set of indicated residues (4, 5). Gly and Pro at position 2 are depicted in a different color because these residues, in contrast to other (indicated) small residues, are very rarely Nt-acetylated after the removal of N-terminal Met (3).

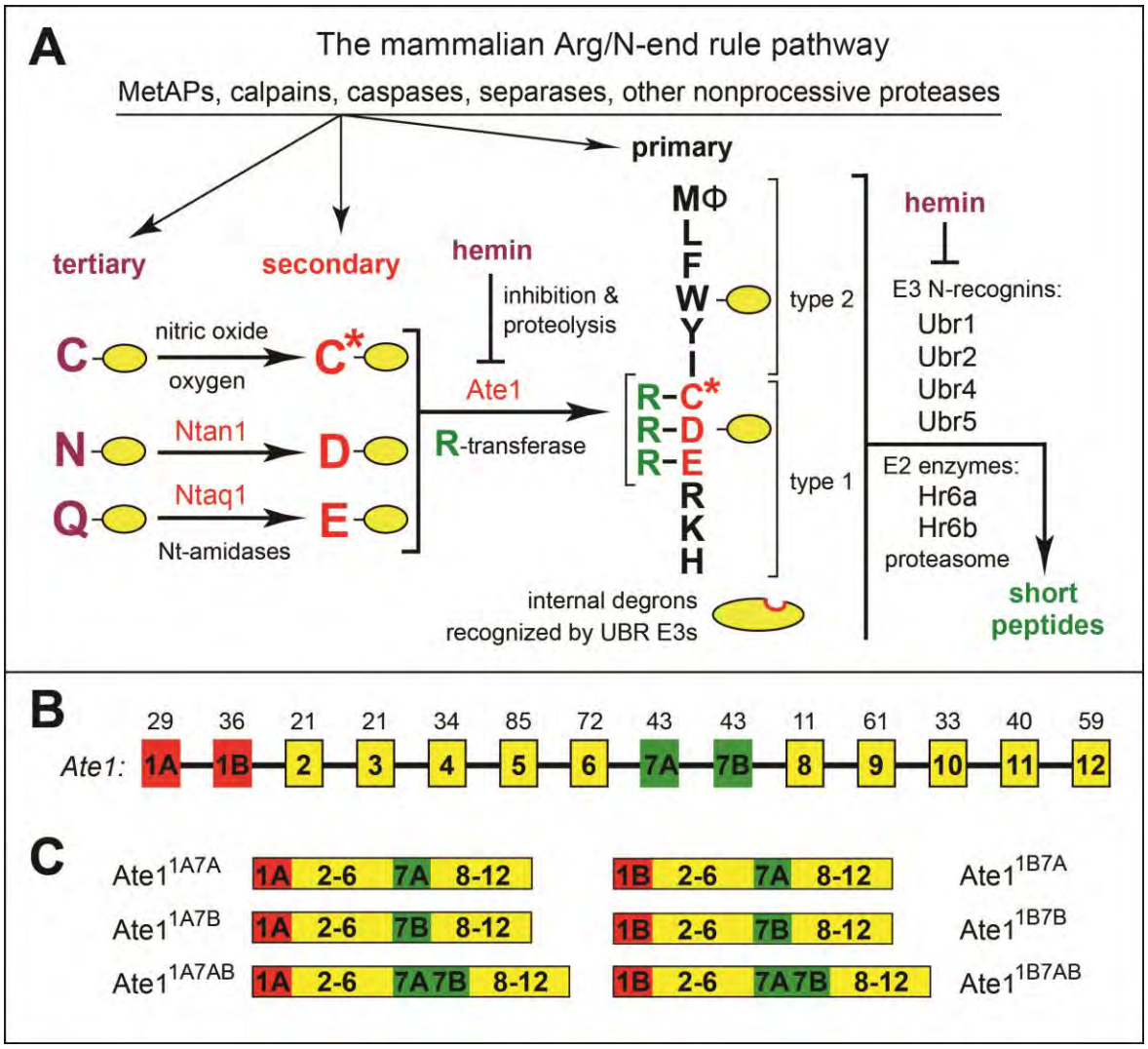

Fig. S2. The mammalian Arg/N-degron pathway and isoforms of the mouse Ate1 R-transferase. Amino acid residues are denoted by single-letter abbreviations. A yellow oval denotes the rest of a protein substrate.

(*A*) The mammalian Arg/N-end rule pathway. It targets proteins for degradation through their specific unacetylated N-terminal (Nt) residues. "Primary", "secondary", and "tertiary" refer to mechanistically distinct classes of destabilizing Nt-residues. Ntan1 and Ntaq1 are N-terminal amidases (Nt-amidases) that convert, respectively, the tertiary destabilizing Nt-residues Asn and Gln to Asp and Glu. The Ate1 R-transferase (arginyltransferase or Arg-tRNA-protein transferase) conjugates Arg, a primary destabilizing residue, to N-terminal Asp, Glu and (oxidized) Cys. "Type 1" and "type 2" refer, respectively, to two sets of primary destabilizing N-terminal residues, basic (Arg, Lys, His) and bulky hydrophobic (Leu, Phe, Trp, Tyr, Ile, and also Met, if the latter is followed by a bulky hydrophobic residue (Φ)). These sets of N-terminal residues are recognized by two distinct substrate-binding sites of N-recognins, the pathway's E3 ubiquitin ligases Ubr1, Ubr2, Ubr4, and Ubr5. Also indicated is the ability of Arg/N-recognin E3s to recognize specific internal (non-N-terminal) degrons in proteins that lack Arg/N-degrons. See the main text for references and further details. Yet another

branch of this proteolytic system, denoted as the Arg/N-degron^{p62} pathway and discovered by Kwon and colleagues, involves p62/Sqstm1, a non-E3 Arg/N-recognin and a component of autophagy-lysosome pathways (see the main text and Fig. S4) ((6-15) and refs. therein).

(*B*) The exons, including alternative exons (1A/1B and 7A/7B) of the mouse *Ate1* gene, which encodes alternative splicing-derived isoforms of R-transferase. The deduced lengths (in amino acid residues) of the encoded *Ate1* exons are indicated on top.

(*D*) Mouse R-transferase isoforms that are produced through alternative splicing of *Ate1* pre-mRNA. The terminology of these isoforms (Ate1^{1A7A}, Ate1^{1A7B}, Ate1^{1B7A}, Ate1^{1B7B}, Ate1^{1A7AB}, Ate1^{1B7AB}) is based on the presence or absence of the alternative exons 1A/1B and 7A/7B. Ate1^{1A7AB} and Ate1^{1B7AB} are minor Ate1 isoforms that result from the retention of both variants of exon 7 (16, 17).

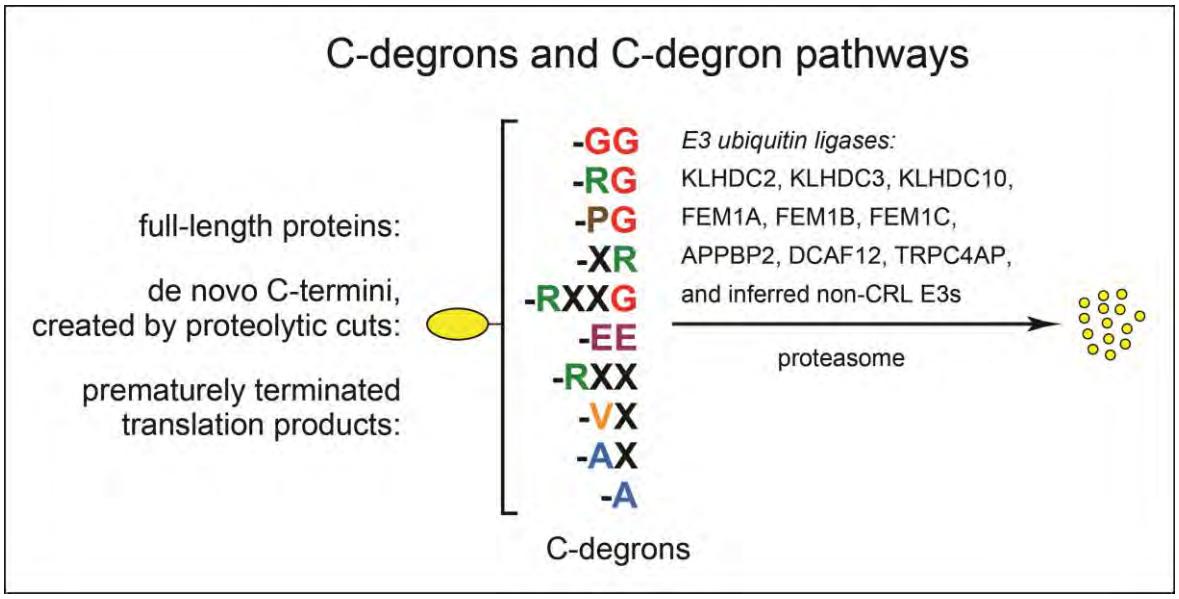

Fig. S3. C-degrons and C-degron pathways in human cells. This diagram is a simplified summary of the 2018 discovery, by the laboratories of Elledge and Yen, of a large set of C-degrons in human proteins (18-20). Amino acid residues are denoted by single-letter abbreviations. A yellow oval denotes a protein substrate upstream of its C-terminus. The indicated C-terminal (Ct) sequences and individual Ct-residues, referred to as C-degrons, are targeted, in conjunction with internal Lys residues of individual C-degron substrates, by a broad range of Ub ligases, largely but not solely of the CRL class (18-20). See the main text for a brief discussion of C-degron pathways and the term "C-degron" for denoting this class of degradation signals.

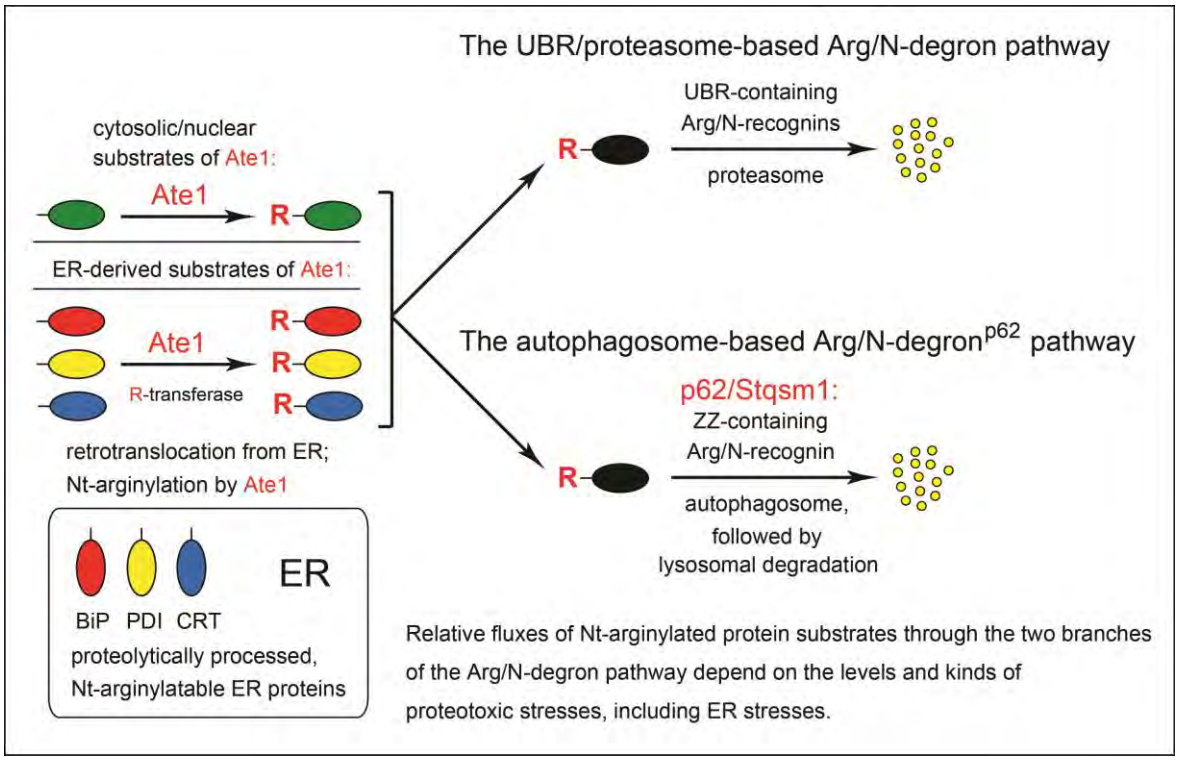

Fig. S4. Autophagy-mediated versus proteasome-mediated degradation of Arg/N-degron substrates. Amino acid residues are denoted by single-letter abbreviations. A yellow oval denotes the rest of a protein substrate.

Kwon and colleagues discovered that p62/Sqstm1 (called p62 below), a component of the autophagy-lysosome system, is also a non-E3 Arg/N-recognin that binds to cytosolic proteins bearing Nt-Arg or specific hydrophobic Nt-residues. p62 mediates the capture of these proteins by autophagy and their subsequent destruction in lysosomes (6-15). Either a proteasome inhibitor or natural stresses can up-regulate the p62/autophagy branch of the Arg/N-degron pathway (6-15). BiP (one of Hsp70 chaperones), calreticulin (CRT, another ER chaperone), and protein disulfide isomerase (PDI) are among ER-resident proteins that bear Nt-arginylatable Nt-residues such as Nt-Asp or Nt-Glu. Upon stresses, including heat stress and unfolded protein response (UPR), a fraction of these ER proteins is transferred to the cytosol, followed by their Nt-arginylation. The resulting Nt-Arg-bearing proteins are captured either by the p62 Arg/N-recognin and/or by E3 Arg/N-recognins, and are destroyed, respectively, by the autophagy-lysosome system and/or by the 26S proteasome (Fig. S2A) (6-15). In sum, the Arg/N-degron pathway is a major functional link between UPS and autophagy.

The affinity of p62 for Nt-Arg (K_d of ~44 nM) is at least 50-fold higher than that of Ubr1/Ubr2 (12). If so, how can intracellular Ubr1, Ubr2, Ubr4 and Ubr5 E3s (Fig. S2A) compete, in vivo, with the relatively abundant p62 for their binding to proteins bearing Nt-Arg? One possibility is that the affinity of p62 for Nt-Arg may be regulated in living cells. Both basic and aromatic Nt-residues of polypeptides that interact with p62 bind to its ~50-residue zinc finger-like ZZ domain, which is sequelogous (21) to the ~80-residue UBR domain of Ubr1/Ubr2 (6).

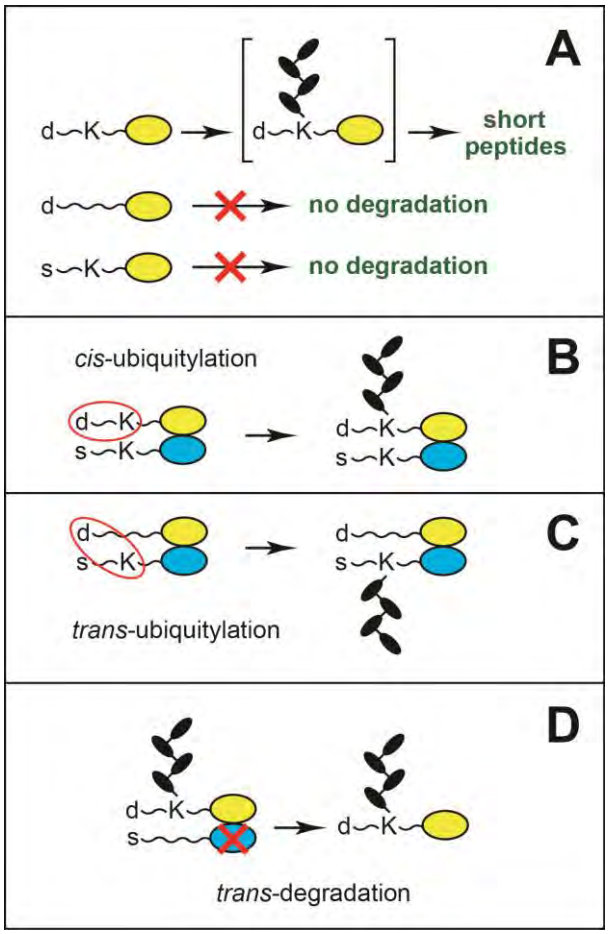

Fig. S5. Organization and cis-trans targeting of eukaryotic N-degrons. Amino acid residues are denoted by single-letter abbreviations. A yellow oval denotes the rest of a protein substrate.

(*A*) Determinants of a eukaryotic N-degron. *d*, a destabilizing Nt-residue. *K*, a "ubiquitylatable" internal Lys residue. See also the main text and refs. (17, 22, 23).

(*B* and *C*) *Cis* versus *trans* polyubiquitylation of an oligomeric Arg/N-degron substrate that can result in the degradation of a subunit that becomes linked to a poly-Ub chain (24).

(D) *Trans*-degradation, in which a specific subunit of an oligomeric protein is polyubiquitylated but is not degraded by the 26S proteasome, for example, because that subunit lacks an unstructured region that is required for the initiation of degradation. Instead, a subunit-selective degradation another, nonubiquitylated subunit takes place. This mode of degradation was discovered by Matouschek and colleagues for oligomeric substrates of the Ub-fusion-degradation (UFD) pathway (25, 26). It remains to be determined whether an analogous *trans*-degradation of an oligomeric Arg/N-degron substrate can also occur.

Caspase-generated proapoptotic Ct-fragments bearing destabilizing Nt-residues

Protein	Cleavage site	Nt-residue	C-terminal fragment
Proapoptotic protein fragments that are experimentally confirmed Arg/N-degron substrates			
Mm **Ripk1**	321 SLQHD↓**C**VPL...S 656	Nd^t	Proapoptotic fragment, generated by caspase-8, of the RIPK1 kinase. Cys-RIPK1 is an N-end rule substrate.
Mm **Traf1**	152 DLEVD↓**C**YRA...A 409	Nd^t	Proapoptotic fragment, generated by caspase-8, of TRAF1, a regulator of apoptosis. Cys-TRAF1 is an N-end rule substrate.
Mm **Brca1**	1118 DDLLD↓**D**VEI...D 1812	Nd^s	Proapoptotic fragment, generated by caspase-3, of the tumor suppressor BRCA1. Asp-BRCA1 is an N-end rule substrate.
Hs **Limk1**	236 LDEID↓**L**LIQ...D 647	Nd^p	Proapoptotic fragment, generated by caspase-3, of LIMK1, a Ser/Thr protein kinase. Leu-LIMK1 is an N-end rule substrate.
Hs **Nedd9**	626 MDDYD↓**Y**VHL...F 834	Nd^p	Proapoptotic fragment, generated by caspase-3, of NEDD9, a regulator of cell adhesion. Tyr-NEDD9 is an N-end rule substrate.
Hs-**Bid**	66 HSRLG↓**R**IEA...D 195	Nd^p	Proapoptotic fragment, generated by calpains, of BID, a regulator of apoptosis. Arg-BID is an N-end rule substrate.
Mm **Bcl$_{XL}$**	56 SWHLA↓**D**SPA...K 233	Nd^s	Proapoptotic fragment, generated by calpains, of the antiapoptotic regulator BCL$_{XL}$. Asp-BCL$_{XL}$ is an N-end rule substrate.
Mm **Bim$_{EL}$**	9 SSECD↓**R**EGG...H 196	Nd^p	Proapoptotic fragment, generated by caspase-3, of the apoptosis regulator BIM$_{EL}$. Arg-BIM$_{EL}$ is an N-end rule substrate.
Mm **Epha4**	769 RVLED↓**D**PEA...V 986	Nd^s	Proapoptotic fragment, generated by caspase-3, of the dependence receptor EPHA4. Cys-EPHA4 is an N-end rule substrate.
Mm **Met**	996 NESVD↓**Y**RAT...S 1379	Nd^p	Proapoptotic fragment, generated by caspase-3, of the dependence receptor MET. Tyr-MET is an N-end rule substrate.
Proapoptotic protein fragments that remain to be verified as Arg/N-degron substrates			
Hs **PkCδ**	325 EDMQD↓**N**SGT...D 676	Nd^t	Proapoptotic fragment, generated by caspase-3, of the antiapoptotic protein kinase Cδ. Asn-PKCδ is a likely N-end rule substrate.
Hs **PkCθ**	350 LDEVD↓**K**MCH...S 706	Nd^p	Proapoptotic fragment, generated by caspase-3, of the antiapoptotic protein kinase Cθ. Lys-PKCθ is a likely N-end rule substrate.
Hs **Etk**	236 EDFPD↓**W**WQV...H 675	Nd^p	Proapoptotic fragment, generated by caspase-3, of the antiapoptotic ETK/BMX tyrosine kinase. Trp-ETK is a likely N-end rule substrate.
Mm **Slk**	432 PDTQD↓**Q**QTV...S 1202	Nd^t	Proapoptotic fragment, generated by caspase-3, of the SLK kinase, a regulator of actin. Gln-SLK is a likely N-end rule substrate.
Hs **Hpk1**	381 YDDVD↓**I**PTP...L 833	Nd^p	Proapoptotic fragment, generated by caspase-3, of the Ser/Thr kinase HPK1. Ile-HPK1 is a likely N-end rule substrate.
Hs-**Mlh1**	414 EDKTD↓**I**ISS...C 756	Nd^p	Proapoptotic fragment, generated by caspase-3, of the mismatch repair protein MLH1. Ile-MLH1 is a likely N-end rule substrate.

Fig. S6. Natural proapoptotic protein C-terminal (Ct) fragments that are experimentally confirmed substrates of the Arg/N-degron pathway, and some proapoptotic Ct-fragments that remain to be verified as such substrates. The list of either confirmed or putative Arg/N-degron substrates in this and other figures (Figs. S7 and S8) is but a subset of the known list of putative Arg/N-degron substrates, most of which remain to be experimentally

verified as such. The name of a protein on the left is preceded by the acronym of a species (Hs, *Homo sapiens*; Mm, *Mus musculus*). Amino acid residues are indicated by single-letter abbreviations. Arrowheads and enlarged residues, in red, indicate P1' residues of cleavage sites, i.e., the Nt-residues of Ct-fragments of the indicated cleaved full-length proteins. Primary, secondary and tertiary destabilizing N-terminal residues are denoted as Nd^p, Nd^s and Nd^t, respectively (17, 27). If a depicted cleavage site is the one recognized by a caspase (the exceptions, in the present list, are the calpain cleavage sites that yield Asp-Bcl_{XL} and Arg-Bid), the site is underlined. The residue number on the left is of the first shown residue of uncleaved protein, numbered as in a full-length protein. The number on the right is of the last residue of a full-length protein.

The first section of this figure describes 10 previously identified natural proapoptotic Ct-fragments, specifically Cys-Ripk1, Cys-Traf1, Asp-Brca1, Leu-Limk1, Tyr-Nedd9, Arg-Bid, Asp-Bcl_{XL}, Arg-Bim_{EL}, Asp-Epha4, and Tyr-Met. They were found to be short-lived substrates of the Arg/N-end rule pathway (28). The figure's second section describes six other previously identified proapoptotic Ct-fragments (all of them are produced by caspases) bearing destabilizing Nt-residues that can be recognized by the Arg/N-degron pathway. These fragments remain to be verified as Arg/N-degron substrates.

Asn-PkCδ is the Ct-fragment of the protein kinase Cδ (PkCδ) that can be generated by (in particular) caspase-3. This fragment bears Nt-Asn and is proapoptotic, in contrast to the full-length PkCδ kinase (29-33). Lys-PkCθ is the Ct-fragment of the protein kinase Cθ (PkCθ). This fragment can be generated by (in particular) caspase-3, bears Nt-Lys, and is proapoptotic, in contrast to the full-length PkCθ kinase (34). Trp-Etk is the Ct-fragment of the Etk/Bmc tyrosine kinase, a member of the Btk/Tek family of kinases, at least some of which regulate apoptosis. The Trp-Etk fragment can be generated by (in particular) caspase-3, bears Nt-Trp, and is proapoptotic, in contrast to the full-length Etk kinase (35). Gln-Slk is the Ct-fragment of Slk, a Ste20-related protein kinase that plays a role in regulation of actin fibers. The Gln-Slk fragment can be generated by (in particular) caspase-3, bears Nt-Gln, and is proapoptotic. The concomitantly produced Nt-fragment of Slk is also proapoptotic (36). Ile-Hhp1 is the Ct-fragment of the hematopoietic progenitor kinase 1 (Hpk1), a Ste20-related protein kinase whose functions include stimulation of the stress-activated protein kinases SAPKs/JNKs and the NF-κB transcriptional regulon. The Ile-Hpk1 fragment can be generated by (in particular) caspase-3, bears Nt-Ile, and is proapoptotic, in contrast to the full-length Hk1 kinase (37). Ile-Mlh1 is the Ct-fragment of the mismatch repair Mhl1 protein that can be generated by (in particular) caspase-3, bears Nt-Ile, resides in the cytosol (unlike the full-length nuclear MLH1) and is proapoptotic, in contrast to full-length MLH1 (38).

Calpain-generated C-terminal (Ct) protein fragments that are either identified or predicted substrates of the Arg/N-degron pathway

	Ct fragment & description of its full-length precursor	Calpain cleavage site
1	***Glu*-Bak**; proapoptotic regulator of apoptosis	11 KVGCD**E**SPS...S 208
2	***Arg*-Bid**; proapoptotic regulator of apoptosis	66 SFNQG**R**IEP...D 195
3	***Asp*-Bcl_{XL}**; antiapoptotic regulator of apoptosis	56 SWHLA**D**SPA...K 233
4	***Arg*-c-Fos**; leucine-zipper transcription factor	86 APSQT**R**APH...L 380
5	***Glu*-IκBα**; inhibitor of the NFκB transcription factor	46 VKELR**E**IRL...L 314
6	***Arg*-Igfbp2**; insulin-like growth factor-binding prot.	176 VNEQH**R**QMG...Q 305
7	***Asp*-Capns1**; calpain, regulatory subunit	138 DGFGI**D**TCR...S 269
8	***Arg*-Atp2b2**; PMCA, pl. membr. Ca^{2+} ATPase	1086 FRGLN**R**IQT...L 1198
9	***Leu*-Capn1**; calpain, catalytic subunit	23 DKELG**L**GRH...A 713
10	***Arg*-Ankrd2**; regulator of muscle differentiation	98 SLDLR**R**EII...Q 358
11	***Tyr*-Grm1**; subunit of metabotropic Glu receptor	932 PLTKS**Y**QGS...L 1199
12	***Lys*-Ica512**; member of receptor phosphatase fam.	604 SRQRD**K**ERL...Q 979
13	***Phe*-GluN2a**; sub. of NMDA Glu receptor, Ca^{2+} channel	1274 NNALQ**F**QKN...V 1464
14	***Asn*-DSCR1 (RCAN1)**; inhibitor of calcineurin	2 HFRNF**N**YSF...S 197
15	***Arg*-GlyT1A**; transmembrane glycine transporter	22 DQNLT**R**GNW...I 633
16	***Asn*-$Ca_V1.1$**; transmembr. voltage-gated Ca^{2+} channel	1660 NTNNA**N**ANV...E 1880
17	***Lys*-cortactin**; actin-binding cytoskeletal protein	354 FENLA**K**ERE...Q 546
18	***Asn*-Bfl-1**; antiapoptotic member of Bcl-2 family	67 ARIIF**N**QVM...K 172
19	***Arg*-dystrophin**; cytoskeletal protein	688 TMVTT**R**EQI...M 3678
20	***Gln*-Ryr1**; Ca^{2+} channel (ryanodine receptor)	1397 VAMMT**Q**PPS...S 5035
21	***Arg*-Mef2d**; enhancer factor (transcription)	74 EPHES**R**TNA...K 514
22	***Gln*-talin**; cytoskeletal protein	429 STVLQ**Q**QYN...H 2541
23	***Arg*-p39**; activator of Cdk5 cyclin-dependent kinase	104 LLRKG**R**DGP...R 369
24	***Gln*-Egfr**; transmembrane receptor for EGF	681 LRRLL**Q**ERE...A 1210
25	***Leu*-β-catenin**; cytoskeletal and transcription factor	26 QQQSY**L**DSG...L 781
26	***Leu*-Nf2** (merlin); tumor suppressor, cytoskeleton	294 KLILQ**L**CIG...L 596
27	***Arg*-caspase-9**; caspase, inactivated by calpain	111 LRPEI**R**KPE...S 416
28	***Leu*-troponin T2**; cytoskeletal protein	77 PKPSR**L**FMP...K 301
29	***Lys*-PKCα**; protein kinase C-alpha	320 PSEDR**K**QPS...V 672
30	***Leu*-RAD21**; subunit of chromosomal cohesin	188 ASNLL**L**EPE...I 635
31	***Phe*-PKCγ**; protein kinase C-gamma	334 SKRCF**F**GAS...M 697
32	***Leu*-$STEP_{33}$**; striatal-enriched tyrosine phosphatase	220 GSNVS**L**TLD...E 541
33	***Leu*-CAMK-IV**; Ca^{2+}/calmodulin kinase IV	19 ASTEN**L**VPD...Y 469
34	***Leu*-vimentin**; intermediate filament protein	260 VSKPD**L**TAA...E 466

Fig. S7. Calpain-generated C-terminal (Ct) fragments of mammalian proteins that are either identified or predicted substrates of the Arg/N-degron pathway. The entries whose numbers are colored in green are the experimentally identified substrates of the Arg/N-degron pathway (39). The entries whose numbers are colored in black are predicted Arg/N-degron substrates. Each entry cites a calpain-generated Ct-fragment of a protein and

the fragment's Nt-residue (in red, using three-letter abbreviations for amino acids), followed by a brief description of uncleaved (full-length) precursor protein. A calpain cleavage site, denoted by an arrowhead, is shown using single-letter abbreviations for amino acids. An enlarged P1' residue (in red) becomes N-terminal upon the cleavage. The indicated residue numbers are the number of the first shown residue of a full-length protein and the number of its last residue, respectively. All entries are mouse proteins, with the exception of #14 and #27, which are human proteins.

#1. *Glu*-Bak is the proapoptotic Ct-fragment of the apoptotic regulator BAK. Glu-BAK is generated by calpain-1 in vitro and is apparently formed in vivo as well (40). #2. *Arg*-Bid. Bid is a 22 kDa member of the BCL-2 family of apoptosis regulators (41-43). Although full-length Bid is already a proapoptotic protein, its Ct-fragments, which can be naturally produced by activated caspases, calpains or granzyme B, can be even more active than intact Bid as proapoptotic protein fragments (44). The cleavage of Bid by calpains produces the 14 kDa Arg^{71}-Bid fragment (44-47) that has been shown to be a short-lived substrate of the Arg/N-degron pathway (48). #3. *Asp*-Bcl_{XL}. Bcl_{XL} is a 26 kDa antiapoptotic regulatory protein (41, 49). Under conditions that include glucose and oxygen deprivation, Bcl_{XL} can be cleaved by activated calpain-1, resulting in the 21 kDa Asp^{61}-Bcl_{XL} Ct-fragment. In contrast to full-length Bcl_{XL}, Asp^{61}-Bcl_{XL} has proapoptotic activity (50), and has been shown to be a short-lived substrate of the Arg/N-degron pathway (48). #4. *Arg*-c-Fos is the Ct-fragment of the c-Fos transcriptional regulator. c-Fos is targeted for conditional degradation through more than one degron, including the path that includes the cleavage by calpains (51). #5. *Glu*-IκBα is the Ct-fragment of the IκBα subunit of the NFκB-IκBα complex in which the NFκB transcriptional regulator is inhibited by IκBα. The IκBα subunit is targeted for degradation either through a conditional phosphodegron or through a specific calpain-mediated cleavage (52). #6. *Arg*-Igfbp2 is the calpain-generated Ct-fragment of the insulin-like growth factor binding protein-2 (53). #7. *Asp*-Capns1 is the Ct-fragment of the calpain regulatory subunit that is cleaved by activated calpains (54, 55). #8. *Arg*-Atp2b2 is the Ct-fragment of the transmembrane Atp2b2 plasma membrane Ca^{2+} pump (PMCA) that ejects Ca^{2+} from cells. This pump is activated either by the binding of Ca^{2+}/calmodulin or by the calpain-mediated truncation of Atp2b2 that generates the *Arg*-Atp2b2 fragment and thereby activates the pump (56). #9. *Leu*-Capn1 is the auto-generated, catalytically active Ct-fragment of calpain-1 (57, 58). #10. *Arg*-Ankrd2. Ankrd2 (Marp2, Arpp), a member of the MARP (muscle ankyrin repeat protein) family, functions as a negative regulator of muscle differentiation (59). Calpains can produce the 30 kDa Arg^{103}-Ankrd2 Ct-fragment (60). #11. *Tyr*-Grm1. Grm1 is the Ct-fragment of the mGluR1α transmembrane metabotropic glutamate receptor (61). Receptors containing the calpain-truncated mGluR1α Ct-fragment can elevate cytosolic Ca^{2+} but cannot activate PI_3K-Akt signaling pathways, in contrast to uncleaved receptors (61, 62). #12. *Lys*-Ica512. Ica512 (Ptprn) is a member of the transmembrane receptor protein phosphatase family (63). The 43 kDa calpain-generated mouse Lys^{609}-Ica512 Ct-fragment enters the nucleus and acts as a transcriptional regulator (63, 64). #13. *Phe*-GluN2a. GluN2a (NMDA-R2a) is a subunit of the N-methyl-D-aspartate receptor (NMDAR), a glutamate receptor that can function as a ligand-gated Ca^{2+} channel (65, 66). The GluN2b subunit of NMDAR can also be cleaved by calpains (67). Ct-fragments of NR2A and NR2B contain domains required for the association of these subunits with other synaptic proteins. NMDAR receptors lacking a Ct-region of GluN2a could function as glutamate-gated Ca^{2+} channels but the intracellular traffic of cleaved receptors and their electrophysiological properties were altered (68). #14. *Asn*-DSCR1

(Rcan1) is the calpain-generated Ct-fragment of the Down syndrome critical region 1 protein Dscr1, which binds to Raf1, inhibits the phosphatase activity of calcineurin, and enhances its degradation. The *Asn*-DSCR1 fragment does not bind to the Raf1 kinase (69). #15. *Arg*-Glyt1a is the Ct-fragment of the transmembrane Glyt1a glycine transporter (70). Another Gly transporter, Glyt1b, is also cleaved by calpains, yielding the *Arg*-Glyt1b fragment (70). These Ct fragments are still active as transporters but are impaired in their ability to remove Gly (an inhibitory neurotransmitter) from synaptic clefts (70). #16. *Asn*-$Ca_v1.1$ is the Ct fragment of the voltage-gated transmembrane Ca^{2+} channel. This (apparently) calpain-generated fragment is noncovalently associated with the rest of the channel and can inhibit its activity. Upon dissociation from the channel, the *Asn*-$Ca_v1.1$ fragment enters the nucleus and functions as a transcriptional regulator (71-74). #17. *Lys*-cortactin is the Ct-fragment of cortactin, a protein that regulates actin polymerization (75). #18. *Asn*-Bfl-1. Bfl-1 is antiapoptotic regulatory protein whose cleavage by calpain-1 generates the *Asn*72-Bfl-1 proapoptotic Ct-fragment (76). #19. *Arg*-dystrophin is the calpain-generated Ct-fragment of a major cytoskeletal protein in the skeletal muscle (77). #20. *Gln*-Ryr1 is the Ct-fragment of the Ryr1 ryanodine receptor, a Ca^{2+} channel in the ER (78) that mediates the efflux of Ca^{2+} from the ER into the cytosol. Calpain-mediated cleavage of Ryr1 increases Ca^{2+} efflux (79). #21. *Arg*-Mef2d is the Ct-fragment of the Mef2d myocyte enhancer factor 2d, a transcriptional regulator that contributes to neuronal survival, development, and synaptic plasticity (80). #22. *Gln*-talin is the calpain-generated Ct-fragment of talin, an adaptor protein that interacts with the integrin family of cell adhesion transmembrane proteins (56, 81, 82). #23. *Arg*-p39 is the calpain-generated Ct-fragment of the p39 activator of the Cdk5 protein kinase (83). The indicated cleavage site is located immediately downstream of two other closely spaced (and strongly conserved) calpain cleavage sites in p39. A cleavage at any one of these sites yields a predicted Arg/N-degron substrate. #24. *Gln*-Egfr is one of calpain-generated Ct-fragments of the transmembrane epidermal growth factor (EGF) receptor protein kinase (84). Remarkably, all 7 calpain cleavage sites in the cytosol-exposed domain of the 170-kDa EGFR contain P1' residues that are recognized as destabilizing by the Arg/N-degron pathway (84). #25. *Leu*-β-catenin is the calpain-generated Ct-fragment of β-catenin, a conditionally short-lived cytoskeletal protein and transcriptional regulator. The *Leu*-β-catenin fragment is a nuclear protein that activates specific genes in conjunction with other transcription factors (85). #26. *Leu*-NF2 is the calpain-generated Ct-fragment of NF2 (merlin), a tumor suppressor and cytoskeletal protein. Loss of function NF2 mutants result in autosomal-dominant neurofibromatosis, a predisposition to specific kinds of brain tumors (86). #27. *Arg*-caspase-9 is the Ct-fragment of caspase-9, which can be inactivated by calpains (87), followed by the (predicted) degradation of the Arg-caspase-9 Ct-fragment by the Arg/N-degron pathway. #28. *Leu*-troponin T2 is the Ct-fragment of the cardiac troponin T that is produced by calpain-1 in the troponin-containing cardiac myofibril complex (88). #29. *Lys*-PkCα is the calpain-generated Ct-fragment of PkCα, a broadly expressed Ser/Thr kinase of the PKC family (89). Being catalytically active but no longer controlled by the regulatory Nt-domain of the full-length PkCα, the Lys-PkCα fragment can be toxic, for example, upon its formation in an ischemic heart (90). #30. *Leu*-Rad21 is the calpain-generated Ct-fragment of the Scc1/Rad21 subunit of the chromosome-associated cohesin complex (91). The calpain-mediated generation of *Leu*-Rad21 contributes to the control of chromosome cohesion/segregation, together with processes that include the separase-mediated cleavage of the same Rad21 subunit of cohesin (91-94). #31. *Phe*-PkCγ is the calpain-generated Ct-fragment of PkCγ, a Ser/Thr kinase of the PKC

family (89). The *Phe*-PkCγ fragment is constitutively active as a kinase, because it lacks the regulatory Nt-domain of the full-length PkCγ kinase (89). #32. *Leu*-$STEP_{33}$ is the Ct-fragment of the striatal-enriched $STEP_{61}$ phosphatase, a brain-specific Tyr-phosphatase whose substrates include the MAPK-family kinases Erk1/2 and p38. The calpain-generated *Leu*-$STEP_{33}$ fragment lacks phosphatase activity (95). #33. *Leu*-Camk-IV is the calpain-generated Ct-fragment of the Ca^{2+}/calmodulin-dependent kinase-IV. This fragment lacks kinase activity (96). #34. *Leu*-vimentin is the calpain-generated Ct-fragment of vimentin, a component of intermediate filaments (97).

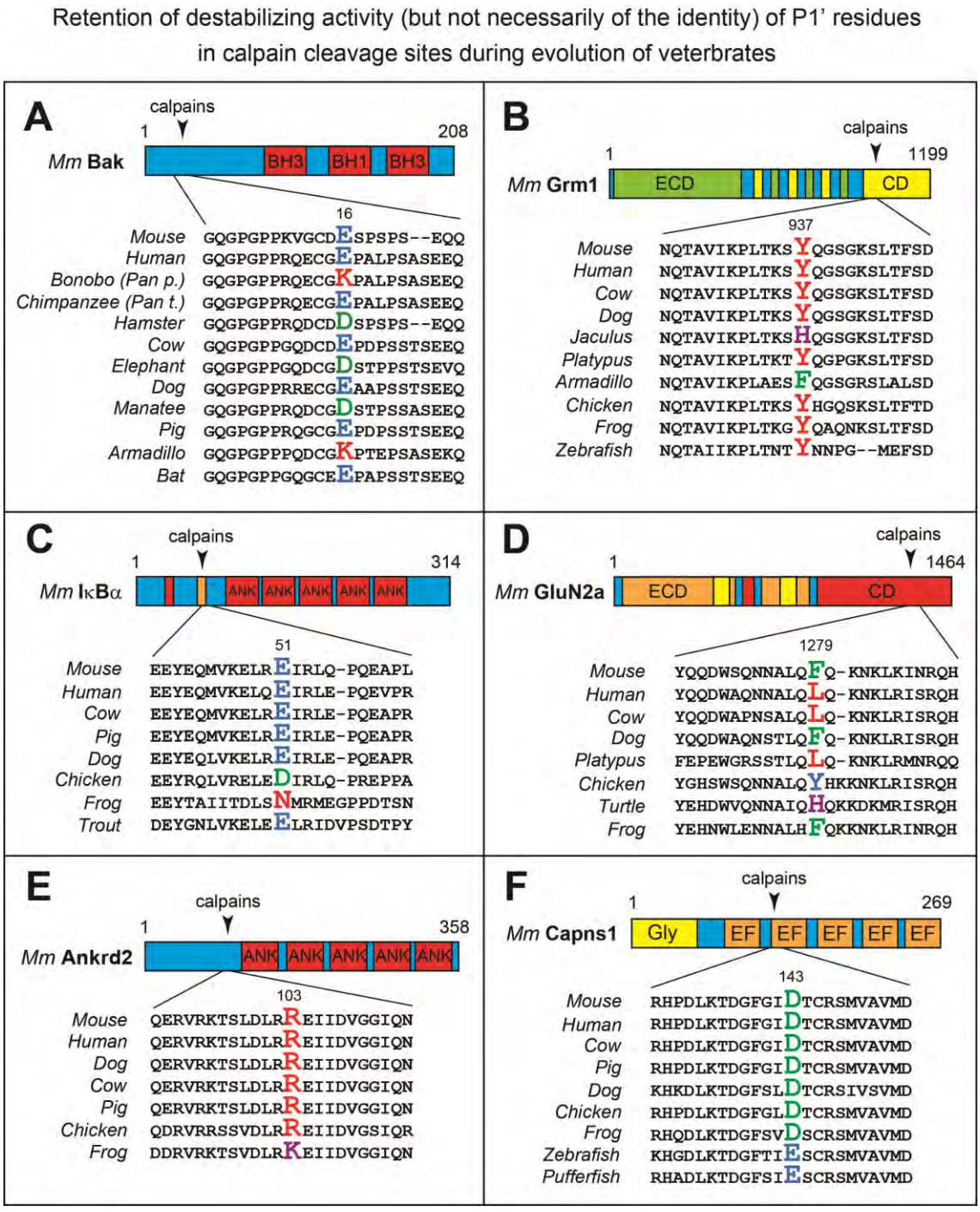

Fig. S8. Retention of destabilizing activity (but not necessarily of the identity) of P1' residues in calpain cleavage sites during evolution of vertebrates. Arrowheads indicate calpain cleavage sites. P1' residues, which become N-terminal upon the cleavage, are larger and colored. Domain organization and approximate location of a calpain cleavage site are indicated for each protein. The diagrams and indicated residue numbers are of mouse (*Mus musculus* (*Mm*)) proteins. (*A*) Bak. (*B*) Grm1. (*C*) IκBα. (*D*) GluN2a. (*E*) Ankrd2. (*F*) Capsn1. Note the retention, during evolution, of destabilizing activity (but not necessarily of the identity) of P1' residues in these proteins. See also Fig. S7.

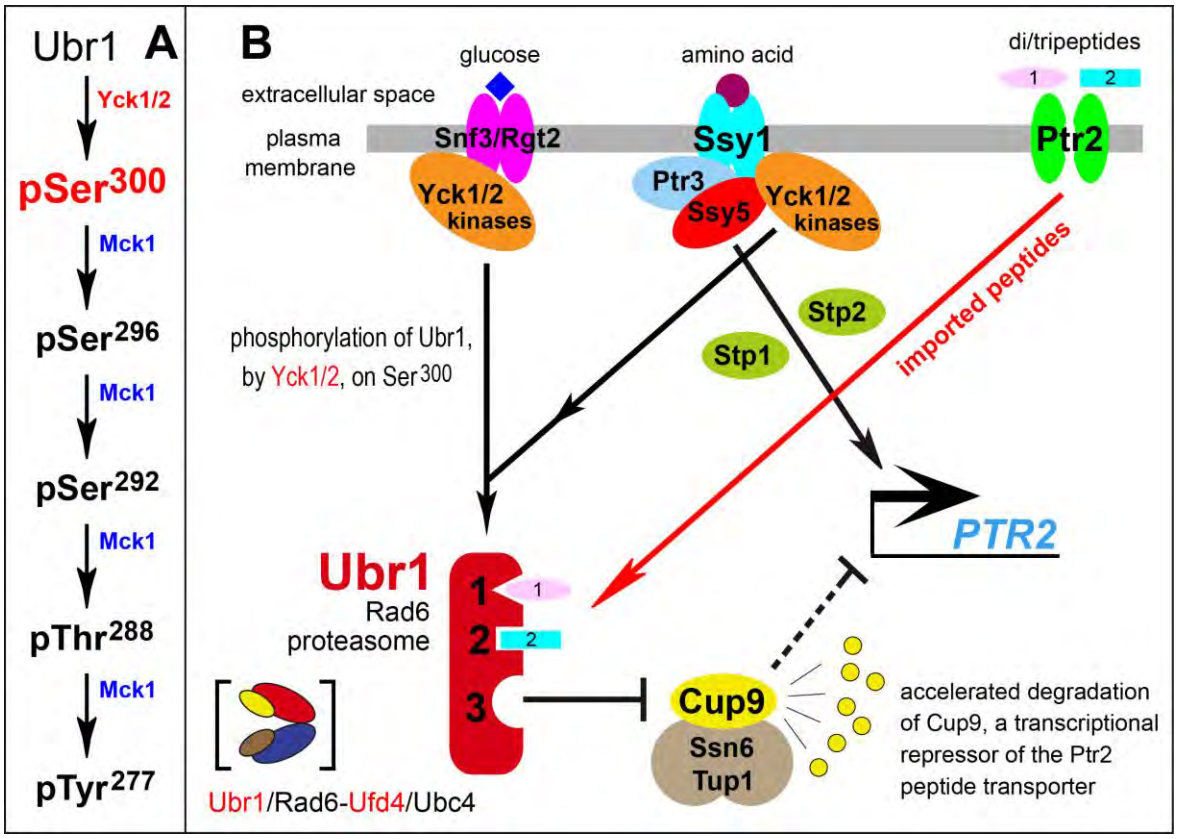

Fig. S9. Regulation of peptide import by the Arg/N-degron pathway in *S. cerevisiae*. (*A*) The "primed" cascade of Ubr1 phosphorylation in which the Yck1/Yck2-mediated phosphorylation on Ser^{300} of Ubr1 is essential for the normal regulation of peptide import (98). (*B*) Ubr1-mediated regulation of peptide import, and the involvement of the SPS pathway (99-103). Cup9 is a transcriptional repressor of the regulon that includes *PTR2*, which encodes the major importer of di/tripeptides. In the absence of Ubr1 (in *ubr1Δ* cells), Cup9 becomes relatively long-lived, accumulates to higher levels, and extinguishes expression of Ptr2. Therefore *ubr1Δ* cells cannot import di/tripeptides. In wild-type (*UBR1*) cells growing in the absence of extracellular di/tripeptides, a low but non-zero number of Ubr1 molecules have their third substrate-binding site "open" (not autoinhibited) and therefore can target Cup9 for degradation ($t_{1/2} \sim 5$ min) via its internal degron, resulting in a weak but non-zero expression of the Ptr2 transporter. In wild-type (*UBR1*) cells growing in the presence of extracellular di/tripeptides (some of which bear type-1 and type-2 destabilizing Nt-residues), the imported di/tripeptides interact with the type-1 and type-2 binding sites of Ubr1. This binding allosterically increases the fraction of Ubr1 molecules whose third (Cup9-specific) site is "open" (active). The result is a decrease in the $t_{1/2} \sim$ of Cup9 from ~5 min to 1 min or less, leading to negligible steady-state levels of Cup9, and consequently to a strong induction of the Ptr2 transporter (101-103). Also shown is the amino acid-sensing SPS pathway which can influence the import of peptides at least in part through the Yck1/Yck2-mediated phosphorylation of Ubr1 on Ser^{300}. This phosphorylation is required (through a mechanism that remains to be understood) for normal levels of Ubr1 activity in the Ptr2-Cup9-Ubr1 circuit (98).

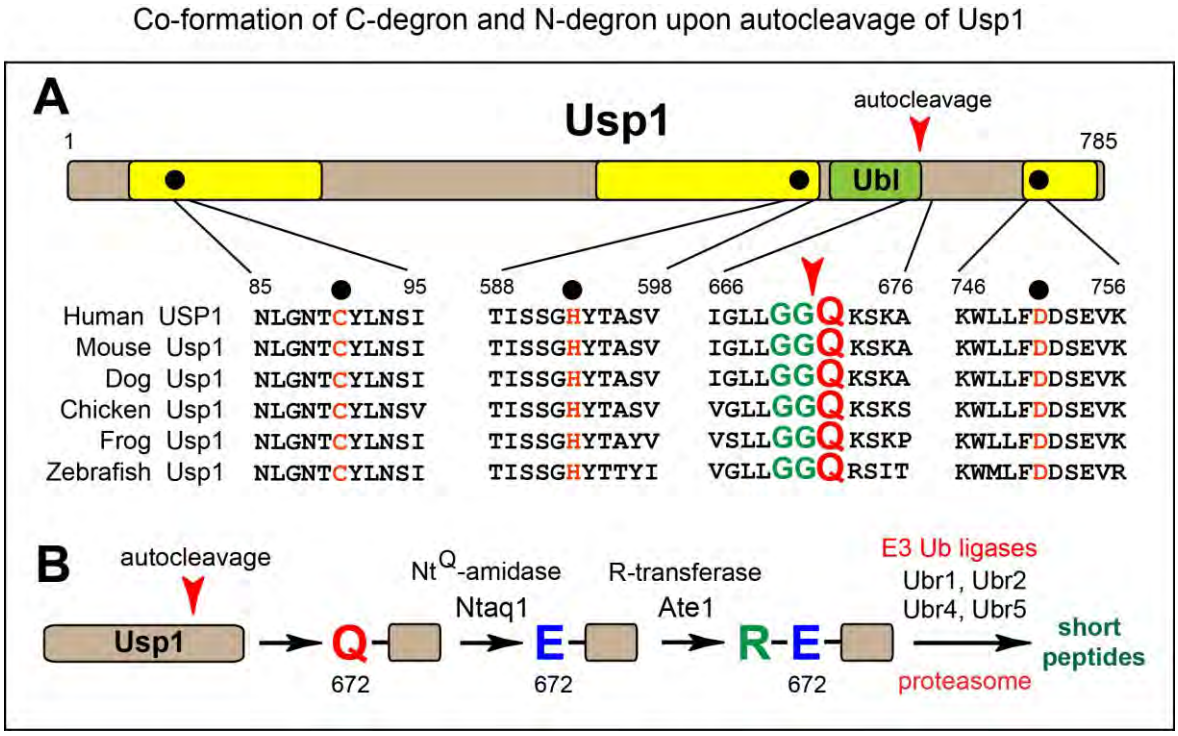

Fig. S10. Co-formation of C-degron and N-degron upon Usp1 autocleavage (18, 19, 28).

(*A*) Sequence alignments, among vertebrates, of the catalytic domains of Usp1 and its autocleavage site, with amino acid residues numbered as in human Usp1. Key residues of the catalytic triad are in red, and are also marked by black circles. The Ub-like domain (Ubl, in green), with its last Gly-Gly 2-residue sequence (in green and a larger font size) abuts the autocleavage site of Usp1. The Gln (Q) P1' residue of the autocleavage site is in red and a larger font size. This residue, #672 in human Usp1 and #670 in mouse Usp1, becomes N-terminal after the autocleavage of Usp1.

(*B*) The enzymatic cascade that includes the autocleavage of Usp1 and the Arg/N-degron pathway, which destroys the *Gln*-Usp1^{Ct} fragment (28).

As described in the main text, the Usp1 deubiquitylase (DUB) forms a heterodimer with Uaf1, a non-DUB protein. Usp1 can autocleave immediately after its internal Gly-Gly sequence (104). Nevertheless, the DUB activity of autocleaved Usp1 can be transiently maintained, inasmuch as Usp1^{Nt}, the Nt-fragment of autocleaved Usp1, can remain bound to the Gln-Usp1^{Ct} Ct-fragment within the cleaved Usp1-Uaf1 heterodimer. The Ct-sequence Gly-Gly of the Usp1^{Nt} Nt-fragment can act as a C-degron, which is recognized by the Kldch2 adaptor subunit of the Clr2 Ub ligase. In the resulting mechanism (not yet analyzed in detail), Uaf1 can hold together two Usp1 fragments, allowing them to function, temporarily, as a DUB enzyme, until successful attacks, on both fragments, by the N-degron and C-degron pathways. Usp1 is the first experimentally addressed setting in which an N-degron and a C-degron can be co-created upon a cleavage (self-cleavage, in this case) of a full-length protein((18, 19, 28, 104) and refs. therein)..

SI References

1. Bachmair A, Finley D, Varshavsky A (1986) In vivo *h*alf-life of a protein is a function of its amino-terminal residue. *Science* 234:179-186.
2. Varshavsky A (2005) Ubiquitin fusion technique and related methods. *Meth. Enzymol.* 399:777-799.
3. Aksnes H, Drazic A, Marie M, Arnesen T (2016) First things first: vital protein marks by N-terminal acetyltransferases. *Trends Biochem. Sci.* 9:746-760.
4. Frottin F*, et al.* (2006) The proteomics of N-terminal methionine cleavage. *Mol. Cell. Proteomics* 5:2336-2349.
5. Xiao Q, Zhang F, Nacev BA, Liu JO, Pei D (2010) Protein N-terminal processing: substrate specificity of Escherichia coli and human methionine aminopeptidases. *Biochemistry* 49:5588-5599.
6. Ji CH, Kwon YT (2017) Crosstalk and interplay between the ubiquitin-proteasome system and autophagy. *Mol. Cells* 40:441-449.
7. Grumati P, Dikic I (2018) Ubiquitin signaling and autophagy. *J. Biol. Chem.* 293:5404-5413.
8. Shim SM*, et al.* (2018) The endoplasmic reticulum-residing chaperone BiP is short-lived and metabolized through N-terminal arginylation. *Sci. Signal.* 11:eaan0630.
9. Yoo YD*, et al.* (2018) N-terminal arginylation generates a bimodal degron that modulates autophagic proteolysis. *Proc. Natl. Acad. Sci. USA* 115:E2716-E2724.
10. Cha-Molstad H*, et al.* (2018) Regulation of autophagic proteolysis by the N-recognin SQSTM1/p62 of the N-end rule pathway. *Autophagy* 14:359-361.
11. Cha-Molstad H*, et al.* (2015) Amino-terminal arginylation targets endoplasmic reticulum chaperone BiP for autophagy through p62 binding. *Nat. Cell Biol.* 17:917-929.
12. Cha-Molstad H*, et al.* (2017) p62/SQSTM1/sequestosome-1 is an N-recognin of the N-end rule pathway which modulates autophagosome biogenesis. *Nature communications* 8:102.
13. Cha-Molstad H*, et al.* (2016) Modulation of SQSTM1/p62 activity by N-terminal arginylation of the endoplasmic reticulum chaperone HSPA5/GRP78/BiP. *Autophagy* 12:426-428.
14. Lee DH*, et al.* (2018) PARK7 modulates autophagic proteolysis through binding to the N-terminally arginylated form of the molecular chaperone HSPA5. *Autophagy* 14:1870-1885.
15. Kwon DH*, et al.* (2018) Insights into degradation mechanism of N-end rule substrates by p62/SQSTM1 autophagy adapter. *Nature communications* 9:3291.
16. Hu R-G*, et al.* (2006) Arginyl-transferase, its specificity, putative substrates, bidirectional promoter, and splicing-derived isoforms. *J. Biol. Chem.* 281:32559-32573.
17. Varshavsky A (2011) The N-end rule pathway and regulation by proteolysis. *Prot. Sci.* 20:1298-1345.
18. Koren I*, et al.* (2018) The eukaryotic proteome is shaped by E3 ubiquitin ligases targeting C-terminal degrons. *Cell* 173:1622-1635.e1614.
19. Lin HC*, et al.* (2018) C-terminal end-directed protein elimination by CRL2 ubiquitin ligases. *Mol. Cell* 70:602-613.e603.
20. Chatr-Aryamontri A, van der Sloot A, Tyers M (2018) At long last, a C-terminal bookend for the ubiquitin code. *Mol. Cell* 70:568-571.

21. Varshavsky A (2004) Spalog and sequelog: neutral terms for spatial and sequence similarity. *Curr. Biol.* 14:R181-R183.
22. Bachmair A, Varshavsky A (1989) The degradation signal in a short-lived protein. *Cell* 56:1019-1032.
23. Inobe T, Fishbain S, Prakash S, Matouschek A (2011) Defining the geometry of the two-component proteasome degron. *Nat. Chem. Biol.* 7:161-167.
24. Johnson ES, Gonda DK, Varshavsky A (1990) Cis-trans recognition and subunit-specific degradation of short-lived proteins. *Nature* 346:287-291.
25. Prakash S, Inobe T, Hatch AJ, Matouschek A (2009) Substrate selection by the proteasome during degradation of protein complexes. *Nat. Chem Biol.* 5:29-36.
26. Schrader EK, Harstad KG, Matouschek A (2009) Targeting proteins for degradation. *Nat. Chem Biol.* 5:815-822.
27. Sriram SM, Kim BY, Kwon YT (2011) The N-end rule pathway: emerging functions and molecular principles of substrate recognition. *Nat. Rev. Mol. Cell Biol.* 12:735-747.
28. Piatkov KI, Colnaghi L, Bekes M, Varshavsky A, Huang TT (2012) The auto-generated fragment of the Usp1 deubiquitylase is a physiological substrate of the N-end rule pathway. *Mol. Cell* 48:926-933.
29. Emoto Y, Kisaki H, Manome Y, Kharbanda S, Kufe D (1996) Activation of protein kinase C-delta in human myeloid leukemia cells treated with 1-beta-D-arbabinofuranosylcytosine. *Blood* 87:1990-1996.
30. Denning MF, Wang Y, Nickoloff B, Wrone-Smith T (1998) Protein kinase Cdelta is activated by caspase-dependent proteolysis during ultraviolet radiation-induced apoptosis of human keratinocytes. *J. Biol. Chem.* 273:29995-30002.
31. Leverrier S, Vallentin A, Joubert D (2002) Positive feedback of protein kinase C proteolytic activation during apoptosis. *Biochem. J.* 368:905-913.
32. Reyland ME (2007) Protein kinase Cdelta and apoptosis. *Biochem. Soc. Trans.* 35:1001-1004.
33. LaGory EL, Sitailo LA, Denning MF (2010) The protein kinase Cdelta catalytic fragment is critical for maintenance of the G2/M DNA damage checkpoint. *J. Biol. Chem.* 285:1879-1887.
34. Datta R, Kojima H, Yoshida K, Kufe D (1997) Caspase-3-mediated cleavage of protein kinase C-theta in induction of apoptosis. *J. Biol. Chem.* 272:20317-20320.
35. Wu YM, Huang CL, Kung HJ, Huang CYF (2001) Proteolytic activation of Etk/Bmx tyrosine kinase by caspases. *J. Biol. Chem.* 276:17672-17678.
36. Sabourin LA, Seale P, Wagner J, Rudnicki MA (2000) Caspase 3 cleavage of the Ste20-related kinase SLK releases and activates an apoptosis-inducing kinase domain and an actin-disassembling region. *Mol. Cell. Biol.* 20:684-696.
37. Arnold R, Liou J, Drexler HCA, Weiss A, Friedemann K (2001) Caspase-mediated cleavage of hematopoietic progenitor kinase 1 (HPK1) converts an activator of NF-kappaB into an inhibitor of NF-kappaB. *J. Biol. Chem.* 276:14675-14684.
38. Chen F, Arseven OK, Cryns VL (2004) Proteolysis of the mismatch repair protein MLH1 by caspase-3 promotes DNA damage-induced apoptosis. *J. Biol. Chem.* 279:27542-27548.
39. Piatkov KI, Oh J-H, Liu Y, Varshavsky A (2014) Calpain-generated natural protein fragments as short-lived substrates of the N-end rule pathway. *Proc. Natl. Acad. Sci. USA* 111:E817-E826.

40. Moldoveanu T, *et al.* (2006) The X-ray structure of a BAK homodimer reveals an inhibitory zinc binding site. *Mol. Cell* 24:677-688.
41. Green DR (2011) *Means to an End: Apoptosis and Other Cell Death Mechanisms.* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).
42. Yin X-M (2006) Bid, a BH3-only multifunctional molecule, is at the cross road of life and death. *Gene* 369:7-19.
43. Billen LP, Shamas-Din A, Andrews DW (2009) Bid: a Bax-like BH3 protein. *Oncogene* 27:S93-S104.
44. Chen M, *et al.* (2001) Bid is cleaved by calpain to an active fragment in vitro and during myocardial ischemia/reperfusion. *J. Biol. Chem.* 276:30724-30728.
45. Mandic A, *et al.* (2002) Calpain-mediated Bid cleavage and calpain-independent Bak modulation: two separate pathways in cisplatin-induced apoptosis. *Mol. Cell. Biol.* 22:3003-3013.
46. Gil-Parrado S, *et al.* (2002) Ionomycin-activated calpain triggers apoptosis. A probable role for Bcl-2 family members. *J. Biol. Chem.* 277:27217-27226.
47. Cabon L, *et al.* (2012) BID regulates AIF-mediated caspase-independent necroptosis by promoting BAX activation. *Cell Death Diff.* 19:245-256.
48. Piatkov KI, Brower CS, Varshavsky A (2012) The N-end rule pathway counteracts cell death by destroying proapoptotic protein fragments. *Proc. Natl. Acad. Sci. USA* 109:E1839-E1847.
49. García-Sáez AJ, Ries J, Orzáez M, Pérez-Payà E, Schwille P (2009) Membrane promotes tBID interaction with BCL(XL). *Nat. Struct. Mol. Biol.* 16:1178-1185.
50. Nakagawa T, Yuan J (2000) Cross-talk between two cysteine protease families: activation of caspase-12 by calpain in apoptosis. *J. Cell Biol.* 150:887-894.
51. Pariat M, *et al.* (2000) The sensitivity of c-Jun and c-Fos proteins to calpains depends on conformational determinants of the monomers and not on formation of dimers. *Biochem. J.* 345:129-138.
52. Schaecher K, Goust JM, Banik NL (2004) The effects of calpain inhibition on IkB alpha degradation after activation of PBMCs: identification of the calpain cleavage sites. *Neurochem. Res.* 29:1443-1451.
53. Berg U, Bang P, Carlsson-Skwirut C (2007) Calpain proteolysis of insulin-like growth factor binding protein (IGFBP) -2 and -3, but not of IGFBP-1. *Biol. Chem.* 388:859-863.
54. Ono Y, Sorimachi H (2012) Calpains: an elaborate proteolytic system. *Biochim. Biophys. Acta* 1824:224-236.
55. Goll DE, Thompson VF, Li H, Wei W, Cong J (2003) The calpain system. *Physiol. Rev.* 83:731-801.
56. Di Leva F, Domi T, Fedrizzi L, Lim d, Carafoli E (2008) The plasma membrane Ca2+ ATPase of animal cells: Structure, function and regulation. *Arch. Biochem. Biophys.* 476:65-74.
57. Brown N, Crawford C (1993) Structural modifications associated with the change in Ca2+ sensitivity on activation of m-calpain. *FEBS Lett.* 322(1):65-68.
58. Molinari M, Anagli J, Carafoli E (1994) Ca^{2+}-activated neutral protease is active in the erythrocyte membrane in its noautolized 80-kDa form. *J. Biol. Chem.* 269:27992-27995.
59. Bean C, Facchinello N, Faulkner G, Lanfranchi G (2008) The effects of Ankrd2 alteration indicate its involvement in cell cycle regulation during muscle differentiation. *Biochim. Biophys. Acta* 1783:1023-1035.

60. Hayashi C, *et al.* (2008) Multiple molecular interactions implicate the connectin/titin N2A region as a modulating scaffold for p94/calpain 3 activity in skeletal muscle. *J. Biol. Chem.* 283:14801-14814.
61. Xu W, *et al.* (2007) Calpain-mediated mGluR1alpha truncation: a key step in excitotoxicity. *Neuron* 53:399-412.
62. Niswender CM, Conn PJ (2010) Metabotropic glutamate receptors: physiology, pharmacology, and disease. *Annu. Rev. Pharmacol. Toxicol.* 50:295-322.
63. Trajkovski M, *et al.* (2004) Nuclear translocation of an ICA512 cytosolic fragment couples granule exocytosis and insulin expression in beta-cells. *J. Cell Biol.* 167:1063-1074.
64. Ort T, *et al.* (2001) Dephosphorylation of beta2-syntrophin and Ca2+/mu-calpain-mediated cleavage of ICA512 upon stimulation of insulin secretion. *EMBO J.* 20:4013-4023.
65. Bi R, Rong Y, Bernard A, Khrestchatisky M, Baudry M (2000) Src-mediated tyrosine phosphorylation of NR2 subunits of N-methyl-D-aspartate receptors protects from calpain-mediated truncation of their C-terminal domains. *J. Biol. Chem.* 275:26477-26483.
66. Collingridge GL, Olsen RW, Peters J, Spedding M (2009) A nomenclature for ligand-gated ion channels. *Neuropharmacology* 56:2-5.
67. Simpkins KL, *et al.* (2003) Selective activation induced cleavage of the NR2B subunit by calpain. *J. Neurosci.* 23:11322-11331.
68. Steigerwald F, *et al.* (2000) C-Terminal truncation of NR2A subunits impairs synaptic but not extrasynaptic localization of NMDA receptors. *J. Neurosci.* 20:4573-4581.
69. Cho Y-J, Abe M, Kim SY, Sato Y (2005) Raf-1 is a binding partner of DSCR1. *Arch. Biochem. Biophys.* 439:121-128.
70. Baliova M, Jursky F (2005) Calpain sensitive regions in the N-terminal cytoplasmic domains of glycine transporters GlyT1A and GlyT1B. *Neurochem. Res.* 30:1093-1100.
71. Catterall WA (2011) Voltage-gated calcium channels. *Cold Spring Harb. Perspect. Biol.* 3:a003947.
72. Hulme JT, *et al.* (2005) Sites of proteolytic processing and noncovalent association of the distal C-terminal domain of CaV1.1 channels in skeletal muscle. *Proc. Natl. Acad. Sci. USA* 102:5274–5279.
73. Hulme JT, Yarov-Yarovoy V, Lin TWC, Scheuer T, Catterall WA (2006) Autoinhibitory control of the CaV1.2 channel by its proteolytically processed distal C-terminal domain. *J. Physiol.* 576:87-102.
74. Gomez-Ospina N, Tsuruta F, Barreto-Chang O, Hu L, Dolmetsch R (2006) The C-terminus of the L-type voltage-gated calcium channel Ca(V)1.2 encodes a transcription factor. *Cell* 127:591-606.
75. Perrin BJ, Amann KJ, Huttenlocher A (2006) Proteolysis of cortactin by calpain regulates membrane protrusion during cell migration. *Mol. Biol. Cell* 17:239–250.
76. Valero JG, *et al.* (2012) μ-Calpain conversion of antiapoptotic Bfl-1 (BCL2A1) into a prodeath factor reveals two distinct alpha-helices inducing mitochondria-mediated apoptosis. *PLoS One* 7:e38620.
77. Yoshida M, Suzuki A, Shimizu T, Ozawa E (1992) Proteinase-sensitive sites on isolated rabbit dystrophin. *J. Biochem.* 112:433-439.

78. Wu Y, *et al.* (1997) Functional interactions between cytoplasmic domains of the skeletal muscle Ca2+ release channel. *J. Biol. Chem.* 272:25051-25061.
79. Shoshan-Barmatz V, Weil S, Meyer H, Varsanyi M, Heilmeyer LM (1994) Endogenous, Ca(2+)-dependent cysteine-protease cleaves specifically the ryanodine receptor/Ca2+ release channel in skeletal muscle. *J. Membr. Biol.* 142:281-288.
80. Wei G, *et al.* (2012) Calpain-mediated degradation of myocyte enhancer factor 2D contributes to excitotoxicity by activation of extrasynaptic NMDA receptors. *J. Biol. Chem.* 287:5797-5805.
81. Rees DJG, Ades SE, Singer SJ, Hynes RO (1990) Sequence and domain structure of talin. *Nature* 347:685-689.
82. Bate N, *et al.* (2012) Talin contains a C-terminal calpain-2 cleavage site important in focal adhesion dynamics. *PLoS One* 7:e34461.
83. Patzke H, Tsai LH (2002) Calpain-mediated cleavage of the cyclin-dependent kinase-5 activator p39 to p29. *J. Biol. Chem.* 277:8054-8060.
84. Gregoriou M, Willis AC, Pearson MA, Crawford C (1994) The calpain cleavage sites in the epidermal growth factor receptor kinase domain. *Eur. J. Biochem.* 223:455-464.
85. Abe K, Takeichi M (2007) NMDA-receptor activation induces calpain-mediated beta-catenin cleavages for triggering gene expression. *Neuron* 53:387-397.
86. Kimura Y, ., *et al.* (1998) The involvement of calpain-dependent proteolysis of the tumor suppressor NF2 (merlin) in schwannomas and meningiomas. *Nat. Med.* 4:915-922.
87. Chua BT, Guo K, Li P (2000) Direct cleavage by the calcium-activated protease calpain can lead to inactivation of caspases. *J. Biol. Chem.* 275:5131-5135.
88. Zhang Z, Biesiadecki BJ, Jin JP (2006) Selective deletion of the NH2-terminal variable region of cardiac troponin T in ischemia reperfusion by myofibril-associated mu-calpain cleavage. *Biochemistry* 45:11681-11694.
89. Kishimoto A, *et al.* (1989) Limited proteolysis of protein kinase C subspecies by calcium-dependent neutral protease (calpain). *J. Biol. Chem.* 264:4088-4092.
90. Kang MY, *et al.* (2010) Receptor-independent cardiac protein kinase Calpha activation by calpain-mediated truncation of regulatory domains. *Circ. Res.* 107:903-912.
91. Panigrahi AK, Zhang N, Mao Q, Pati D (2011) Calpain-1 cleaves Rad21 to promote sister chromatid separation. *Mol. Cell. Biol.* 31:4335-4447.
92. Rao H, Uhlmann F, Nasmyth K, Varshavsky A (2001) Degradation of a cohesin subunit by the N-end rule pathway is essential for chromosome stability. *Nature* 410:955-960.
93. Hauf S, Waizenegger IC, Peters J-M (2001) Cohesin cleavage by separase required for anaphase and cytokinesis in human cells. *Science* 293:1320-1323.
94. Varshavsky A (2008) Discovery of cellular regulation by protein degradation. *J. Biol. Chem.* 283:34469-34489.
95. Xu J, *et al.* (2009) Extrasynaptic NMDA receptors couple preferentially to excitotoxicity via calpain-mediated cleavage of STEP. *J Neurosci.* 29:9330-9343.
96. McGinnis KM, Whitton MM, Gnegy ME, Wang KK (1998) Calcium/calmodulin-dependent protein kinase IV is cleaved by caspase-3 and calpain in SH-SY5Y human neuroblastoma cells undergoing apoptosis. *J. Biol. Chem.* 273:19993-20000.

97. Fischer S, Vandekerckhove J, Ampe C, Traub P, Weber K (1986) Protein-chemical identification of the major cleavage sites of the Ca2+ proteinase on murine vimentin, the mesenchymal intermediate filament protein. *Biol. Chem. Hoppe Seyler* 367:1147-1152.
98. Hwang CS, Varshavsky A (2008) Regulation of peptide import through phosphorylation of Ubr1, the ubiquitin ligase of the N-end rule pathway. *Proc. Natl. Acad. Sci. USA* 105:19188-19193.
99. Alagramam K, Naider F, Becker JM (1995) A recognition component of the ubiquitin system is required for peptide transport in Saccharomyces cerevisiae. *Mol. Microbiol.* 15:225-234.
100. Byrd C, Turner GC, Varshavsky A (1998) The N-end rule pathway controls the import of peptides through degradation of a transcriptional repressor. *EMBO J.* 17:269-277.
101. Turner GC, Du F, Varshavsky A (2000) Peptides accelerate their uptake by activating a ubiquitin-dependent proteolytic pathway. *Nature* 405:579-583.
102. Du F, Navarro-Garcia F, Xia Z, Tasaki T, Varshavsky A (2002) Pairs of dipeptides synergistically activate the binding of substrate by ubiquitin ligase through dissociation of its autoinhibitory domain. *Proc. Natl. Acad. Sci. USA* 99:14110-14115.
103. Xia Z, Turner GC, Hwang C-S, Byrd C, Varshavsky A (2008) Amino acids induce peptide uptake via accelerated degradation of CUP9, the transcriptional repressor of the PTR2 peptide transporter. *J. Biol. Chem.* 283:28958-28968.
104. Huang TT, *et al.* (2006) Regulation of monoubiquitinated PCNA by DUB autocleavage. *Nat. Cell Biol.* 8:339-347.