Title
End-Binding E3 Ubiquitin Ligases Enable Protease Signaling

Journal
ACS Chemical Biology, 16(11)

ISSN
1554-8929

Authors
Ravalin, Matthew
Basu, Koli
Gestwicki, Jason E
et al.

Publication Date
2021-11-19

DOI
10.1021/acschembio.9b00621

Peer reviewed

Published in final edited form as:

ACS Chem Biol. 2021 November 19; 16(11): 2047–2056. doi:10.1021/acschembio.9b00621.

End-binding E3 ubiquitin ligases enable protease signaling

Matthew Ravalin[1], **Koli Basu**[1], **Jason E. Gestwicki**[1,2], **Charles S. Craik**[1,3]

[1]Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, CA, USA.

[2]Institute for Neurodegenerative Diseases, University of California at San Francisco, San Francisco, CA, USA.

Abstract

Post-translational modifications (PTMs) direct the assembly of protein complexes. In this context, proteolysis is a unique PTM because it is irreversible; the hydrolysis of the peptide backbone generates separate fragments bearing a new N and C terminus. Proteolysis can “re-wire” protein-protein interactions (PPIs) via the recruitment of end-binding proteins to new termini. In this review, we focus on the role of proteolysis in specifically creating complexes by recruiting E3 ubiquitin ligases to new N and C termini. These complexes potentiate proteolytic signaling by “erasing” proteolytic modifications. This activity tunes the duration and magnitude of protease signaling events. Recent work has shown that the stepwise process of proteolysis, end-binding by E3 ubiquitin ligases, and fragment turnover is associated with both the nascent N terminus (*i.e.* N-degron pathways) and the nascent C terminus (*i.e.* the C-degron pathways). Here, we discuss how these pathways might harmonize protease signaling with protein homeostasis (*i.e.* proteostasis).

Keywords

End-binding: Specific recognition of the unique physiochemical properties of a protein terminus by a protein or multiprotein complex.; Modular Recognition Element (MRE): The structural feature in a protein that recognizes a PTM modified peptide; Terminal Recognition Element (TRE): The structural feature of a protein of multiprotein complex responsible for end-binding.; Post translational modification (PTM): A covalent chemical modification made to a polypeptide following biosynthesis.; Protease: An enzyme that catalyzes the hydrolysis of a peptide bond.; E3 ubiquitin ligase: An enzyme that mediates the ubiquitination of proteins; Degron: A protein feature that is important for regulating protein turnover.; Proteostasis: The cellular pathways that control protein fate.; Apoptosis: A form of programmed cell death that occurs in multicellular organisms that avoids an inflammatory response.; Pyroptosis: A form of programmed cell death that is lytic and pro-inflammatory.

I. Proteolytic control of PPIs

The biophysical underpinnings of protein-protein interactions (PPIs) are diverse.[1,2] A subset of important PPIs occur between a globular protein domain and a peptide. These

[3]Correspondence to: Charles.Craik@ucsf.edu.

interactions have garnered significant interest because of the role they play in signaling and their amenability to chemical modulation.[3,4] This class of PPI is often regulated by post-translational modifications (PTMs). In this way, the PPI is a composite readout of the competing activities of enzymes that install the PTM (E) and those that remove it (E*), as well the specificity of the binding domain that ultimately interacts with the modified peptide (modular recognition element, MRE).[5] For example, in the field of chromatin modification, PPIs are controlled by the interplay between “readers” (MRE), “writers” (E), and “erasers” (E*).[6,7]

Proteolysis is a PTM involving the hydrolysis of an amide bond in a polypeptide chain. Proteolytic activity is capable of transmitting biological signals by destroying entire proteins, removing whole domains, facilitating the maturation of structural features, and modulating PPIs (Figure 1).[8] Proteolysis is a unique PTM because it is irreversible, and thus represents a committed biological step. Given this irreversibility, proteolytic signals exist in flux rather than in a dynamic equilibrium.[9,10] Each proteolytic event generates two products, a nascent N and C terminus with opposing charges (Figure 2, panel A). The only way to recover the full-length protein is through re-synthesis, so the “eraser” function is necessarily served by degrading the cleavage products. This feature positions the ubiquitin proteasome system (UPS) to be a major contributor to protease signaling.[11] Briefly, the UPS involves a cascade of ubiquitin transfer steps that append a polyubiquitin chain to the protein and enhance its affinity for the proteasome. In mammals, the specificity of ubiquitination is controlled by the activity of 500 to 1000 different E3 ubiquitin ligases.[12] Important for this discussion, a subset of these E3 ubiquitin ligases recognize protein termini (i.e. are end-binding) and, thus, are poised to regulate the stability of the products of proteolysis (Figure 2, panel B and C).[13,14]

II. Specificity links protease and E3 function

The unique features of proteolysis as a PTM have enabled the evolution of protease-driven signaling cascades.[15] Links between terminal recognition and protein degradation have been characterized in bacteria, plants, yeast, and mammals.[16] In eukaryotes, E3 ubiquitin ligases that contain a terminal recognition element (TRE) couple the terminal chemical composition of a polypeptide substrate to its degradation. An implicit connection between a protease and E3 ligase arises from the physical interaction of each with a shared substrate (Figure 2, panel D). Thus, the steps of proteolysis and turnover are driven by overlapping specificity of the protease and the TRE for a common polypeptide sequence. For recognition of new C termini, the TRE recognizes the same sequence motif that is N-terminal to the cleaved bond (the non-prime sites). This is referred to as a C degron. Likewise, recognition at new N termini depends on the complementarity between the TRE and the amino acids C-terminal to the cleaved bond (the prime sites) (Figure 2, panel E). This motif is termed the N degron. Thus, an understanding of the molecular recognition in these systems can emerge from mapping the sequence specificity of the protease and the E3 ubiquitin ligase’s TRE. Focusing on cleavage events that generate degrons is likely to facilitate the prioritization of functional cleavages. With this framework in hand, we can leverage an understanding of overlapping specificity *in vitro* to illuminate biological signaling pathways *in vivo*.

II. The role of N degrons in protease signaling

The N-degron pathways broadly link the stability of a protein to the identity of its N terminus.[17] In eukaryotes, these pathways were classically divided into two arms. In the Ac/N-degron pathways, protein degradation is trigged by the recognition of an acetylated N-terminal sequence by dedicated E3 ubiquitin ligases.[18] The bulk of eukaryotic proteins are irreversibly acetylated co-translationally.[19,20] Acetylation can be preceded by the removal of the N-terminal Met generated by the start codon via the activity of methionine aminopeptidases (MetAPs).[21,22] The propensity for an encoded N-terminal sequence to be acetylated or demethionylated is dictated by the specificity of the MetAPs and N-terminal acetyltransferases. In general, the removal of methionine from the N terminus is favored in instances when a sterically small (Ala, Val, Ser, Thr, Cys, Gly, and Pro) residue is present at the position directly adjacent to the N-terminal Met. N-terminal acetylation subsequent to demethionylation is favored for all of these residues with the important exception of Gly and Pro.[23]

The other classical arm of the N-degron pathways are the Arg/N-degron pathways. Here, N termini bearing positively charged (Arg, Lys, His, e.g. Type I degrons) or large hydrophobic (Leu, Phe, Trp, Tyr, Ile, e.g. Type II degrons) residues are ubiquitinated by a subset of the UBR family of E3 ubiquitin ligases (UBR 1, 2, 4, and 5 in mammals).[24] These E3 ligases depend on the specific recognition of the positively charged N-terminal amine to bind to their substrates. Type 1 degrons are recognized by the UBR domain that contains a binding site capable of coordinating the positive charges of both the N terminus and the side chain of the N-terminal Arg, Lys, or His (Figure 3, panel A).[25] The structural aspects of the recognition of Type II degrons is less understood, however a discrete recognition domain in UBR1 and UBR2 is responsible for binding these sequences.[26] This domain is homologous to a TRE found in bacteria, ClpS, which binds to N-terminal hydrophobic and aromatic amino acids, directing degradation by ClpXP (Figure 3, panel B).[27] Collectively, N termini that are directly recognized and ubiquitinated by the UBR E3 ligases are referred to as “primary” N degrons.

Some N termini require chemical modification prior to being shunted through the Arg/N-degron pathways. For example, N-terminal Asp and Glu residues must first be modified with an N-terminal Arg by the Ate family of Arg transferases.[28] Given the necessity for an additional modification to sequences bearing terminal Asp and Glu residues they are referred to as “secondary” degrons. N-terminal Cys residues require oxidation by nitric oxide prior to argenylation by Ate.[29] N-terminal Asn and Gln are converted to Asp and Glu respectively via side chain deamidation by Ntaq1 before argenylation can occcur.[30] As they require two modifications prior to recognition by the Arg/N-degron E3 ligases, sequences bearing terminal Cys, Asn, or Gln are referred to as “tertiary” degrons. In the context of this discussion, the modifications that produce secondary and tertiary degrons provide a bridge between proteolysis and binding to UBR domains. A poignant example of this logic is evident in the processsing of the Ubiquitin specific peptidase 1 (USP1). USP1 auto-proteolyzes to yield a nascent N terminus bearing a conserved Gln residue. The degradation of this USP1 fragment requires the activity of both Ntaq1 and Ate1.[31] Therefore, although

USP1 does not generate a N-terminal Arg directly, its prime-side specificity still links it to the UBR E3 ligases.

Although N-degron pathways were first described over thirty years ago, the biological functions that they direct are still being uncovered. For example, N-terminal Gly and Pro residues are generated by MetAP activity but resist acetylation and degradation by both the Ac/N-degron and Arg/N-degron pathways. Recently it was discovered that GID4, a component of the GID E3 ubiquitin ligase complex in *S. cerevisiae*, specifically binds proteins bearing a N-terminal Pro.[32] This pathway is critical for the rapid degradation of gluconeogenic enzymes when *S. cerevisiae* are transitioned from acetate- or ethanol-based growth media to glucose-based media. Structural and biophysical evidence demonstrated that the human GID4 ortholog specifically recognized N-terminal proline and required a free amine at the N terminus (Figure 3, panel C).[33]

An N-degron pathway for unacetylated Gly residues has also recently been described. Briefly, a functional genomics screen identified $Cul2^{ZYG11b}$ and $Cul2^{ZER1}$, as the E3 ubiquitin ligase complexes responsible for the degradation of polypeptides bearing a N-terminal Gly.[34] A proximal Phe, Gly, His, Leu, Met, or Tyr residue was found to further favor degradation. These Gly/N degrons were relatively underrepresented in the human proteome, suggesting that few full-length proteins use this pathway. In contrast these Gly/N degrons were enriched in the N termini produced by caspase cleavages. Furthermore, N myristoylation of N-terminal Gly was shown to compete with E3 ligase activity so that proteins or protein fragments bearing N-terminal Gly which do not readily undergo myristoylation and subsequent membrane localization are likely to be destroyed. When considered in context with previous data indicating that N degrons are conserved in toxic caspase-generated fragments, this data points to intricate crosstalk between caspase activity and N-degron pathways.[35]

Pathogens also employ proteolysis to hijack degradation pathways and these systems often provide key insight into mechanism. For example, it was recently demonstrated that Anthrax Lethal Factor (LF), a secreted metalloprotease from *Bacillus anthracis,* relies on the N-end rule E3 ubiquitin ligases: UBR2 and UBR4. It was known that cell killing by LF depended on the activation of the NLRP1B inflammasome.[36] More recently, it was shown that, in the cytosol, LF cleaves NLRPB1 between Lys44 and Leu45, creating a type-II Arg/N degron.[37,38] NLRP1B normally exists as an autoinhibited heterodimer. However, the unmasking of Leu45 by LF destabilizes the autoinhibitory fragment, leading to the oligomerization of the NLRP1B CARD (CAspase-Recruitment Domain) domain-containing fragment, the activation of caspase-1, and pyroptosis. In essence, LF reprograms the stability of a host NLRP1B protein via proteolysis to elicit a cellular response that favors cell death and pathogenesis. Given that many pathogens have evolved secreted proteolytic effectors, it seems probable that many of these could reshape the stability of host proteins through N-degron pathways.

III. Caspase regulation by IAP N-end-binding E3s

The regulation of protease activity is often redundant and multi-faceted. Many proteases are produced in a zymogen form to restrict their activity.[39] Thus, these proteases must mature into an active enzyme prior to efficient cleavage of their biological substrates. This maturation usually requires proteolysis, such that the active enzyme contains N and/or C termini not present in the zymogen. The caspase family of cysteine proteases are well understood in this respect. The inflammatory caspases (−1, −4, and −5 in humans) and initiator caspases (−2, −8, −9, and −10), must dimerize in order to mature.[40] Dimerization is mediated by signaling events that cue the assembly of multicomponent complexes (i.e. inflammasome, apoptosome, DISC complex, etc.) which facilitate dimerization. Upon dimerization, autoproteolysis leads to the release of the mature enzyme bearing new N and C termini on both the large and small subunits. While dimerization and activation of the inflammatory and initiator caspases requires stimulation, the executioner caspases (−3, −6, and −7) exist as constitutive homodimers in their zymogen form. For those caspases, maturation is canonically induced by the upstream activity of the initiator or inflammatory caspases.[41]

The maturation of caspases-3, −7, and −9 yields a new N terminus on the small subunit that engages in a regulatory interaction with the X-linked inhibitor of apoptosis (XIAP). XIAP consists of three Baculovirus IAP Repeat (BIR) domains and a RING domain with E3 ubiquitin ligase activity. BIR1 is devoid of caspase binding activity and seems to support dimerization.[42] In contrast, BIR2 binds and inhibits the active forms of caspases-3 and −7. This activity further involves the peptide linker region between BIR1 and BIR2, which acts as a substrate mimic to bind the caspase active site.[43,44] The structure of the BIR2 domain bound to the N terminus of the small caspase subunit (Figure 4, panel A) has shown that Asp214 and Glu219 in the BIR2 domain are critical to recognition. Further, structural and mutagenesis studies have confirmed that both the end-binding activity and substrate mimicry are important for BIR2 inhibition of caspase-3 and caspase-7.[45] Finally, the BIR3 domain inhibits caspase-9 in a similar fashion, with the distinction that binding of the caspase N-terminus is sufficient to inhibit (Figure 4, panel B), without a role for the linker.[46] It should be noted that the role of XIAP's ubiquitin-ligase activity in regulating caspase degradation remains unclear, but it is clear that end-binding is functionally important.

N-terminal binding of caspases to XIAP needs to be neutralized during apoptosis. This activity is achieved by the generation of nascent N termini that compete with caspases for the N-end binding sites in XIAP. For example, pro-apoptotic proteins released from the mitochondria, including Smac, HTRA2, and PGAM5, contain IAP Binding Motifs (IBMs) at their nascent N termini.[47] In a healthy mitochondrion, Smac is tethered to the inter-membrane by an N-terminal transmembrane anchor. Upon outer-membrane permeabilization, it is cleaved by the integral-membrane serine protease PARL, revealing an N-terminal IBM capable of binding BIR2 and BIR3 in XIAP (Figure 4, panel C and D).[48–50] Through this competition between N termini and XIAP, cell survival is carefully tuned. Knowledge of this pathway has also facilitated creation of peptidomimetics with activity in cells and mice. Specifically, these compounds act at distinct BIR domains to potentiate cell death in cancer cells.[51,52] The competition between caspases and cleaved Smac for XIAP

effectively links the two signaling modules and creates a higher-order signaling network (Figure 4, panel E).

IV. C ends: new opportunities in protease signaling

Until recently, knowledge of end-binding E3 ubiquitin ligases was largely restricted to the N terminus. In 2018 Koren *et al* and Lin *et al* used global protein-stability (GPS) profiling and CRISPR-based functional genomics screens to systematically identify C degrons and match them to the components of the UPS that mediated their degradation.[53–55] This global profiling effort led to the identification of eleven C-end binding E3 ubiquitin ligases with unique specificities. The bulk of these E3 ligases were in the Cullin-RING E3 ligase family (CRLs e.g. Cul2 or Cul4). These ligases do not bind the C terminus directly but rather they use adaptors. Thus far, these adaptors all seem to contain repeats of protein-binding domains, such as Kelch, Ankyrin, Tetratricopeptide, WD40, and armadillo-like repeats. Structures of the Kelch domain adaptor, KLHDC2, bound to C-degron sequences were solved by Rusnac *et al*, showing that binding involves a Gly-Gly sequence at the C terminus (Figure 5, panel A and B).[56] Interestingly, one such Gly-Gly/C degron in USP1 arises from the same autoproteolytic cleavage event that also generates a tertiary Type-1 N degron (via an N-terminal Asn residue). This coincidence of N- and C-degron pathways on a single cleavage site highlights the amount of information that can be unmasked by a single proteolytic event (Figure 5, panel C).

The regulatory logic of C-degron pathways is likely quite different from N-degron pathways. One reason is that free C termini are far more abundant than free N termini in the cell. While C termini are sometimes modified by methylation, amidation and adenylation, most seem to be free.[57] Additionally, it seems likely that C degrons might be more likely than N degrons to co-evolve with proteases because soluble, mammalian protease specificity is often dominated by non-prime residues, particularly in serine and cysteine proteases, where C-end binding would arise. However, this idea has not been rigorously tested and further work is required to understand the molecular and cellular determinants of C-degron pathway function.

V. Crosstalk between CHIP and caspases

N and C degrons link proteolysis with proteostasis. In that context, another recently described C degron is particularly interesting. The carboxy terminus of Hsp70 interacting protein (CHIP) is a dimeric, Ubox domain-containing E3 ubiquitin ligase.[58,59] CHIP also contains a tetratricopeptide repeat (TPR) domain, which is best known for binding to the C terminus of cytosolic heat shock protein 70 (Hsp70) and heat shock protein 90 (Hsp90). Hsp70 and Hsp90 are molecular chaperones that facilitate protein folding, trafficking and function. These chaperones end in similar C-terminal sequences: Hsp70s (IEEVD-CO_2H) and Hsp90s (MEEVD-CO_2H). Binding of these sequences to CHIP facilitates transfer of ubiquitin from E2 to the chaperone-bound client (Figure 6, panel A and B), likely serving to remove misfolded proteins and maintain proteostasis. At a molecular level, CHIP binds C-terminal Asp residues in Hsp70/Hsp90 through a conserved carboxylate clamp in its TPR domain.[60–62]

It was previously assumed that the majority of CHIP functions are a result of its cooperation with Hsp70 and Hsp90 chaperones. However, we recently reported that the CHIP TPR domain interacts with a far broader range of C termini than previously thought.[63] Briefly we used a positional scanning synthetic combinatorial library (PSSCL) of peptides to assign the subsite specificity of the CHIP TPR domain at the four positions (P2, P3, P4, P5) preceding the C-terminal Asp (P1). This approach defined the specificity of the CHIP TPR domain and led to development of a peptide optimized at all five residues (CHIPOpt, Ac-LWWPD). Indeed, CHIPOpt exhibited superior binding properties to the previously best-known sequences, such as Ac-IEEVD (Figure 6, panel C). Binding information from the PSSCL library also created the opportunity to predict other interactors of CHIP in the human proteome using a sequence-based approach. While we predicted that few C termini in full-length proteins would bind well to CHIP, we found ~2,700 putative CHIP binding sequences that were likely to arise from caspase activity. Two of these putative interactions, with tau and caspase-6, were verified in follow-up binding studies. Tau cleaved at Asp 421 (tau^{D421}) and caspase-6 cleaved at Asp179 (caspase-6^{D179}), are both known to accumulate in Alzheimer's disease, and we found that both bound directly to CHIP's TPR domain. Moreover, the direct binding to CHIP could rapidly ubiquitinate tau^{D421} in the absence of a chaperone. This finding greatly expands the potential, chaperone-independent roles for CHIP in regulating proteostasis. Interestingly, binding of CHIP to caspase-6^{D179} did not lead to ubiquitination. Rather, this interaction inhibited caspase-6^{D179} enzyme activity, reminiscent of how XIAP inhibits caspases-3, −7 and −9. For Alzheimer's Disease these observations suggest a key role for this axis in regulating tau accumulation. More broadly, this C-degron pathway suggests a complex interplay among CHIP, caspases, and caspase substrates (Figure 6, panel D).

It is interesting to consider the co-evolution of caspases and CHIP in the context of multicellularity. The TPR family of co-chaperones and the corresponding C-terminal interaction motif in Hsp70s and Hsp90s are absent in prokaryotes but present in yeast. However, there is no yeast ortholog of CHIP. Yeast are also devoid of caspases or similar proteases with specificity for a P1 Asp. It is enticing to consider that CHIP and caspases evolved together in higher organisms to facilitate the transfer of information between caspases and proteostasis networks.

VI. Implications of end-binding in the proteome

It has become clear that new protein termini are a hotspot for PPIs and that end-binding E3 ubiquitin ligases are intimately intertwined with protease signaling. The well characterized proteases involved in this process to date include MetAPs and caspases, but further work is essential to understanding which other endo- and exo- proteases might be involved. These efforts will likely need to integrate modern proteomic methods for detecting protein termini in cells, as well as activity-based protein profiling (ABPP) methods.[64–69] Given the relatively recent description of C-degron pathways, this might be a particularly fruitful area of focus for future work because it seems likely that many other proteases will be involved. At the same time, continued use of modern functional genomics (*e.g.* CRISPRi screens) is poised to uncover which of the hundreds of E3 ubiquitin ligases and adapter proteins might be involved.

We also suggest that the chemical biology community could additionally contribute to this field through the development of high-quality probes for individual PPIs. Success in creating potent, selective Smac mimetics provides motivation and a possible framework for this idea. The interactions of N- or C-terminal peptides with TREs typically involve a relatively small surface area, which is favorable for creating competitive inhibitors (Ran and Gestwicki, 2018 ref 4). However, to gain selectivity, any resulting inhibitors will likely need to mimic the free amine or carboxylate. This constraint may require innovations in isosteres and/or prodrugs to effectively mimic these moieties while maintaining favorable physiochemical properties for cellular activity. However, this goal seems worthwhile because a broader collection of N- and C-degron pathway inhibitors are expected to be useful in both understanding biology and creating translational opportunities.

Funding Statement

This work was supported by grants from the NIH no. R01059690 (to J.E.G.),nos. P41CA196276 and P50GM082250 (to C.S.C.). Additional support included an ARCS Foundation fellowship (to M.R.) and funding from the Sandler Foundation.

References

(1). Nooren IMA; Thornton JM Structural Characterization and Functional Significance of Transient Protein-Protein Interactions. J. Mol. Biol 2003, 325, 991–1018. [PubMed: 12527304]

(2). Perkins JR; Diboun I; Dessailly BH; Lees JG; Orengo C Transient Protein-Protein Interactions: Structural, Functional, and Network Properties. Structure 2010, 18, 1233–1243. [PubMed: 20947012]

(3). Scott DE; Bayly AR; Abell C; Skidmore J Small Molecules, Big Targets: Drug Discovery Faces the Protein-Protein Interaction Challenge. Nat. Rev. Drug Discov 2016, 15, 533–550. [PubMed: 27050677]

(4). Ran X; Gestwicki JE Inhibitors of Protein–Protein Interactions (PPIs): An Analysis of Scaffold Choices and Buried Surface Area. Curr. Opin. Chem. Biol 2018, 44, 75–86. [PubMed: 29908451]

(5). Deribe YL; Pawson T; Dikic I Post-Translational Modifications in Signal Integration. Nat. Struct. Mol. Biol 2010, 17, 666–672. [PubMed: 20495563]

(6). Torres IO; Fujimori DG Functional Coupling between Writers, Erasers and Readers of Histone and DNA Methylation. Curr. Opin. Struct. Biol 2015, 35, 68–75. [PubMed: 26496625]

(7). Hyun K; Jeon J; Park K; Kim J Writing, Erasing and Reading Histone Lysine Methylations. Exp. Mol. Med 2017, 49, e324. 10.1038/emm.2017.11. [PubMed: 28450737]

(8). Klein T; Eckhard U; Dufour A; Solis N; Overall CM Proteolytic Cleavage - Mechanisms, Function, and “Omic” Approaches for a Near-Ubiquitous Posttranslational Modification. Chem. Rev 2018, 118, 1137–1168. [PubMed: 29265812]

(9). Thornberry NA; Lazebnik. Caspases: Enemies Within. Science 1998, 281, 1312–1317. [PubMed: 9721091]

(10). Overall CM; Kleifeld O Validating Matrix Metalloproteinases as Drug Targets and Anti-Targets for Cancer Therapy. Nat. Rev. Cancer 2006, 6, 227–239. [PubMed: 16498445]

(11). Smith DM; Benaroudj N; Goldberg A Proteasomes and Their Associated ATPases: A Destructive Combination. J. Struct. Biol 2006, 156, 72–83. [PubMed: 16919475]

(12). Nakayama KI; Nakayama K Ubiquitin Ligases: Cell-Cycle Control and Cancer. Nat. Rev. Cancer 2006, 6, 369–381. [PubMed: 16633365]

(13). Varshavsky A N-Degron and C-Degron Pathways of Protein Degradation. Proc. Natl. Acad. Sci 2019, 116, 358–366. [PubMed: 30622213]

(14). Salvesen GS; Duckett CS IAP Proteins: Blocking the Road to Death’s Door. Nat. Rev. Mol. Cell Biol 2002, 3, 401–410. [PubMed: 12042762]

(15). Turk B; Turk D; Turk V Protease Signaling: The Cutting Edge. EMBO J. 2012, 31, 1630–1643. [PubMed: 22367392]

(16). Varshavsky A The N-End Rule Pathway and Regulation by Proteolysis. Protein Sci. 2011, 20, 1298–1345. [PubMed: 21633985]

(17). Bachmair A; Finley D; Varshavsky A In Vivo Half-Life of a Protein Is a Function of Its Amino-Terminal Residue. Science 1986, 234, 179–186. [PubMed: 3018930]

(18). Dougan DA; Micevski D; Truscott KN The N-End Rule Pathway: From Recognition by N-Recognins, to Destruction by AAA+proteases. Biochim. Biophys. Acta - Mol. Cell Res 2012, 1823, 83–91.

(19). Mischerikow N; Heck AJR Targeted Large-Scale Analysis of Protein Acetylation. Proteomics 2011, 11, 571–589. [PubMed: 21246731]

(20). Arnesen T; Damme P Van; Polevoda B; Helsens K; Evjenth R; Colaert N; Erik J Proteomics Analyses Reveal the Evolutionary Conservation and Divergence of N-Terminal Acetyltransferases from Yeast and Humans. Proc. Natl. Acad. Sci 2009, 106, 8157–8162. [PubMed: 19420222]

(21). Xiao Q; Zhang F; Nacev BA; Liu JO; Pei D Protein N-Terminal Processing: Substrate Specificity of Escherichia Coli and Human Methionine Aminopeptidases. Biochemistry 2010, 49, 5588–5599. [PubMed: 20521764]

(22). Frottin F; Martinez A; Peynot P; Mitra S; Holz RC; Giglione C; Meinnel T The Proteomics of N-Terminal Methionine Cleavage. Mol. Cell. Proteomics 2006, 5, 2336–2349. [PubMed: 16963780]

(23). Aksnes H; Drazic A; Marie M; Arnesen T First Things First: Vital Protein Marks by N-Terminal Acetyltransferases. Trends Biochem. Sci 2016, 41, 746–760. [PubMed: 27498224]

(24). N-degrons R; Tasaki T; Mulder LCF; Iwamatsu A; Lee MJ; Davydov IV; Varshavsky A; Muesing M; Kwon YT A Family of Mammalian E3 Ubiquitin Ligases That Contain the UBR Box Motif and Recognize N-Degrons. Mol. Cell. Biol 2005, 25, 7120–7136. [PubMed: 16055722]

(25). Choi WS; Jeong B-C; Joo YJ; Lee M-R; Kim J; Eck MJ; Song HK Structural Basis for the Recognition of N-End Rule Substrates by the UBR Box of Ubiquitin Ligases. Nat. Struct. Mol. Biol 2010, 17, 1175–1181. [PubMed: 20835240]

(26). Tasaki T; Zakrzewska A; Dudgeon DD; Jiang Y; Lazo JS; Kwon YT The Substrate Recognition Domains of the N-End Rule Pathway. J. Biol. Chem 2009, 284, 1884–1895. [PubMed: 19008229]

(27). Román-Hernández G; Hou JY; Grant RA; Sauer RT; Baker TA The ClpS Adaptor Mediates Staged Delivery of N-End Rule Substrates to the AAA+ ClpAP Protease. Mol. Cell 2011, 43, 217–228. [PubMed: 21777811]

(28). Hu RG; Brower CS; Wang H; Davydov IV; Sheng J; Zhou J; Yong TK; Varshavsky A Arginyltransferase, Its Specificity, Putative Substrates, Bidirectional Promoter, and Splicing-Derived Isoforms. J. Biol. Chem 2006, 281, 32559–32573. [PubMed: 16943202]

(29). Hu RG; Sheng J; Qi X; Xu Z; Takahashi TT; Varshavsky A The N-End Rule Pathway as a Nitric Oxide Sensor Controlling the Levels of Multiple Regulators. Nature 2005, 437, 981–986. [PubMed: 16222293]

(30). Wang H; Piatkov KI; Brower CS; Varshavsky A Glutamine-Specific N-Terminal Amidase, a Component of the N-End Rule Pathway. Mol. Cell 2009, 34, 686–695. [PubMed: 19560421]

(31). Piatkov KI; Colnaghi L; Békés M; Varshavsky A; Huang TT The Auto-Generated Fragment of the Usp1 Deubiquitylase Is a Physiological Substrate of the N-End Rule Pathway. Mol. Cell 2012, 48, 926–933. [PubMed: 23159736]

(32). Chen S-J; Wu X; Wadas B; Oh J-H; Varshavsky A An N-End Rule Pathway That Recognizes Proline and Destroys Gluconeogenic Enzymes. Science 2017, 355, eaal3655. 10.1126/science.aal3655. [PubMed: 28126757]

(33). Dong C; Zhang H; Li L; Tempel W; Loppnau P; Min J Molecular Basis of GID4-Mediated Recognition of Degrons for the Pro/N-End Rule Pathway Article. Nat. Chem. Biol 2018, 14, 466–473. [PubMed: 29632410]

(34). Timms RT; Zhang Z; Rhee DY; Harper JW; Koren I; Elledge SJ A Glycine-Specific N-Degron Pathway Mediates the Quality Control of Protein N-Myristoylation. Science 2019, 364 (6448).

(35). Piatkov KI; Brower CS; Varshavsky A The N-End Rule Pathway Counteracts Cell Death by Destroying Proapoptotic Protein Fragments. Proc. Natl. Acad. Sci 2012, 109, E1839–E1847. 10.1073/pnas.1207786109. [PubMed: 22670058]

(36). Chavarría-Smith J; Vance RE Direct Proteolytic Cleavage of NLRP1B Is Necessary and Sufficient for Inflammasome Activation by Anthrax Lethal Factor. PLoS Pathog. 2013, 9, 1–4.

(37). Chui AJ; Okondo MC; Rao SD; Gai K; Griswold AR; Johnson DC; Ball DP; Taabazuing CY; Orth EL; Vittimberga BA; Bachovchin DA N-Terminal Degradation Activates the NLRP1B Inflammasome. Science 2019, 364, 82–85. [PubMed: 30872531]

(38). Sandstrom A; Mitchell PS; Goers L; Mu EW; Lesser CF; Vance RE Functional Degradation: A Mechanism of NLRP1 Inflammasome Activation by Diverse Pathogen Enzymes. Science 2019, 364, eaau1330. 10.1126/science.aau1330. [PubMed: 30872533]

(39). Khan AR; James MNG Molecular Mechanisms for the Conversion of Zymogens to Active Proteolytic Enzymes. Protein Sci. 2008, 7, 815–836.

(40). Stennicke HR; Salvesen GS Caspases - Controlling Intracellular Signals by Protease Zymogen Activation. Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol 2000, 1477, 299–306.

(41). Fuentes-Prior P; Salvesen GS The Protein Structures That Shape Caspase Activity, Specificity, Activation and Inhibition. Biochem. J 2004, 384, 201–232. [PubMed: 15450003]

(42). Lin SC; Huang Y; Lo YC; Lu M; Wu H Crystal Structure of the BIR1 Domain of XIAP in Two Crystal Forms. J. Mol. Biol 2007, 372, 847–854. [PubMed: 17698078]

(43). Riedl SJ; Renatus M; Schwarzenbacher R; Zhou Q; Sun C; Fesik SW; Liddington RC; Salvesen GS Structural Basis for the Inhibition of Caspase-3 by XIAP. Cell 2001, 104, 791–800. [PubMed: 11257232]

(44). Huang Y Structural Basis of Caspase Inhibition by XIAP Differential Roles of the Linker versus the BIR Domain. Cell 2004, 104, 781–790.

(45). Scott FL; Denault JB; Riedl SJ; Shin H; Renatus M; Salvesen GS XIAP Inhibits Caspase-3 and −7 Using Two Binding Sites: Evolutionary Conserved Mechanism of IAPs. EMBO J. 2005, 24, 645–655. [PubMed: 15650747]

(46). Shiozaki EN; Chai J; Rigotti DJ; Riedl SJ; Li P; Srinivasula SM; Alnemri ES; Fairman R; Shi Y Mechanism of XIAP-Mediated Inhibition of Caspase-9. Mol. Cell 2003, 11, 519–527. [PubMed: 12620238]

(47). Shi Y A Conserved Tetrapeptide Motif: Potentiating Apoptosis through IAP-Binding. Cell Death Differ. 2002, 9, 93–95. [PubMed: 11840157]

(48). Saita S; Nolte H; Fiedler KU; Kashkar H; Saskia AV; Zahedi RP; Krüger M; Langer T PARL Mediates Smac Proteolytic Maturation in Mitochondria to Promote Apoptosis. Nat. Cell Biol 2017, 19, 318–328. [PubMed: 28288130]

(49). Wu G; Chai J; Suber TL; Wu J-W; Du C; Wang X; Shi Y Structural Basis of IAP Recognition by Smac/DIABLO. Nature 2000, 408, 1008–1012. [PubMed: 11140638]

(50). Liu Z; Sun C; Olejniczak ET; Meadows RP; Betz SF; Oost T; Herrmann J; Wu JC; Fesik SW Structural Basis for Binding of Smac / DIABLO to the XIAP BIR3 Domain. Nature 2000, 408, 1004–1008. [PubMed: 11140637]

(51). Fulda S; Vucic D Targeting IAP Proteins for Therapeutic Intervention in Cancer. Nature Reviews Drug Discovery 2012, 11, 109–124. [PubMed: 22293567]

(52). Sun H; Lu J; Liu L; Yang CY; Wang S Potent and Selective Small-Molecule Inhibitors of cIAP1/2 Proteins Reveal That the Binding of Smac Mimetics to XIAP BIR3 Is Not Required for Their Effective Induction of Cell Death in Tumor Cells. ACS Chem. Biol 2014, 9, 994–1002. [PubMed: 24521431]

(53). Koren I; Timms RT; Kula T; Xu Q; Li MZ; Elledge SJ The Eukaryotic Proteome Is Shaped by E3 Ubiquitin Ligases Targeting C-Terminal Degrons. Cell 2018, 173, 1622–1635. [PubMed: 29779948]

(54). Chatr-Aryamontri A; van der Sloot A; Tyers M At Long Last, a C-Terminal Bookend for the Ubiquitin Code. Mol. Cell 2018, 70, 568–571 [PubMed: 29775575]

(55). Lin HC; Yeh CW; Chen YF; Lee TT; Hsieh PY; Rusnac DV; Lin SY; Elledge SJ; Zheng N; Yen HCS C-Terminal End-Directed Protein Elimination by CRL2 Ubiquitin Ligases. Mol. Cell 2018, 70, 602–613. [PubMed: 29775578]

(56). Rusnac DV; Lin HC; Canzani D; Tien KX; Hinds TR; Tsue AF; Bush MF; Yen HCS; Zheng N Recognition of the Diglycine C-End Degron by CRL2KLHDC2 Ubiquitin Ligase. Mol. Cell 2018, 72, 813–822. [PubMed: 30526872]

(57). Marino G; Eckhard U; Overall CM Protein Termini and Their Modifications Revealed by Positional Proteomics. ACS Chem. Biol 2015, 10, 1754–1764. [PubMed: 26042555]

(58). Ballinger CA; Connell P; Wu Y; Hu Z; Thompson LJ; Yin LY; Patterson C Identification of CHIP, a Novel Tetratricopeptide Repeat-Containing Protein That Interacts with Heat Shock Proteins and Negatively Regulates Chaperone Functions. Mol. Cell. Biol 1999, 19, 4535–4545. [PubMed: 10330192]

(59). Jiang J; Ballinger CA; Wu Y; Dai Q; Cyr DM; Höhfeld J; Patterson C CHIP Is a U-Box-Dependent E3 Ubiquitin Ligase. J. Biol. Chem 2001, 276, 42938–42944. [PubMed: 11557750]

(60). Zhang M; Windheim M; Roe SM; Peggie M; Cohen P; Prodromou C; Pearl LH Chaperoned Ubiquitylation - Crystal Structures of the CHIP U Box E3 Ubiquitin Ligase and a CHIP-Ubc13-Uev1a Complex. Mol. Cell 2005, 20, 653–659. [PubMed: 16337587]

(61). Qian SB; McDonough H; Boellmann F; Cyr DM; Patterson C CHIP-Mediated Stress Recovery by Sequential Ubiquitination of Substrates and Hsp70. Nature 2006, 440, 551–555. [PubMed: 16554822]

(62). Assimon VA; Southworth DR; Gestwicki JE Specific Binding of Tetratricopeptide Repeat Proteins to Heat Shock Protein 70 (Hsp70) and Heat Shock Protein 90 (Hsp90) Is Regulated by Affinity and Phosphorylation. Biochemistry 2015, 54, 7120–7131. [PubMed: 26565746]

(63). Ravalin M; Theofilas P; Basu K; Opoku-Nsiah KA; Assimon VA; Medina-Cleghorn D; Chen Y-F; Bohn MF; Arkin M; Grinberg LT; Craik CS; Gestwicki JE Specificity for Latent C Termini Links the E3 Ubiquitin Ligase CHIP to Caspases. Nat. Chem. Biol 2019, 15, 786–794. [PubMed: 31320752]

(64). Weeks AM; Wells JA Engineering Peptide Ligase Specificity by Proteomic Identification of Ligation Sites. Nat. Chem. Biol 2018, 14, 50–57. [PubMed: 29155430]

(65). Julien O; Zhuang M; Wiita AP; O'Donoghue AJ; Knudsen GM; Craik CS; Wells JA Quantitative MS-Based Enzymology of Caspases Reveals Distinct Protein Substrate Specificities, Hierarchies, and Cellular Roles. Proc. Natl. Acad. Sci. U. S. A 2016, 113, E2001–E2010. 10.1073/pnas.1524900113. [PubMed: 27006500]

(66). Donoghue AJO; Eroy-reveles AA; Knudsen GM; Ingram J; Zhou M; Statnekov JB; Greninger AL; Hostetter DR; Qu G; Maltby DA; Anderson MO; DeRisi JL; McKerrow JH; Burlingame AL; Craik CS. Global Identification of Peptidase Specificity by Multiplex Substrate Profiling. Nat. Methods 2012, 9, 1095–1100. [PubMed: 23023596]

(67). Griswold AR; Cifani P; Rao SD; Axelrod AJ; Miele MM; Hendrickson RC; Kentsis A; Bachovchin DA A Chemical Strategy for Protease Substrate Profiling. Cell Chem. Biol 2019, 26, 901–907. [PubMed: 31006619]

(68). Fomby P; Cherlin AJ Activity Based Probes as a Tool for Functional Proteomic Analysis of Proteases. Expert. Rev. Proteomics 2011, 72, 181–204.

(69). Kleifeld O; Doucet A; Auf Dem Keller U; Prudova A; Schilling O; Kainthan RK; Starr AE; Foster LJ; Kizhakkedathu JN; Overall CM Isotopic Labeling of Terminal Amines in Complex Samples Identifies Protein N-Termini and Protease Cleavage Products. Nat. Biotechnol 2010, 28, 281–288. [PubMed: 20208520]

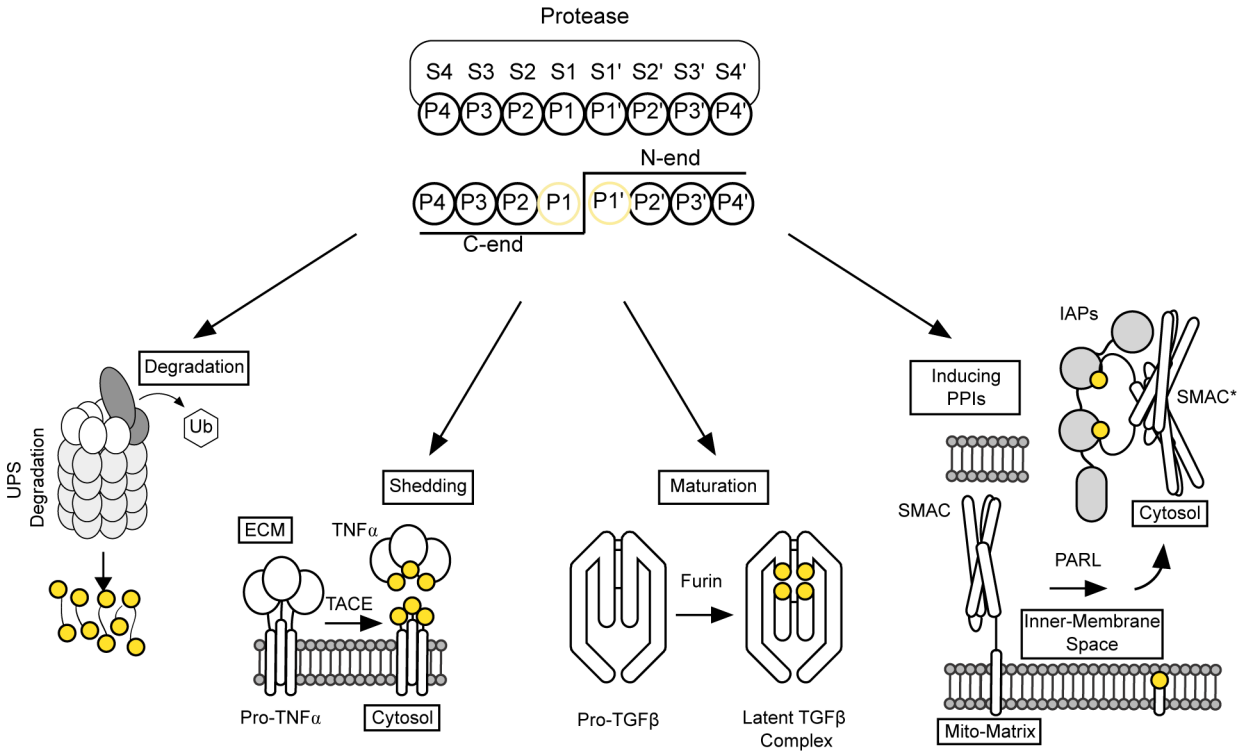

Figure 1.

Protease activity catalyzes the hydrolysis of a polypeptide. The structure of the protease active site forms subsites (S1-S4, S1'-S4') with specificity for amino acids N terminal to the cleaved bond (P1-P4, non-prime) and C terminal to the cleaved bond (P1'-P4', prime). The non-prime residues generate a new C terminus (C-end) and the prime side residues generate a new N terminus (N-end). This specificity is coupled to biological function through a variety of signaling mechanisms including degradation (*e.g.* the ubiquitin proteasome system or UPS), domain liberation (*e.g.* shedding of TNFα by TACE), inducing maturation (*e.g.* activation of TGFβ by Furin), and modulating PPIs (*e.g.* sequestration of IAP's by Smac via PARL). The key challenge in understanding protease signaling is identifying the determinants of protease specificity and connecting them to a signaling mechanism.

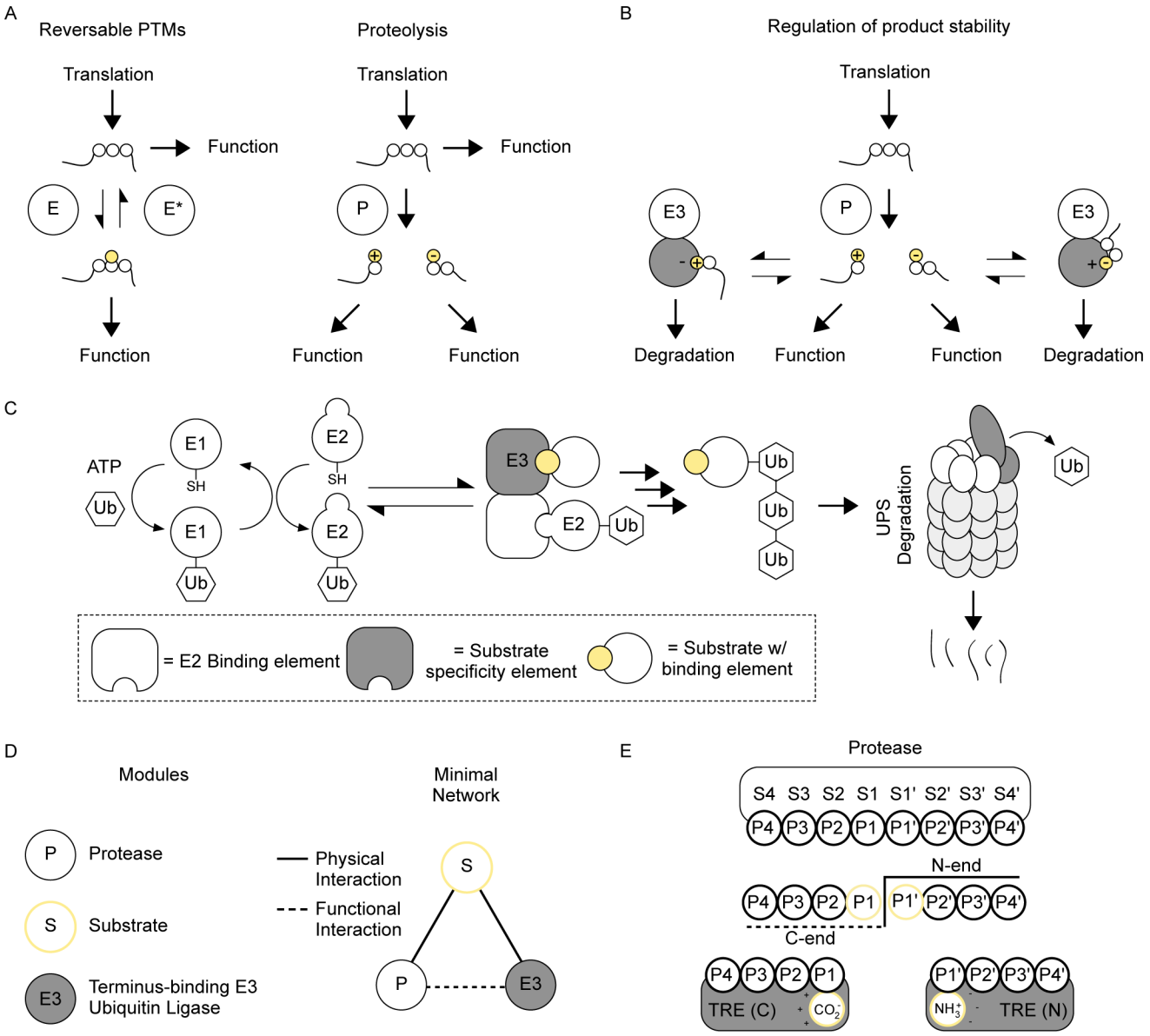

Figure 2.

End-binding E3 ubiquitin ligases regulate protease product stability to enable signaling. (a) The balance of activity between PTM placing (E) and PTM removing (E*) enzymes controls the partitioning of function between modified and unmodified species for most PTMs. The irreversibility of proteolysis requires a different regulatory mechanism. (b) End- binding E3 ubiquitin ligases can degrade the products of proteolysis and control proteolytic signal intensity and duration. (c) E3 ubiquitin ligases provide specificity to the ubiquitin proteasome system by binding to degrons (yellow circles) and scaffolding substrate ubiquitination and subsequent delivery to the proteasome. (d) When a protease (P) cleaves a substrate (S) that is subsequently recognized by and end-binding E3 ubiquitin ligase (E3), a functional interaction is created between protease and E3. These elements form a minimal signaling network. (e) The strength of the functional link between a protease and an end-binding E3 is related to the overlap in the specificity between the protease active

site and the terminal recognition element (TRE) in the E3. This links S1-S4 subsites to C-terminal TREs and S1’-S4’ to N-terminal TREs.

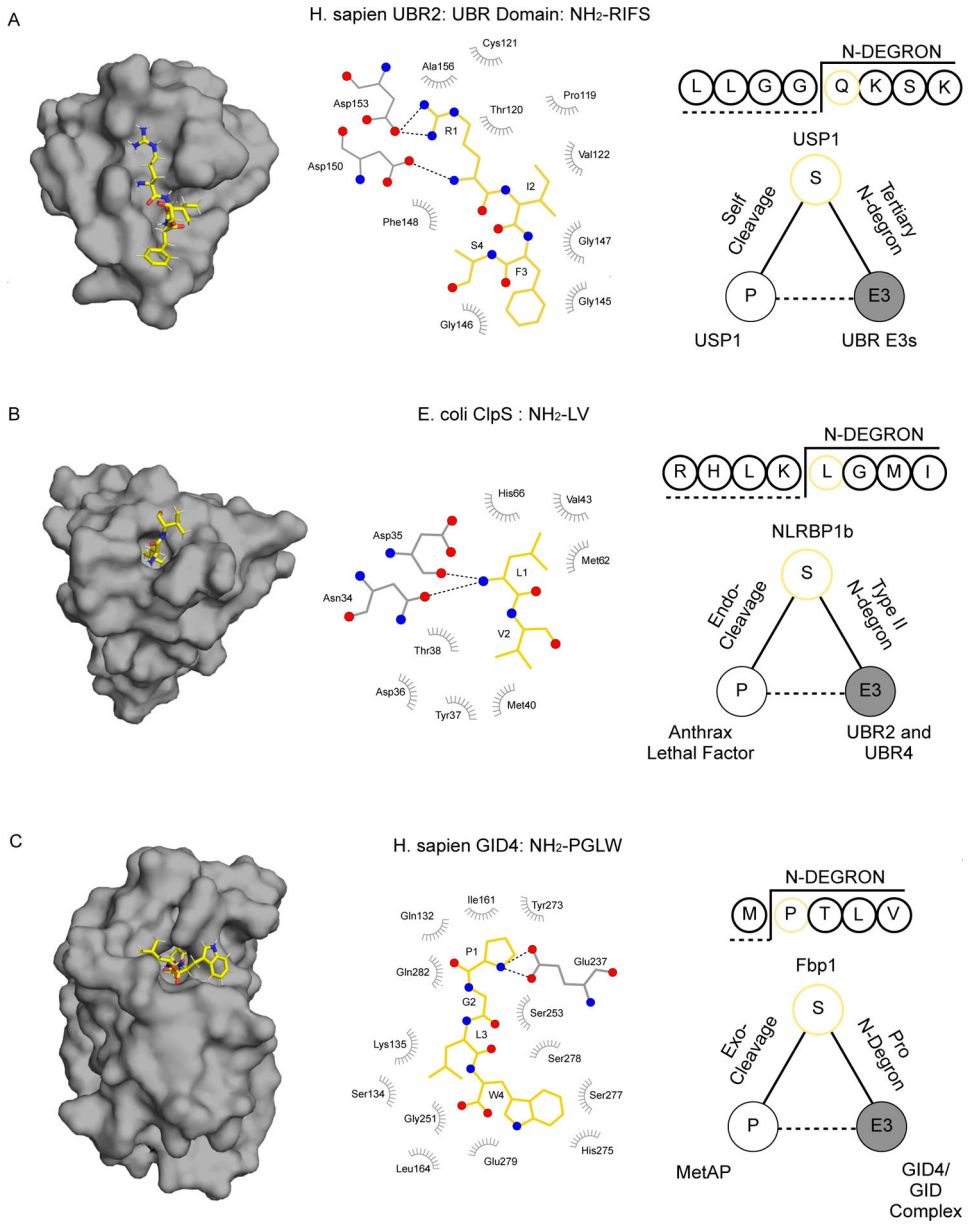

Figure 3.

Specific determinants of N-end binding link protease and E3 function in signaling pathways. (a) The X-ray crystal structure (PDB 3NY3, left panel) of the UBR N-end binding TRE from UBR2 (surface, grey) bound to a tetrapeptide (NH_2-RIFS-CO_2H, sticks, yellow). Specificity for the N-terminal Arg is mediated by Asp150 and Asp 153 (middle panel). UBR binds to the N degron at the nascent N terminus of a self-cleavage product of USP1 after the N-terminal Gln is deamidated and argenylated (right panel). (b) The X-ray crystal structure (PDB 3O2H, left panel) of the *E. coli* N-end binding TRE from ClpS (surface, grey) bound to a peptide (NH_2-LVKSKATNLLYC-CO_2H, sticks, yellow). Specificity for the N-terminal Leu is mediated by coordination of the amine by Asn34, Asp35, and His66 (middle panel). A homologous domain in human UBR 1 and 2 is thought to direct the binding of Type II N degrons. Generation of a Type II N degron by Anthrax LF activity

mediates NLRP1B inflammasome activation and cell death (right panel). (c) The X-ray crystal structure (PDB 6CDC, left panel) of the GID4 N-end binding TRE from GID E3 ubiquitin ligase complex (surface, grey) bound to a tetrapeptide (NH_2-PGLW-CO_2H, sticks, yellow). Specificity for the N-terminal Pro is mediated by coordination of the proline amine by Glu237 and backbone contacts with Gln132, Ser253, Ser278, Gln282 (middle panel). GID4 binds to the N degron at the nascent N terminus of a methionine aminopeptidase cleavage product of gluconeogenic enzymes including Fbp1 to mediate their rapid decay (right panel).

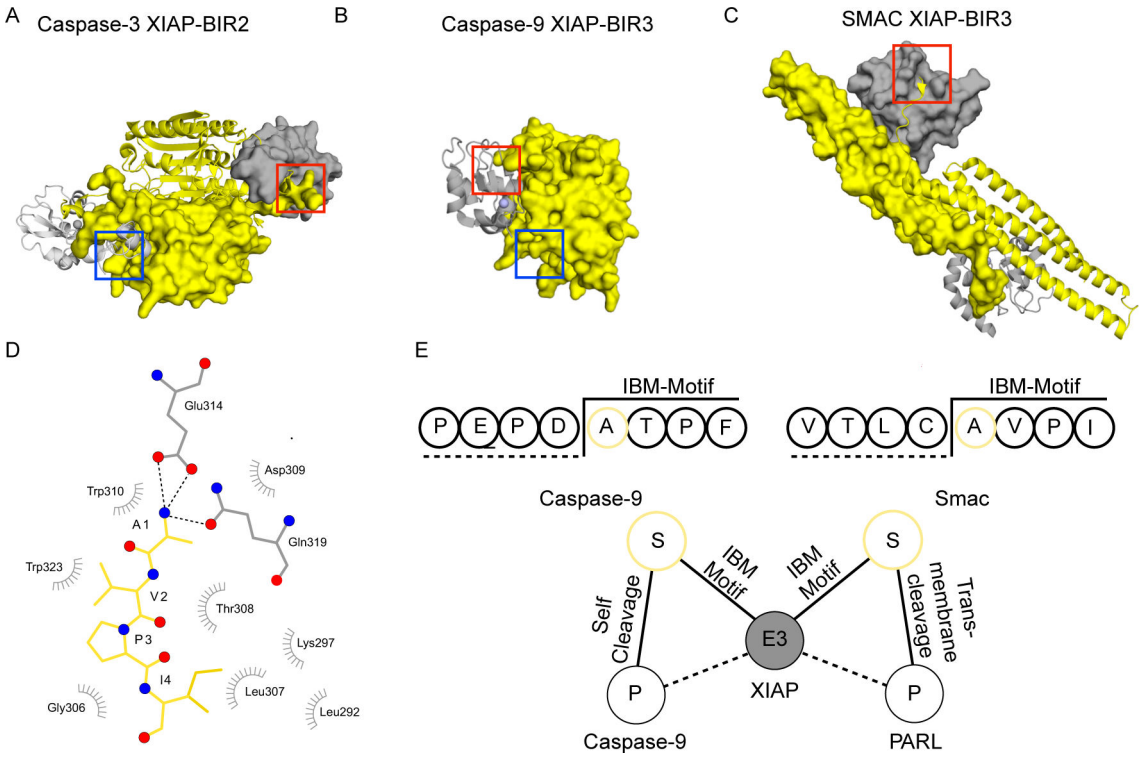

Figure 4.

N-end binding is critical for the regulation of apoptosis. (a) The X-ray crystal structure (PDB 1I3O) of the BIR2 domain of XIAP (grey, surface and cartoon) bound to caspase-3 dimer (yellow, surface and cartoon). BIR2 inhibition of caspase-3 activity depends on its ability to occupy the active site (blue square) and bind the nascent N-terminal IBM motif generated by maturation of the enzyme (red square). (b) The X-ray crystal structure (PDB 1NW9) of the BIR3 domain of XIAP (grey, cartoon) bound to caspase-9 monomer. BIR3 inhibition of caspase-9 depends on its ability to bind the nascent N-terminal IBM motif generated by maturation of the enzyme (red square) but not the active site (blue square). (c) The X-ray crystal structure (PDB 1G73) of Smac dimer (Second mitochondrial activator of caspases, yellow, cartoon and surface) bound to BIR3 domain of XIAP (grey, cartoon and surface). Cleavage of Smac by PARL and release form the mitochondria allows the IBM motif generated by cleavage to sequester XIAP and relieve caspase inhibition. (d) The N-terminal Ala in the IBM motif of Smac (NH_2-AVPI) is specifically coordinated by Glu314 and Gln319. (e) Complementary specificities of caspase-9, PARL, and XIAP establish a network architecture that encodes the signaling function of this system.

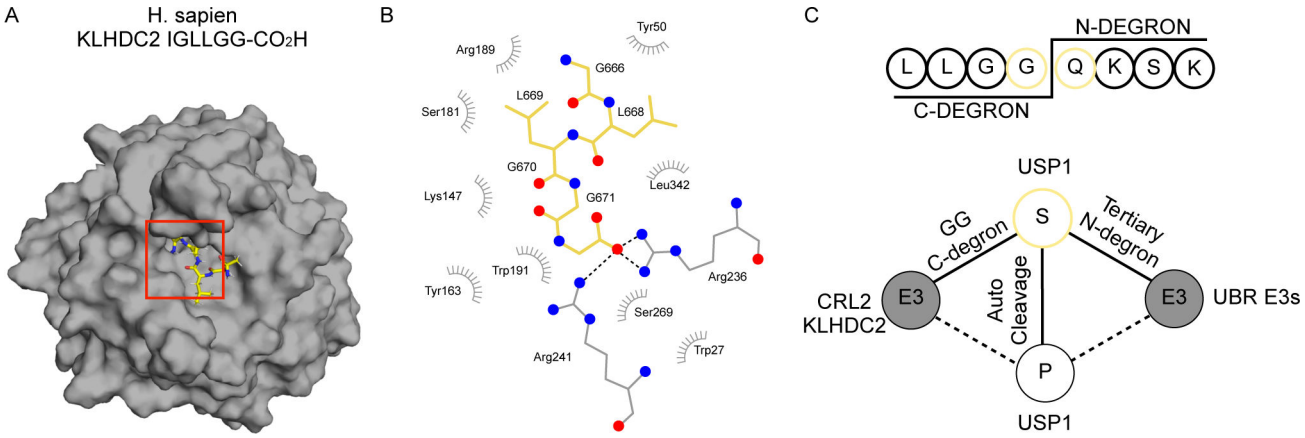

Figure 5.

C-end binding adaptors to Cullin E3 ubiquitin ligases link degradation and non-prime protease specificity. (a) The X-ray crystal structure (PDB 6DO5) the Cul2 adaptor KLHDC2 (grey, surface) bound to a C-terminal peptide derived form a self-cleavage event in USP1 (NH_2-IGLLGG-CO_2H, yellow sticks, red square). (b) The C-terminal Gly is coordinated by Arg236, Arg241, and Ser269. (c) Self-cleavage by USP1 generates coincident N and C degrons with the potential to regulate USP1 function.

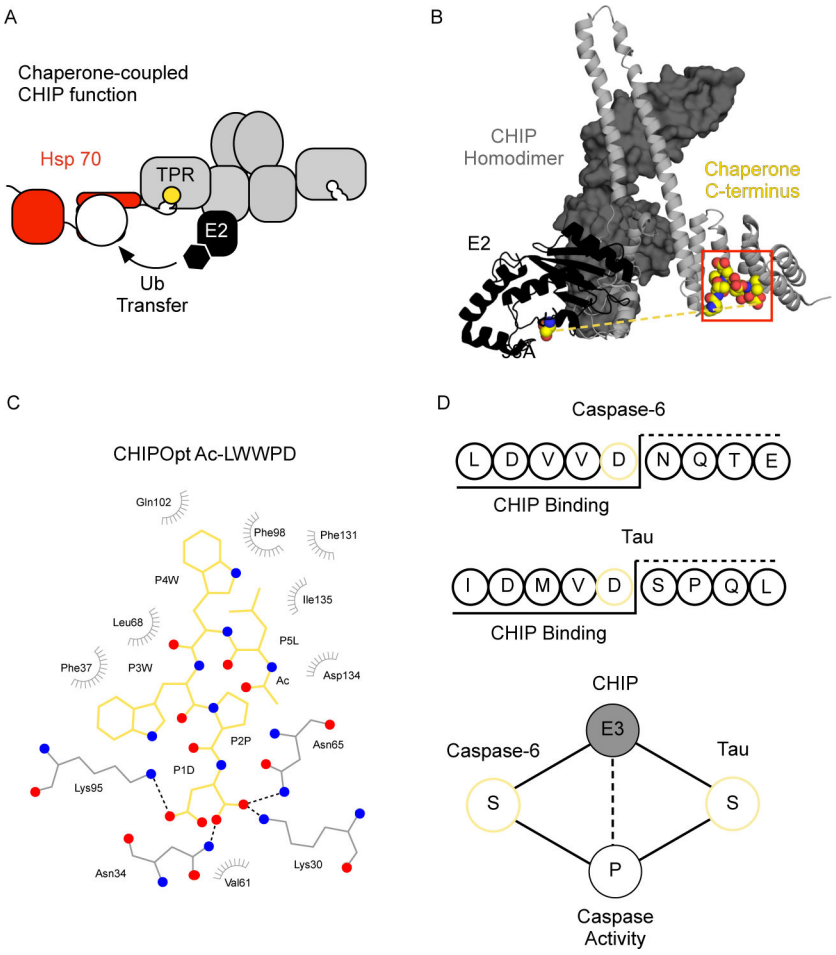

Figure 6.

C-end binding by the E3 ubiquitin ligase CHIP links caspases and protein homeostasis. (a) Canonical CHIP (grey) function stems from an ability to bind to ubiquitin charged E2s (black) and cytosolic chaperones (*i.e.* Hsp70s, red) to mediate ubiquitination of chaperone bound client (white). (b) This function is rooted in the ability of CHIP (PDB 2Cl2, grey, cartoon and surface) to bind a conserved C-terminal motif (IEEVD-CO_2H in Hsp70s and MEEVD-CO_2H in Hsp90s, yellow, spheres) through its TPR domain (red square) and bring it into the proximity of E2 bound to the Ubox domain (PDB 20XQ, black, cartoon). (c) The C-terminal Asp of an optimized peptide bound to the CHIP TPR (PDB 6NSV) is coordinated by Lys30, Asn34, Asn65 and Lys95. (d) Shared specificity with caspases mediates direct binding and inhibition of active capsase-6 (caspase-6^{D179}) and binding and ubiquitination of caspase-cleaved tau (tau^{D421}).