HHS Public Access

Author manuscript

Mol Cell. Author manuscript; available in PMC 2019 May 17.

Published in final edited form as:

Mol Cell. 2018 May 17; 70(4): 602–613.e3. doi:10.1016/j.molcel.2018.04.006.

C-terminal end-directed protein elimination by CRL2 ubiquitin ligases

Hsiu-Chuan Lin[1,2], **Chi-Wei Yehf**[#1], **Yen-Fu Chenf**[#1], **Ting-Ting Lee**[1], **Pei-Yun Hsieh**[1], **Domnita V. Rusnac**[3], **Sung-Ya Lin**[1,2], **Stephen J. Elledge**[4], **Ning Zheng**[3], and **Hsueh-Chi S. Yen**[*,1,2]

[1.]Institute of Molecular Biology, Academia Sinica, Taipei 11529, Taiwan.

[2.]Genome and Systems Biology Degree Program, National Taiwan University and Academia Sinica, Taipei 10617, Taiwan.

[3.]Howard Hughes Medical Institute, Department of Pharmacology, University of Washington, Seattle, WA 98195, USA.

[4.]Howard Hughes Medical Institute, Department of Genetics, Harvard Medical School, Division of Genetics, Brigham and Women's Hospital, Boston, MA 02115, USA.

[#] These authors contributed equally to this work.

SUMMARY

The proteolysis-assisted protein quality control system guards the proteome from potentially detrimental aberrant proteins. How miscellaneous defective proteins are specifically eliminated and which molecular characteristics direct them for removal are fundamental questions. We reveal a mechanism, DesCEND (Destruction *via* the C-END), by which CRL2 ubiquitin ligase uses interchangeable substrate receptors to recognize the unusual C-termini of abnormal proteins, i.e. C-end degrons. C-end degrons are mostly less than ten residues in length and comprise a few indispensable residues along with some rather degenerate ones. The C-terminal end-position is essential for C-end degron function. Truncated selenoproteins generated by translation errors and the USP1 N-terminal fragment from post-translational cleavage are eliminated by DesCEND. DesCEND also targets full-length proteins with naturally-occurring C-end degrons. The C-end degron in DesCEND echoes the N-end degron in the N-end rule pathway, highlighting the dominance of protein "ends" as indicators for protein elimination.

[*]**Corresponding author:** hyen@imb.sinica.edu.tw, Address: Rm. 313, Institute of Molecular Biology, Academia Sinica No. 128, Sec. 2, Academia Rd., Nangang, Taipei 115, Taiwan, TEL: +886-2-2788-3593 (office); +886-2-2789-9212 (lab).

AUTHOR CONTRIBUTIONS

HCY carried out the CRL2 GPS screen, validated the screen, designed experiments, analyzed data, supervised the project and wrote the paper; HCL validated the GPS screen, performed BC-box screens, characterized C-end degrons, performed experiments, analyzed data and wrote the paper; CWY validated the GPS screen; YFC performed bioinformatic analyses; TTL performed BC-box screens; PYH analyzed the sequences of CRL2 substrates; DVR examined the physical interaction between ubiquitin and KLHDC2; SYL characterized the role of APP in CRL2-mediated degradation; SJE supervised the CRL2 GPS screen; NZ designed experiments and supervised the project.

DECLRATION OF INTERESTS

The authors declare no competing interests.

Keywords

Protein quality control; CRL2 ubiquitin ligase; C-end degron

INTRODUCTION

Proteome integrity is of critical importance in almost all cellular processes, yet it is constantly challenged by a diversity of protein aberrations arising from genetic mutations, erroneous transcription and translation, improper folding, faulty targeting, and damage induced by various environmental stresses (Balch et al., 2008; Goldberg, 2003; Wolff et al., 2014). To ensure proteome fidelity, cells have evolved a comprehensive proteolysis-assisted protein quality control (PQC) network to selectively eliminate aberrant proteins (Chen et al., 2011; Goldberg, 2003; Powers and Balch, 2013; Wickner et al., 1999). Accumulation of defective proteins may either dominant-negatively compete with native proteins or form harmful aggregates, which have been implicated in neurodegenerative disorders (Chen et al., 2011; Ross and Poirier, 2004; Shastry, 2003; Skovronsky et al., 2006). Understanding the mechanisms of PQC is therefore of paramount significance in biological science.

Proteolysis-assisted PQC is mainly mediated by the ubiquitin-proteasome system, with substrate selectivity being conducted by ubiquitin ligases (Ciechanover, 1994; Ciechanover et al., 1984; Glickman and Ciechanover, 2002; Hershko and Ciechanover, 1998). However, it remains puzzling how a limited set of ubiquitin ligases can differentiate a whole spectrum of anomalous proteins from their normal counterparts. Abnormal proteins may be detected co- or post-translationally. Known mechanisms include sensing ribosome stalling when translating nascent peptides from problematic mRNAs (Bengtson and Joazeiro, 2010; Brandman et al., 2012), detection of abnormally exposed hydrophobicity of misfolded proteins directly by ubiquitin ligases or indirectly through chaperones (Fang et al., 2014; Murata et al., 2001; Rosenbaum et al., 2011), or recognizing the destabilizing N-terminal residue of proteins by the N-end rule pathway (Bachmair et al., 1986; Shemorry et al., 2013; Sriram et al., 2011; Varshavsky, 2011). However, these strategies cannot capture all types of protein abnormalities.

We previously found that CRL2 ubiquitin ligase specifically eliminates the truncated selenoproteins generated by failed UGA to Sec recoding (Lin et al., 2015). CRL2 distinguishes defective selenoproteins by their “unusual” C-termini (Lin et al., 2015). CRL2 is a modular ubiquitin ligase and belongs to a member of the cullin-RING ligase (CRL) superfamily (Deshaies and Joazeiro, 2009; Petroski and Deshaies, 2005). CRL2 has a Cul2 protein scaffold, a RING domain catalytic subunit, and an Elongin B/C adaptor to recruit around 40 interchangeable sets of BC-box proteins as substrate receptors (Kamura et al., 2004; Mahrour et al., 2008; Petroski and Deshaies, 2005). The most well-characterized BC-box protein is the tumor suppressor von Hippel-Lindau (VHL), which targets hypoxia-inducible factor 1 alpha (HIF1A) for degradation (Ivan et al., 2001; Jaakkola et al., 2001). Compared to other CRL members, the primary function of CRL2 remains elusive (Bennett et al., 2010; Emanuele et al., 2011; Okumura et al., 2012).

In this study, we performed a screen to identify global CRL2 substrates. Our findings suggest that instead of serving as an exclusive selenoprotein quality inspector, CRL2 has a prominent function in PQC. We identified the auto-processed N-terminal fragment of USP1 as a physiological substrate of CRL2. Moreover, we uncovered a novel substrate recognition strategy in PQC, which we denominate as "DesCEND" (Destruction *via* the C-END).

RESULTS

Identified CRL2 substrates are nearly all defective

We applied Global Protein Stability (GPS) profiling to screen for CRL2 substrates (Fig. 1A). GPS is a fluorescence-based system for monitoring the stability of proteins in living cells with single-cell resolution (Yen et al., 2008). The protein of interest is fused at the GFP C-terminus, whereas RFP serves as a control to normalize protein synthesis. The GFP/RFP ratio is a surrogate for protein stability measurements, comparable to traditional half-life assays. By coupling GPS with the hORFeome 5.1 library (Lamesch et al., 2007), microarray deconvolution, and genetic ablation of CRL2 function by dominative-negative Cul2 (DNCul2), we identified 54 proteins having increased stability when CRL2 was inhibited (Fig. 1B, Table S1). We confirmed that the stability and abundance of these identified proteins increased upon different methods to abrogate CRL2 activity (Fig. 1C, 1D, Fig. S1A), and that these proteins are degraded by the proteasome (Fig. S1B). HIF1A and nine BC-box proteins were recovered (Fig. 1B, Table S1), indicative of a successful screen. Similarly, we identified the substrate receptors of CRLs from previous CRL substrate screens (Emanuele et al., 2011; Yen and Elledge, 2008).

We noticed that the sizes of most CRL2 substrates were smaller than predicted, with some being only slightly bigger than GFP (Fig. 1D). We sequenced these CRL2 substrates and found that 93% (42/45) were either truncated or nonsense (NS) peptides resulting from PCR errors or primer mutations during library construction (Fig. 1E, Table S1). We cloned the full-length versions of truncated proteins and found that CRL2 specifically targets truncated proteins but spares full-length ones (Fig. 1F). These data suggest a conspicuous function of CRL2 in the clearance of erroneous proteins.

CRL2 recognizes aberrant proteins through C-end degrons

We investigated how CRL2 identifies defective proteins. Codon replacements or removal of untranslated regions (UTR) did not affect CRL2-mediated degradation, suggesting that CRL2 recognized protein products (Fig. 2A, 2B). To test if active translation is required for substrate recognition, we examined whether CRL2 is capable of degrading accumulated CRL2 substrates released from MLN4924 treatment (a reversible inhibitor of CRLs; (Soucy et al., 2009). As shown in Fig. 2C and Fig. S1C, CRL2 efficiently targeted substrates exempted from MLN4924 treatment, suggesting that CRL2 post-translationally recognizes aberrant proteins.

We suspected that CRL2 might recognize the C-terminal tail (CTT) of abnormal proteins, since many CRL2 substrates are composed of short gibberish sequences fused at the C-terminus of GFP (Fig. 1D, Table S1). Changing the CTT of CRL2 substrates by either

deletion or masking completely inhibited CRL2-dependent degradation (Fig. 2A). Intriguingly, deletion of the last two residues did not always result in protein stabilization since that may create a new destabilizing CTT (e.g. NS2, NS3, NS16 in Fig. 2B), and as shown below, mutating the C-end degron (degradation signal) of one BC-box protein may create another one (Fig. 4D,5B). These data suggest that CRL2 utilizes protein C-termini as a protein quality indicator and supports our conclusion that aberrant proteins are tackled by CRL2 after protein synthesis is complete.

We tested if the CTT of CRL2 substrates include autonomous CRL2 degrons. The 12-residue CTT of CRL2 substrates, but not those of CDC25A or NRF2, directed CRL2-mediated degradation when tagged to the C-termini of unrelated stable proteins involved in diverse metabolic pathways (Fig. 2D, 2E). The degree of CTT-fusion-triggered destabilization was correlated with the CTT used in all tested proteins (i.e. NS12 ≈ NS19>RGAG1*>NS16>NS20), further supporting the notion that these chimeric proteins were degraded *via* a CTT-mediated mechanism (Fig. 2E). Those CTTs functioned as degrons only when placed at protein C-termini, suggesting that they are “C-END” degrons (Fig. 2F, S1D). The minimal length of most C-end degrons is only 6~ 10 residues (Fig. 2G, 2H, S1E). Unlike most known degrons involved in PQC, C-end degrons are not hydrophobic (Fig. 4A, 5A) (Fredrickson et al., 2011; Ravid and Hochstrasser, 2008). We have named this novel protein degradation mechanism DesCEND (Destruction *via* the C-END).

CRL2 targets aberrant proteins through various BC-box proteins

We mapped the BC-box proteins responsible for substrate recognition by overexpression and knockdown screens. The aberrant proteins we identified are mainly targeted by KLHDC3, FEM1C or APPBP2 (Fig. 3A, 3B, 3C, S2A, S2B). Each substrate preferentially associated with the corresponding BC-box protein (Fig. 3D). The 12-residue CTT of the CRL2 substrate was sufficient to confer specific BC-box binding when placed at the C-termini of reporter proteins (Fig. 3E, S2C). We observed similar results in all cell types tested, indicating the universality of CRL2-mediated DesCEND (Fig. 3F, S2D). Substrate information and their corresponding BC-box proteins are detailed in Table S1.

Characterization of C-end degrons

Together with our previous work (Lin et al., 2015), we have identified multiple CRL2 substrates tackled by the BC-box proteins KLHDC2, KLHDC3, FEM1C and APPBP2. We delineated the features of C-end degrons for each BC-box protein through sequence comparison and mutagenesis, and the results are summarized in Fig. 4A and Fig. 5A.

As an example, the last residue of KLHDC3 substrates is always glycine or alanine (Fig. 4A). Deleting the Gly^{-1}, but not other residues, stabilized the substrate (Fig. 4B) and invalidated KLHDC3-mediated degradation (Fig. 4C, 4D). Consistent with the critical role of Gly^{-1} or Ala^{-1}, changing these residues to other amino acids or adding a single amino acid downstream completely revoked KLHDC3-dependent degradation (Fig. 4D, S3A). In some cases Gly^{-1} can be substituted with the cognate small amino acid alanine (NS12, NS19), and replacement of the Ala^{-1} of NS10 with glycine converted NS10 into a stronger KLHDC3 substrate (Fig. S3A), suggesting that Gly^{-1} promoted KLHDC3-mediated

degradation because of its small size. In addition to Gly^{-1}, an upstream arginine is crucial. Most minimal KLHDC3 degrons begin with an arginine (Fig. 4A). Mutating the Arg significantly weakened or completely abolished KLHDC3-mediated degradation (Fig. 4E). Notably, deleting the Gly^{-1} residues of NS12 and NS19 changed them into FEM1C substrates, whereas adding a leucine downstream of Gly^{-1} converted NS19 into a weak APPBP2 substrate (Fig. 4D). Additional C-end degron conversions are reported in Fig. 5B.

KLHDC2 substrates all end with di-glycine (Fig. 4A). Mutating or deleting either glycine, or masking the terminal di-Gly, abrogated KLHDC2-mediated degradation (Fig. 4F).

The C-end degrons of FEM1C are longer than those of the other BC-box proteins we examined (Fig. S1E, see NS11). We identified a critical arginine located at either the −1, −2 or −3 positions of FEM1C substrates (Fig. 4A). Despite lysine also being positively charged, it cannot substitute for this arginine (Fig. 4G, S3B). Unlike the strict terminal location of the Gly or di-Gly in KLHDC3 and KLHDC2 degrons, respectively, the position of Arg in the FEM1C degron is substrate-dependent. In some cases (NS2, NS22 and CASC1*), adding a single amino acid totally abolished FEM1C-mediated degradation. In other cases that allowed degron capping (NS4, NS9 and NS11), the degree of tolerance was negatively correlated with the size of the amino acid added (Fig. 4G, S3B). Supporting the pivotal role of this Arg in FEM1C degrons, deleting Gly^{-1} of NS12 and NS19 exposed Arg^{-2}, thereby transforming them into FEM1C substrates (Fig. 4D).

APPBP2 degrons contain an “RxxG” motif that is similar to that of KLHDC3 degrons (Fig. 5A, 4A). Mutating either Arg or Gly stabilized APPBP2 substrates and abrogated APPBP2-mediated degradation (Fig. 5B, S3C). Consistently, changing the Arg or Gly abolished APPBP2 binding (Fig. 5C). KLHDC3 and APPBP2 degrons have two major differences (Fig. 4A, 5A). Firstly, in contrast to KLHDC3 degrons, the critical Gly in APPBP2 degrons cannot be the last residue. Deleting the amino acids C-terminal to the Gly residue converted APPBP2 targets into KLHDC3 substrates (Fig. 5B, last column). Secondly, whereas the spacing between Arg and Gly is flexible in KLHDC3 degrons, it is stringent for APPBP2 degrons (Fig. 5D). To further characterize APPBP2-C-end degron interactions, we mapped the substrate-binding region in APPBP2. APPBP2 contains a tetratricopeptide repeat (TPR) structural motif that frequently functions to mediate protein-protein interactions (Fig. S3D) (Lamb et al., 1995). We found that mutating several individual TPR repeats of APPBP2 significantly impaired substrate binding (Fig. 5E).

Cytosolic APP fragments inhibit $CRL2^{APPBP2}$-mediated degradation

APPBP2 was originally isolated as an amyloid precursor protein (APP)-binding protein (Zheng et al., 1998). Abnormal cleavage of APP generates extracellular beta amyloid (Aβ) peptide, i.e. the principle constituent of the amyloid plaques found in Alzheimer’s disease, and various intracellular fragments that are potentially cytotoxic (O’Brien and Wong, 2011; Zhang et al., 2011). How intracellular APP fragments influence cellular function is not fully understood.

We found that neither full-length APP nor intracellular APP fragments possess C-end degrons. Consistently, APP and its fragments were not substrates of CRL2 (Fig. S3E).

Instead, two APP fragments, C31 and AICD, but not full-length APP, inhibited APPBP2-mediated degradation (Fig. 5F, S3F). The effect was highly specific, as neither C31 nor AICD affected the stability of APPBP2 substrates when APPBP2 was absent (Fig. 5F, middle row), nor did they interfere with KLHDC3- and FEMlC-dependent degradation (Fig. 5F, bottom). These fragments stabilized APPBP2 substrates by reducing binding between APPBP2 and its substrates, but they did not influence the recruitment of APPBP2 into CRL2 complexes (Fig. 5G). These data indicate a potential cause of APP processing-induced cytotoxicity.

The N-terminal fragment of auto-cleaved USP1 is a physiological substrate of CRL2

We investigated the physiological function of CRL2-mediated DesCEND. Abundances of BC-box proteins remained constant under various proteotoxic treatments. We also did not detect significant growth defects in BC-box knockdown or overexpressed cells. Given the role of CRL2 in clearing proteins with aberrant C-termini, we examined whether CRL2 can alleviate the proteotoxic stress caused by inhibiting the nonsense-mediated decay (NMD) pathway or by facilitating stop codon read-through (Fig. S4A). As expected, inhibition of NMD by knocking down SMG1 or UPF1, or acceleration of stop codon read-through by reducing RF1 or RF3 levels, was toxic to cells (Fig. 6A, S4B). Nevertheless, changing the abundance of BC-box proteins neither rescued nor exacerbated growth defects (Fig. 6A, S4B). These data agree with our conclusion that CRL2 is highly selective for proteins with “specific” C-terminal ends.

We previously showed that CRL2 is responsible for clearance of five prematurely terminated selenoproteins arising from failures in UGA/Sec decoding (Lin et al., 2015). To identify additional physiological CRL2 substrates, we exploited the features of C-end degrons to postulate that truncated selenoprotein SEPW1 and ubiquitin specific peptidase 1 (USP1) are potential candidate targets. The C-terminal end of truncated SEPW1 generated from ambiguous UGA/Sec translation resembles both KLHDC2 and APPBP2 C-end degrons (Fig. 4A, 5A). Indeed, truncated SEPW1 but not its full-length version was specifically targeted by KLHDC2 and APPBP2 (Fig. 6B).

USP1 is a deubiquitinating enzyme that functions as a negative regulator in the DNA damage repair pathway (Cohn et al., 2007; Garcia-Santisteban et al., 2013; Huang et al., 2006; Nijman et al., 2005). To initiate DNA repair, USP1 self-inactivates by an auto-cleavage event, generating an N-terminal domain (NTD) and a C-terminal domain (CTD) (Huang et al., 2006). Both the NTD and CTD have to be removed for complete inactivation of USP1 (Cohn et al., 2007). Notably, USP1-NTD is terminated by a di-Gly, i.e. the characteristic of KLHDC2 degrons (Fig. 4A).

We detected both endogenous full-length and NTD of USP1 in HEK293T cells (Fig. 6C). Overexpression of KLHDC2 facilitated NTD degradation, but spared full-length USP1. Conversely, suppression of $CRL2^{KLHDC2}$ selectively stabilized NTD (Fig. 6C). We noticed that inhibition of $CRL2^{KLHDC2}$ did not completely prevent USP1-NTD degradation and this is likely because of a redundant ubiquitin ligase, APC/C^{Cdh1} (Cotto-Rios et al., 2011). USP1-NTD preferentially associated with KLHDC2 over other BC-box proteins (Fig. 6D). Covering or changing the terminal di-Gly avoided KLHDC2 binding (Fig. 6D) and

prevented CRL2^{KLHDC2}-mediated degradation in both HEK293T and U2OS cells (Fig. 6E, S4C). Ectopic placement of the CTT of USP1-NTD, but not its di-Gly mutants, was able to render irrelevant proteins degraded by CRL2^{KLHDC2} (Fig. 6F, 6G, S4D, S4E). Collectively, these data suggest that USP1-NTD is a *bona fide* CRL2^{KLHDC2} substrate.

There are over 70 human proteins that end with a di-Gly, including ubiquitin, SUMO, Nedd8, FAT10, ISG15 and URM1. We wondered whether di-Gly-terminated proteins are all targeted by KLHDC2. Whereas chimeras containing the CTT of ubiquitin and SUMO are CRL2^{KLHDC2} substrates (Fig. 6G), full-length ubiquitin and SUMO2 are not (Fig. 6H). Consistently, we were unable to detect physical binding between KLHDC2 and ubiquitin. Structurally, both ubiquitin and SUMO contain a globular domain possessing a six-amino acid tail (Fig. 6H), raising the possibility that this feature prevents C-end degrons from being targeted by BC-box proteins. We engineered ubiquitin and SUMO with longer C-terminal extensions by inserting Gly-Ser linkers between the globular domain and the six-amino acid tails (Fig. 6H). Consistent with our hypothesis, adding GS linkers enabled CRL2-mediated degradation (Fig. 6H). Interestingly, engineered ubiquitin was regulated by both KLHDC2 and KLHDC3. These results suggest that besides the intransigent terminal location, the structural accessibility of C-end degrons is critical for execution of CRL2-mediated DesCEND.

Full-length proteins bearing C-end degrons are regulated by DesCEND

We wondered whether full-length human proteins with cognate C-end degrons are potential targets of CRL2-mediated DesCEND. Any full-length protein with C-end degrons would be missing from our screen since the hORFeome collections are composed of open clones devoid of stop codons (Lamesch et al., 2007). We surveyed full-length human proteins with putative C-end degrons and identified PPP1R15A, USP49 and TCAP as CRL2^{KLHDC3} substrates (Fig. 7A). In agreement with C-end degron-driven degradation, changing their CTT stabilized these proteins and completely abolished CRL2-mediated degradation (Fig. 7B, 7C). Mutating all three arginines was required for complete inhibition of TCAP degradation because the spacing between Gly and Arg in KLHDC3 degrons is adjustable (Fig. 7B, 5D). These data suggest that besides clearing incomplete proteins with unusual ends, CRL2-mediated DesCEND may have a broader impact on the proteome. The success of applying C-end degron features to identify novel DesCEND targets demonstrates the potential of using this strategy to capture additional DesCEND substrates. How many native proteins bearing C-end degrons are controlled by CRL2-mediated DesCEND and how DesCEND reshapes the human proteome remain subjects for future study.

DISCUSSON

Several things may go wrong during generation of a single protein. How to capture such a wide spectrum of defective proteins using limited regulators, as well as which molecular characteristics distinguish proteins directed to the “correct” pool versus the “incorrect” pool, are fundamental questions in protein quality surveillance. Here, we define a differentiation mechanism, DesCEND, which utilizes the C-termini of proteins as an indicator of protein quality/integrity.

DesCEND possesses several unique features. In contrast to the rather broad primary sequence specificity of known degrons involved in protein quality control, DesCEND exhibits unprecedented precision. Masking or changing the C-end degron with even a single residue completely abrogated DesCEND. CRL2 exploits interchangeable BC-box proteins to extend the plasticity of substrate recognition. At least five BC-box proteins are involved in DesCEND; namely KLHDC2, KLHDC3, FEM1C, APPBP2 and PRAME members (Lin et al., 2015). BC-box proteins involved in DesCEND are expressed in all cell types and are overexpressed in cancer (Berglund et al., 2008; Forbes et al., 2017; Monni et al., 2001; Uhlen et al., 2015). Substrates of DesCEND are not restricted to any specific cellular compartment. Notably, USP1-CTD is removed by the N-end rule pathway that utilizes the N-terminal residue of a protein as a signal for degradation (N-end degron) (Piatkov et al., 2012). The elimination of USP1-NTD through its C-end degron mirrors the process of USP1-CTD elimination *via* its N-end degron, highlighting the dominance of employing "illegal ends" in reporting protein integrity.

Apart from their absolute C-terminal position, the C-end degrons of different BC-box proteins share common features. C-end degrons are autonomous, portable and mostly only 6~ 10 residues in length. They are neither defined by compositional enrichment nor a fixed sequence of amino acids, but by a few indispensable residues with the remainder being rather degenerate (Fig. 7D). These latter "flexible" residues do not define but rather tune CRL2-mediated degradation. Strikingly, Gly or Arg is consistently the essential residue in different C-end degrons, and the Arg cannot be replaced by Lys. A single protein may be simultaneously a substrate of multiple BC-box proteins (e.g. SELS, SEPX1 and ubiquitin with GS linkers) (Lin et al., 2015), and changing it by as little as a single residue of its C-end degron swaps the respective BC-box protein. Intuitively, it makes sense to assume that C-end degrons are closely related and that their corresponding BC-box proteins recognize them *via* similar molecular mechanisms. Surprisingly, BC-box proteins that function in DesCEND contain disparate structural motifs involved in protein-protein interactions, such as Kelch, LRR, ANK, and TPR repeats. For instance, whereas the KLHDC3 and APPBP2 degrons comprise similar RxxG motifs, KLHDC3 and APPBP2 possess unrelated Kelch and TPR repeats, respectively. Defining the structural basis of C-end degron-BC-box protein interactions should shed light on this perplexing specificity rule concerning the pairing of C-end degrons and BC-box proteins.

An ideal protein surveillance system applies a few regulators to tackle diverse erroneous proteins. As a result, the degrons for protein quality control are expected to be sufficiently lax that most aberrant proteins contain them. In line with targeting proteins possessing unusual C-termini, substrates of DesCEND may be derived from translation errors (e.g. truncated selenoproteins or defective ribosomal products, DRiPs) (Lin et al., 2015; Yewdell et al., 1996), proteolytic events (e.g. USP1-NTD), post-translational damage, or C-terminal modifications. Nevertheless, given the fastidious nature of DesCEND, the broad possibilities of aberrant ends, and the fact that far fewer BC-box proteins exist (<40) and not all of them are involved in DesCEND, it seems unlikely that CRL2-mediated DesCEND serves as a global quality gatekeeper for truncated proteins. Instead, one tempting possibility is that DesCEND is customized for the clearance of a specific source of proteins with definite C-

terminal ends. Alternatively, additional ubiquitin ligases that function in DesCEND may yet be unveiled.

We propose a model to explain the potential physiological functions of CRL2-mediated DesCEND (Fig. 7E). Our data suggests that apart form clearing truncated proteins with acquired aberrant ends, DesCEND also regulates full-length proteins with naturally-occurring C-end degrons. Moreover, structural accessibility is necessary for C-end degrons to trigger degradation. If shielded by the native structure of proteins, end-located C-end degrons do not promote degradation, raising an intriguing possibility that DesCEND plays a further role in protein quality control by degrading misfolded or uncomplexed proteins when the originally buried C-end degron is exposed (Fig. 7E). This idea is consistent with the discovery that conditional N-end degrons regulate subunit stoichiometries by steric sequestration (Shemorry et al., 2013).

Paralleling our study, C-terminal tails have been shown to serve as the degradation signal for many proteins, including c-FLIPs, Noxa, SMNΔ7, mouse ornithine decarboxylase, and hantavirus G1 protein (Cho and Dreyfuss, 2010; Loetscher et al., 1991; Pang et al., 2014; Poukkula et al., 2005; Sen et al., 2007). Moreover, aberrant proteins resulting from stop codon read-through were shown to be more labile than their wild-type counterparts, with the aberrant C-terminal extensions hosting degrons (Arribere et al., 2016; Shibata et al., 2015). Interestingly, in bacteria, proteins synthesized from mRNAs without stop codons are marked by a C-terminal addition of an 11-residue ssM-encoded peptide tail and are subsequently degraded by C-terminal-specific proteases (Gottesman et al., 1998; Karzai et al., 2000; Keiler et al., 1996). Although it may operate according to distinct mechanisms, C-terminal tail-specific recognition for protein degradation is likely more widespread than heretofore acknowledged.

STAR METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Hsueh-Chi S. Yen (hyen@imb.sinica.edu.tw).

EXPERIMENTAL MODEL AND SUBJECT DETAILS Cell lines

HEK293T, HeLa and HepG2 cells were maintained in DMEM with 10% FBS and antibiotics at 37°C in a 6% CO2 atmosphere. U2OS cells were cultured in McCoy's 5A medium with 10% FBS and antibiotics. Tissue culture media and supplements were from Gibco®| Life Technologies (Carlsbad, CA, *USA*). Experiments were performed in HEK293T cells unless otherwise indicated.

METHOD DETAILS

Tissue culture—To generate GPS reporter cells, cells were infected with lentiviruses carrying GPS reporter constructs at low multiplicity of infection (MOI~0.1) and infected cells were selected by puromycin (1 μg/mL, Clontech Laboratories, Mountain View, CA, USA). To block CRL2 function, cells were either treated with 1 μM MLN4924 (Active

Biochem, Maplewood, NJ, USA) for 6 hrs, infected with viruses carrying DNCul2 (MOI~10) for 40 hrs, or treated with shCul2 (Santa Cruz Biotechnology, Dallas, TX, USA) for 88 hrs. To block proteasome function, cells were either treated with 1 μM Bortezomib (Biovision, Milpitas, CA, USA) or 10 μM MG-132 (Merck Millipore, Billerica, MA, USA). To block the autophagy-lysosomal pathway, cells were either treated with 5 mM 3-methyladenine (Sigma-Aldrich, St. Louis, MO, USA) or 10 nM Bafilomycin A1 (LC Laboratories, Woburn, MA, USA). To study the function BC-box proteins, cells were infected with viruses expressing BC-box proteins or shRNAs against BC-box proteins and analyzed 20 hrs or 88 hrs thereafter, respectively.

To prepare lentiviruses, HEK293T cells were transfected with pHAGE, pRev, pTat, pHIV gag/pol and pVSVG using TransIT-293 transfection reagent (Mirus Bio LLC, Madison, WI, USA). To produce retroviruses, HEK293T cells were transfected with pMSCV, pGag/pol and pVSVG. Viruses were harvested 48 hrs after transfection and infection was conducted in medium containing 8 μg/mL polybrene (Sigma-Aldrich, St. Louis, MO, USA).

DNA cloning—To generate GPS reporter constructs, genes of interest were cloned into pLenti-GPS by Gateway recombination (Invitrogen, Carlsbad, CA, USA). A GPS v5.1 library was created by using the hORFeome v.5.1 library that contains 15,483 human ORFs (Lamesch et al., 2007). To examine the function of CTTs when placed in the middle of a protein, a HindIII fragment encoding the CTT of interest was cloned into the internal HindIII site of GAPDH, SARS and EEF1A1 genes. To investigate the role of APP in CRL2-mediated degradation, full-length, AICD and C31 APP were cloned into pLenti-GPS for stability analysis or into pHAGE for functional studies. The AICD and C31 fragments contain the last 50 and last 31 amino acids of APP, respectively.

GPS assay and GPS screen for CRL2 substrates—GPS reporter cells were analyzed by the LSR Fortessa system (BD Biosciences, San Jose, CA, USA). Multiple GFP/RFP ratio scaling was recorded for optimal resolution and to avoid off-scaling.

The CRL2 GPS screen was performed as previously described (Emanuele et al., 2011; Yen and Elledge, 2008; Yen et al., 2008). In brief, HEK293T cells carrying the GPS v5.1 library (Emanuele et al., 2011) were infected with either control lentiviruses or lentiviruses expressing DNCul2 at high MOI (~10) for 24 hrs. The cells were then harvested and sorted using FACSAria (BD Biosciences) based on the GFP/RFP ratio. Protein stability information was deconvoluted by microarray using ORFs amplified from the genome of GPS reporter cells as probes. The screen was validated using MLN4924 and shCul2 treatments as alternative approaches to blocking CRL2 function. We originally reported that we identified 102 substrates (Lin et al., 2015), but later found that 49 of them represent identical proteins and, therefore, the correct number of substrates identified from our GPS screen is 54.

Cycloheximide-chase, GST pull-down and Western blotting—Cycloheximide (CHX)-chase experiments were conducted by treating cells with 100 Mg/mL cycloheximide (Calbiochem, San Diego, CA, USA), followed by sample collection at multiple time-points. Samples were collected by directly lysing cultured cells in 1× SDS sample buffer followed

by sonication. Protein abundance was then subjected to Western blot analysis and quantified using the ImageJ software (National Institute of Health, Bethesda, MD, USA).

For GST pull-down assays, GPS reporter cells were transfected with GST-tagged constructs for 42 hrs followed by treatment with 1 μM of the proteasome inhibitor Bortezomib (Biovision, Milpitas, CA, USA) and 1 μM MLN4924 for 6 hrs to block protein degradation. Cells were then lysed in RIPA buffer (150 mM NaCl, 1.0% IGEPAL® CA-630, 0.5% sodium deoxycholate, 0.1% SDS, and 50 mM Tris, pH 8.0) supplemented with protease and phosphatase inhibitors (Roche, Basel, CH), mixed with binding buffer (0.5% CA630 in 1× PBS), and incubated with Glutathione Sepharose 4B (GE Healthcare Life Science, Little Chalfont, UK) for 16 hrs at 4°C. The corresponding Sepharose was washed three times in binding buffer, boiled in 2× SDS sample buffer, and subjected to Western blot analysis. We only detected BC-box protein-substrate binding when GST-tagged BC-box proteins and corresponding substrates were co-expressed in the same cell. Antibodies to the following epitopes and proteins were purchased from the indicated vendors: HA (16B12, Covance, Princeton, NY, USA); GFP (JL-8, Clontech); Tubulin (Ab-2, Thermo Fisher Scientific, Waltham, MA, USA); GST (27457701, GE Healthcare); Flag (M2, Sigma-Aldrich); Cul2 (C-4, Santa Cruz Biotechnology); and USP1 (A301–669A, Bethyl Laboratories, Montgomery, TX, USA).

Competition assay—HEK293T cells were either stably overexpressed or knocked-down with BC-box proteins followed by infection of lentiviruses carrying pGIPZ-RFP-control vector or pGIPZ-GFP-shRNAs against SMG1, UPF1, RF1 or RF3. RFP- and GFP-expressing cells were then mixed in equal amounts before being subcultured and subjected to FACS analysis every two days.

Targeting sequences for shRNAs—KLHDC3: TGGAAAAAGATTGAACCGA (#1); CCAATGACATTCACAAGCT (#2); GAGATGAATTTGACCTTAT (#3); TGCTGTATTGTTGGTGACA (#4)

FEM1C: GTAACAGTTGTTTCATAAA (#1); ACCAAATTGTTGGCAAGCA (#2); GAGCTACATTTGTAGACAA (#3); CTCTTACTATATTAGATAT (#4)

APPBP2: GCCTTCAGTTGTGTACTCT (#1); GACATCTGGCTTCTTTATA (#2); TGATGGATCATGGTGTTAA (#3); CAGTTTGATGTTTACTACA (#4)

VHL: TGGCTCAACTTCGACGGCG (#1); AGATCTGGAAGACCACCCA (#2)

KLHDC2 CTTGGTGTCTGGGTATATA

ZER1: CTCTCTTCTACCTAACAAA

LRR1:ATATGGCTCTCATATCATT

KLHDC1: GGGTATATAAAGACAGACT

KLHDC4: TCAGACATGTTCCTGCTGA

KLHDC10: GGTGTCACAGTTGTGTTCA

FEM1A: GGGAGCTACGTATGTGGAT

FEM 1B:AGAAGATCAGTGCAAAATT

SMG1: CAGAAGCACTTCGATGTTT (#1); GATTCCATTTAAAGAGATA (#2)

UPF1: GAGTCCCAGACTCAAGATA (#1); CTACCAGTACCAGAACATA (#2)

RF1: AGGACGATACTTTGATGAA (#1); GCACTTCTCACATGAGAAA (#2)

RF3: AGAAAGAGCATGTAAATGT (#1); ACGTGGAAGTTCTTGGAAT (#2)

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank DE Hill, M Vidal and JW Harper for reagents, PH Hsu for mass spectrometry analysis, and YN Chen, WC Hsu, SC Ho, WC Huang, J O'Brien, E Yeh and the RNAi Core Facility at Academia Sinica for suggestions and assistance. This work was supported by Career Development Award 101-CDA-L05 from Academia Sinica and MOST grant 106–2321-B-001–045 awarded to HCY, and NIH grant AG11085 to SJE. SJE and NZ are Howard Hughes Medical Institute Investigators.

REFERENCES

Arribere JA, Cenik ES, Jain N, Hess GT, Lee CH, Bassik MC, and Fire AZ (2016). Translation readthrough mitigation. Nature 534, 719–723. [PubMed: 27281202]

Bachmair A, Finley D, and Varshavsky A (1986). In vivo half-life of a protein is a function of its amino-terminal residue. Science 234, 179–186. [PubMed: 3018930]

Balch WE, Morimoto RI, Dillin A, and Kelly JW (2008). Adapting proteostasis for disease intervention. Science 319, 916–919. [PubMed: 18276881]

Bengtson MH, and Joazeiro CA (2010). Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. Nature 467, 470–473. [PubMed: 20835226]

Bennett EJ, Rush J, Gygi SP, and Harper JW (2010). Dynamics of cullin-RING ubiquitin ligase network revealed by systematic quantitative proteomics. Cell 143, 951–965. [PubMed: 21145461]

Berglund L, Bjorling E, Oksvold P, Fagerberg L, Asplund A, Szigyarto CA, Persson A, Ottosson J, Wernerus H, Nilsson P, et al. (2008). A genecentric Human Protein Atlas for expression profiles based on antibodies. Mol Cell Proteomics 7, 2019–2027. [PubMed: 18669619]

Brandman O, Stewart-Ornstein J, Wong D, Larson A, Williams CC, Li GW, Zhou S, King D, Shen PS, Weibezahn J, et al. (2012). A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. Cell 151, 1042–1054. [PubMed: 23178123]

Chen B, Retzlaff M, Roos T, and Frydman J (2011). Cellular strategies of protein quality control. Cold Spring Harb Perspect Biol 3, a004374. [PubMed: 21746797]

Cho S, and Dreyfuss G (2010). A degron created by SMN2 exon 7 skipping is a principal contributor to spinal muscular atrophy severity. Genes Dev 24, 438–442. [PubMed: 20194437]

Ciechanover A (1994). The ubiquitin-proteasome proteolytic pathway. Cell 79, 13–21. [PubMed: 7923371]

Ciechanover A, Finley D, and Varshavsky A (1984). The ubiquitin-mediated proteolytic pathway and mechanisms of energy-dependent intracellular protein degradation. J Cell Biochem 24, 27–53. [PubMed: 6327743]

Cohn MA, Kowal P, Yang K, Haas W, Huang TT, Gygi SP, and D'Andrea AD (2007). A UAF1-containing multisubunit protein complex regulates the Fanconi anemia pathway. Mol Cell 28, 786–797. [PubMed: 18082604]

Cotto-Rios XM, Jones MJ, Busino L, Pagano M, and Huang TT (2011). APC/CCdh1-dependent proteolysis of USP1 regulates the response to UV-mediated DNA damage. J Cell Biol 194, 177–186. [PubMed: 21768287]

Deshaies RJ, and Joazeiro CA (2009). RING domain E3 ubiquitin ligases. Annu Rev Biochem 78, 399–434. [PubMed: 19489725]

Emanuele MJ, Elia AE, Xu Q, Thoma CR, Izhar L, Leng Y, Guo A, Chen YN, Rush J, Hsu PW, et al. (2011). Global identification of modular cullin-RING ligase substrates. Cell 147, 459–474. [PubMed: 21963094]

Fang NN, Chan GT, Zhu M, Comyn SA, Persaud A, Deshaies RJ, Rotin D, Gsponer J, and Mayor T (2014). Rsp5/Nedd4 is the main ubiquitin ligase that targets cytosolic misfolded proteins following heat stress. Nat Cell Biol 16, 1227–1237. [PubMed: 25344756]

Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, Cole CG, Ward S, Dawson E, Ponting Lv et al. (2017). COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res 45, D777–D783. [PubMed: 27899578]

Fredrickson EK, Rosenbaum JC, Locke MN, Milac TI, and Gardner RG (2011). Exposed hydrophobicity is a key determinant of nuclear quality control degradation. Mol Biol Cell 22, 2384–2395. [PubMed: 21551067]

Garcia-Santisteban I, Peters GJ, Giovannetti E, and Rodriguez JA (2013). USP1 deubiquitinase: cellular functions, regulatory mechanisms and emerging potential as target in cancer therapy. Mol Cancer 12, 91. [PubMed: 23937906]

Glickman MH, and Ciechanover A (2002). The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. Physiol Rev 82, 373–428. [PubMed: 11917093]

Goldberg AL (2003). Protein degradation and protection against misfolded or damaged proteins. Nature 426, 895–899. [PubMed: 14685250]

Gottesman S, Roche E, Zhou Y, and Sauer RT (1998). The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. Genes Dev 12, 1338–1347. [PubMed: 9573050]

Hershko A, and Ciechanover A (1998). The ubiquitin system. Annu Rev Biochem 67, 425–479. [PubMed: 9759494]

Huang TT, Nijman SM, Mirchandani KD, Galardy PJ, Cohn MA, Haas W, Gygi SP, Ploegh HL, Bernards R, and D'Andrea AD (2006). Regulation of monoubiquitinated PCNA by DUB autocleavage. Nat Cell Biol 8, 339–347. [PubMed: 16531995]

Ivan M, Kondo K, Yang H, Kim W, Valiando J, Ohh M, Salic A, Asara JM, Lane WS, and Kaelin WG, Jr. (2001). HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. Science 292, 464–468. [PubMed: 11292862]

Jaakkola P, Mole DR, Tian YM, Wilson MI, Gielbert J, Gaskell SJ, von Kriegsheim A, Hebestreit HF, Mukherji M, Schofield CJ, et al. (2001). Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. Science 292, 468–472. [PubMed: 11292861]

Kamura T, Maenaka K, Kotoshiba S, Matsumoto M, Kohda D, Conaway RC, Conaway JW, and Nakayama KI (2004). VHL-box and SOCS-box domains determine binding specificity for Cul2-Rbx1 and Cul5-Rbx2 modules of ubiquitin ligases. Genes Dev 18, 3055–3065. [PubMed: 15601820]

Karzai AW, Roche ED, and Sauer RT (2000). The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue. Nat Struct Biol 7, 449–455. [PubMed: 10881189]

Keiler KC, Waller PR, and Sauer RT (1996). Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. Science 271, 990–993. [PubMed: 8584937]

Lamb JR, Tugendreich S, and Hieter P (1995). Tetratrico peptide repeat interactions: to TPR or not to TPR? Trends Biochem Sci 20, 257–259. [PubMed: 7667876]

Lamesch P, Li N, Milstein S, Fan C, Hao T, Szabo G, Hu Z, Venkatesan K, Bethel G, Martin P, et al. (2007). hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes. Genomics 89, 307–315. [PubMed: 17207965]

Lin HC, Ho SC, Chen YY, Khoo KH, Hsu PH, and Yen HC (2015). SELENOPROTEINS. CRL2 aids elimination of truncated selenoproteins produced by failed UGA/Sec decoding. Science 349, 91–95. [PubMed: 26138980]

Loetscher P, Pratt G, and Rechsteiner M (1991). The C terminus of mouse ornithine decarboxylase confers rapid degradation on dihydrofolate reductase. Support for the pest hypothesis. J Biol Chem 266, 11213–11220. [PubMed: 2040628]

Mahrour N, Redwine WB, Florens L, Swanson SK, Martin-Brown S, Bradford WD, Staehling-Hampton K, Washburn MP, Conaway RC, and Conaway JW (2008). Characterization of Cullin-box sequences that direct recruitment of Cul2-Rbx1 and Cul5-Rbx2 modules to Elongin BC-based ubiquitin ligases. J Biol Chem 283, 8005–8013. [PubMed: 18187417]

Monni O, Barlund M, Mousses S, Kononen J, Sauter G, Heiskanen M, Paavola P, Avela K, Chen Y, Bittner ML, et al. (2001). Comprehensive copy number and gene expression profiling of the 17q23 amplicon in human breast cancer. Proc Natl Acad Sci U S A 98, 5711–5716. [PubMed: 11331760]

Murata S, Minami Y, Minami M, Chiba T, and Tanaka K (2001). CHIP is a chaperone-dependent E3 ligase that ubiquitylates unfolded protein. EMBO Rep 2, 1133–1138. [PubMed: 11743028]

Nijman SM, Huang TT, Dirac AM, Brummelkamp TR, Kerkhoven RM, D'Andrea AD, and Bernards R (2005). The deubiquitinating enzyme USP1 regulates the Fanconi anemia pathway. Mol Cell 17, 331–339. [PubMed: 15694335]

O'Brien RJ, and Wong PC (2011). Amyloid precursor protein processing and Alzheimer's disease. Annu Rev Neurosci 34, 185–204. [PubMed: 21456963]

Okumura F, Matsuzaki M, Nakatsukasa K, and Kamura T (2012). The role of Elongin BC-containing ubiquitin ligases. Front Oncol 2, 10. [PubMed: 22649776]

Pang X, Zhang J, Lopez H, Wang Y, Li W, O'Neill KL, Evans JJ, George NM, Long J, Chen Y, et al. (2014). The carboxyl-terminal tail of Noxa protein regulates the stability of Noxa and Mcl-1. J Biol Chem 289, 17802–17811. [PubMed: 24811167]

Petroski MD, and Deshaies RJ (2005). Function and regulation of cullin-RING ubiquitin ligases. Nat Rev Mol Cell Biol 6, 9–20. [PubMed: 15688063]

Piatkov KI, Colnaghi L, Bekes M, Varshavsky A, and 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. [PubMed: 23159736]

Poukkula M, Kaunisto A, Hietakangas V, Denessiouk K, Katajamaki T, Johnson MS, Sistonen L, and Eriksson JE (2005). Rapid turnover of c-FLIPshort is determined by its unique C-terminal tail. J Biol Chem 280, 27345–27355. [PubMed: 15886205]

Powers ET, and Balch WE (2013). Diversity in the origins of proteostasis networks--a driver for protein function in evolution. Nat Rev Mol Cell Biol 14, 237–248. [PubMed: 23463216]

Ravid T, and Hochstrasser M (2008). Diversity of degradation signals in the ubiquitin-proteasome system. Nat Rev Mol Cell Biol 9, 679–690. [PubMed: 18698327]

Rosenbaum JC, Fredrickson EK, Oeser ML, Garrett-Engele CM, Locke MN, Richardson LA, Nelson ZW, Hetrick ED, Milac TI, Gottschling DE, et al. (2011). Disorder targets misorder in nuclear quality control degradation: a disordered ubiquitin ligase directly recognizes its misfolded substrates. Mol Cell 41, 93–106. [PubMed: 21211726]

Ross CA, and Poirier MA (2004). Protein aggregation and neurodegenerative disease. Nat Med 10 Suppl, S10–17. [PubMed: 15272267]

Sen N, Sen A, and Mackow ER (2007). Degrons at the C terminus of the pathogenic but not the nonpathogenic hantavirus G1 tail direct proteasomal degradation. J Virol 81, 4323–4330. [PubMed: 17267477]

Shastry BS (2003). Neurodegenerative disorders of protein aggregation. Neurochem Int 43, 1–7. [PubMed: 12605877]

Shemorry A, Hwang CS, and Varshavsky A (2013). Control of protein quality and stoichiometries by N-terminal acetylation and the N-end rule pathway. Mol Cell 50, 540–551. [PubMed: 23603116]

Shibata N, Ohoka N, Sugaki Y, Onodera C, Inoue M, Sakuraba Y, Takakura D, Hashii N, Kawasaki N, Gondo Yv et al. (2015). Degradation of stop codon read-through mutant proteins via the ubiquitin-proteasome system causes hereditary disorders. J Biol Chem 290, 28428–28437. [PubMed: 26442586]

Skovronsky DM, Lee VM, and Trojanowski JQ (2006). Neurodegenerative diseases: new concepts of pathogenesis and their therapeutic implications. Annu Rev Pathol 1, 151–170. [PubMed: 18039111]

Soucy TA, Smith PG, Milhollen MA, Berger AJ, Gavin JM, Adhikari S, Brownell JE, Burke KE, Cardin DP, Critchley S, et al. (2009). An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. Nature 458, 732–736. [PubMed: 19360080]

Sriram SM, Kim BY, and Kwon YT (2011). The N-end rule pathway: emerging functions and molecular principles of substrate recognition. Nat Rev Mol Cell Biol 12, 735–747. [PubMed: 22016057]

Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson A, Kampf C, Sjostedt E, Asplund A, et al. (2015). Proteomics. Tissue-based map of the human proteome. Science 347, 1260419. [PubMed: 25613900]

Varshavsky A (2011). The N-end rule pathway and regulation by proteolysis. Protein Sci 20, 1298–1345. [PubMed: 21633985]

Wickner S, Maurizi MR, and Gottesman S (1999). Posttranslational quality control: folding, refolding, and degrading proteins. Science 286, 1888–1893. [PubMed: 10583944]

Wolff S, Weissman JS, and Dillin A (2014). Differential scales of protein quality control. Cell 157, 52–64. [PubMed: 24679526]

Yen HC, and Elledge SJ (2008). Identification of SCF ubiquitin ligase substrates by global protein stability profiling. Science 322, 923–929. [PubMed: 18988848]

Yen HC, Xu Q, Chou DM, Zhao Z, and Elledge SJ (2008). Global protein stability profiling in mammalian cells. Science 322, 918–923. [PubMed: 18988847]

Yewdell JW, Anton LC, and Bennink JR (1996). Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules? J Immunol 157, 1823–1826. [PubMed: 8757297]

Zhang YW, Thompson R, Zhang H, and Xu H (2011). APP processing in Alzheimer's disease. Mol Brain 4, 3. [PubMed: 21214928]

Zheng P, Eastman J, Vande Pol S, and Pimplikar SW (1998). PAT1, a microtubule-interacting protein, recognizes the basolateral sorting signal of amyloid precursor protein. Proc Natl Acad Sci U S A 95, 14745–14750. [PubMed: 9843960]

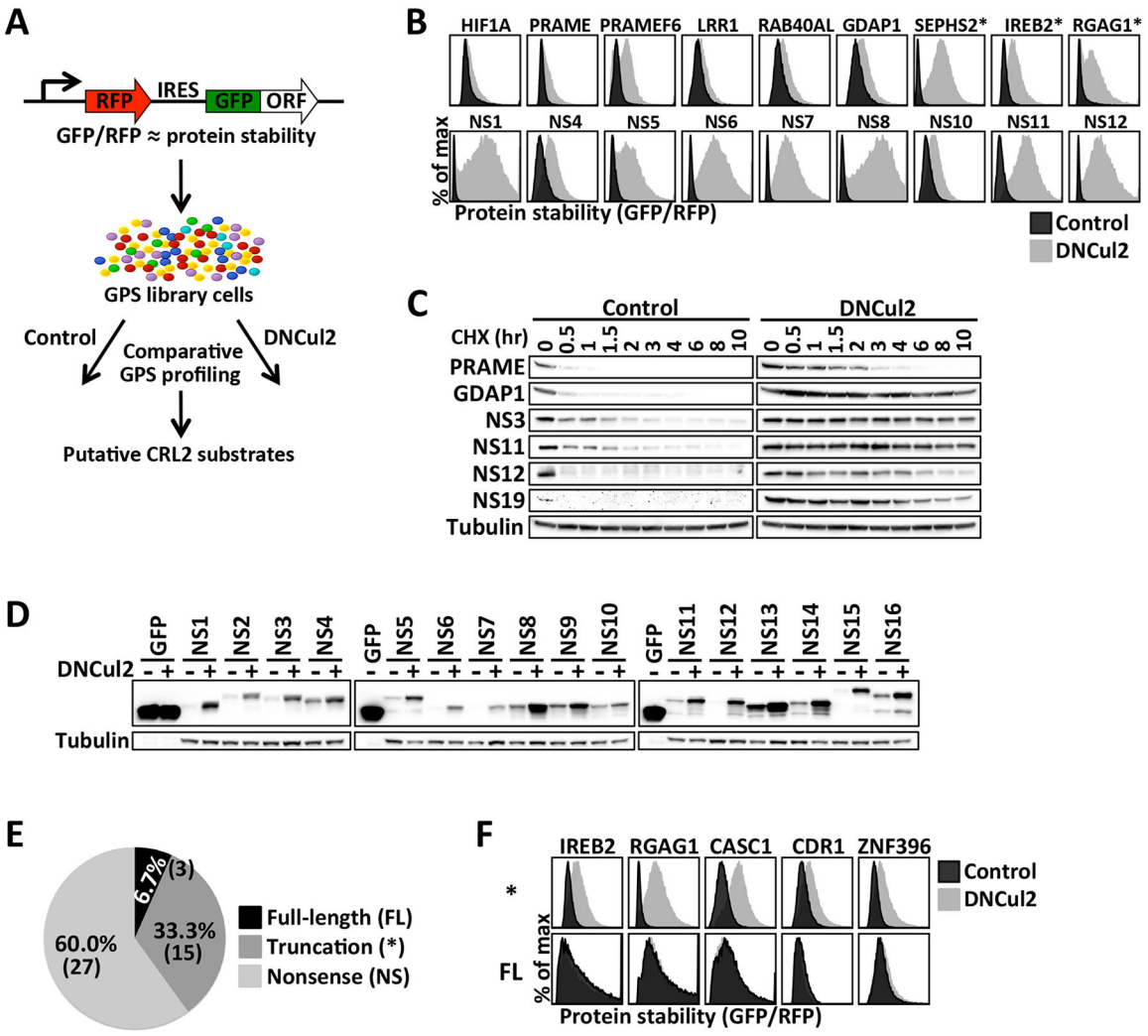

Figure 1.

CRL2 functions in protein quality control. (A) A schematic representation of the CRL2 GPS screen. The GPS reporter system is based on the expression of two fluorescent proteins from a single promoter enabled by an internal ribosome entry site (IRES). GFP is fused to the N-terminus of the protein of interest, whereas RFP serves as an internal control. The GFP/RFP ratio represents protein stability. To search for CRL2 substrates, the GPS v5.1 HEK293T cell library was treated or not treated with DNCul2 and compared. (B) GPS reporter cells expressing putative CRL2 substrates were treated or not treated with DNCul2 and analyzed. Truncated and nonsense proteins are marked as * and NS, respectively. (C) Cycloheximide (CHX)-chase analysis of identified CRL2 substrates with or without DNCul2 treatment. (D) Western blot analysis of identified CRL2 substrates with or without DNCul2 treatment. Tubulin serves as a loading control. (E) Sequence analysis of CRL2 substrates. Protein numbers are indicated in parentheses. (F) GPS assay of truncated CRL2 substrates and their corresponding full-length versions. See also Figure S1 and Table S1.

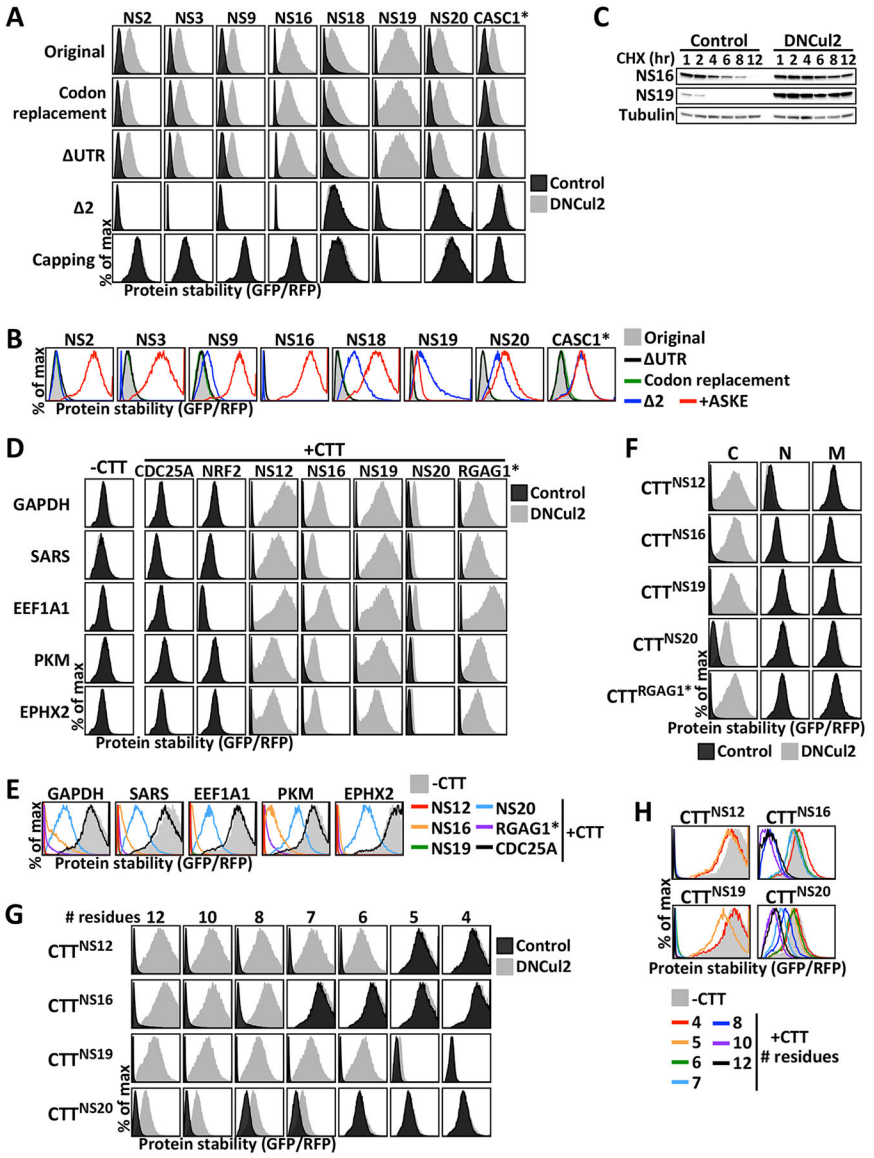

Figure 2.

CRL2 recognizes the C-termini of aberrant proteins. (A) GPS cells expressing substrates labeled at the top with mutations indicated at left were treated or not treated with DNCul2 and analyzed. “Original” represents the original clone identified from the GPS screen. To change the protein C-terminus, the last two amino acids of substrates were deleted (A2) or the last four residues of GAPDH (ASKE) were added for capping. Due to large variations in protein stability, each GPS plot is presented with distinct ratio scaling for better resolution. As a result, the GFP/RFP ratios from different plots cannot be compared. (B) Stability comparison among proteins in (A). (C) GPS cells with or without DNCul2 virus infection were pretreated with MLN4924 for 8 hours, released, and the stability of accumulated CRL2 substrates were analyzed by CHX-chase assay. (D) Stability analysis of proteins indicated at left with or without C-terminal-tagging of the 12-residue CTT of proteins indicated at the top. (E) Stability comparisons among proteins in (D). (F) GPS assay of GAPDH with

various CTTs added at its C-terminus (C), N-terminus (N) or middle (M). (G, H) Stability analysis of GAPDH C-terminally tagged with various lengths of CTTs. See also Figure S1 and Table S1.

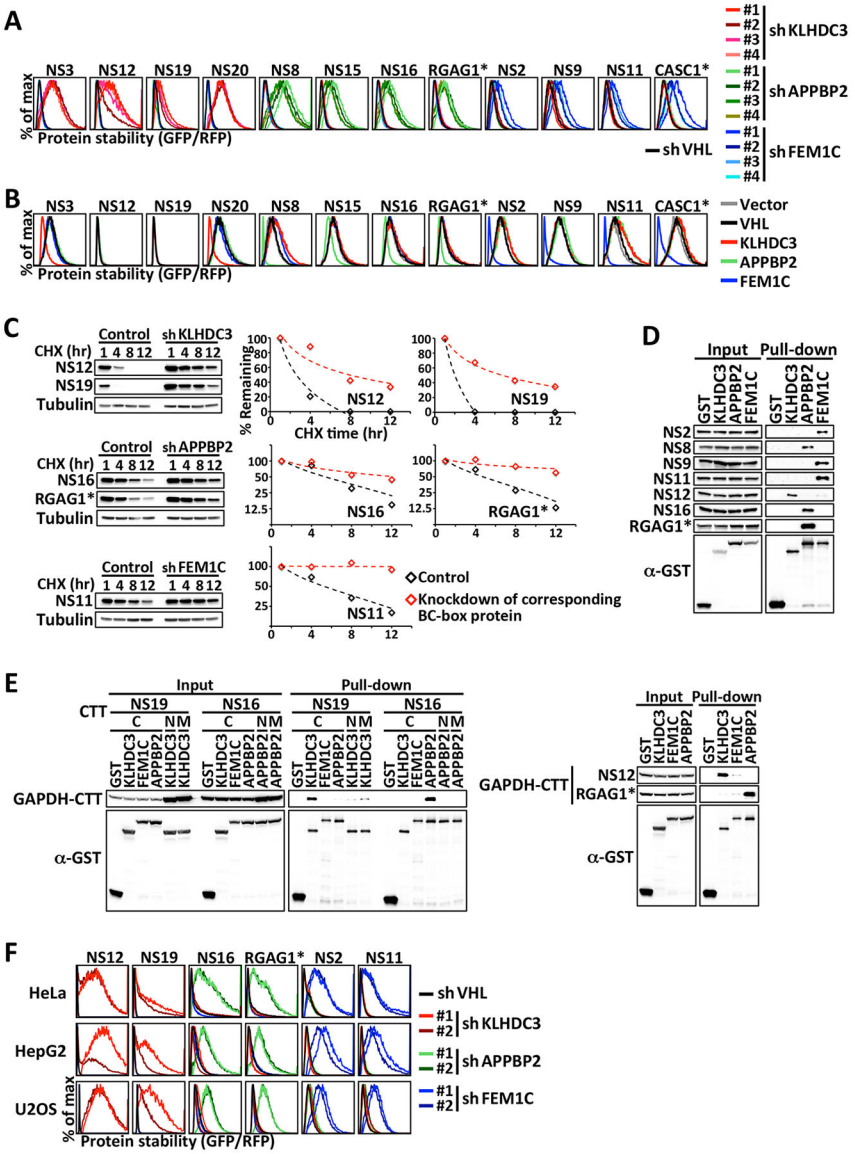

Figure 3.

CRL2 targets aberrant proteins through various BC-box proteins. (A) GPS cells carrying indicated CRL2 substrates were treated with shRNAs against various BC-box proteins and analyzed. (B) GPS assay for cells infected with viruses expressing various BC-box proteins. (C) CHX-chase analysis of CRL2 substrates in cells with or without BC-box protein knockdown. (D) GST pull-down assay using cells expressing GST or GST-tagged BC-box proteins and GFP-tagged CRL2 substrates. (E) Physical interaction between GST-tagged BC-box proteins and GAPDH with the 12-residue CTT from various CRL2 substrates fused at its C-terminus (C), N-terminus (N) or middle (M). (F) Stability analysis of CRL2 substrates in various cells. See also Figure S2.

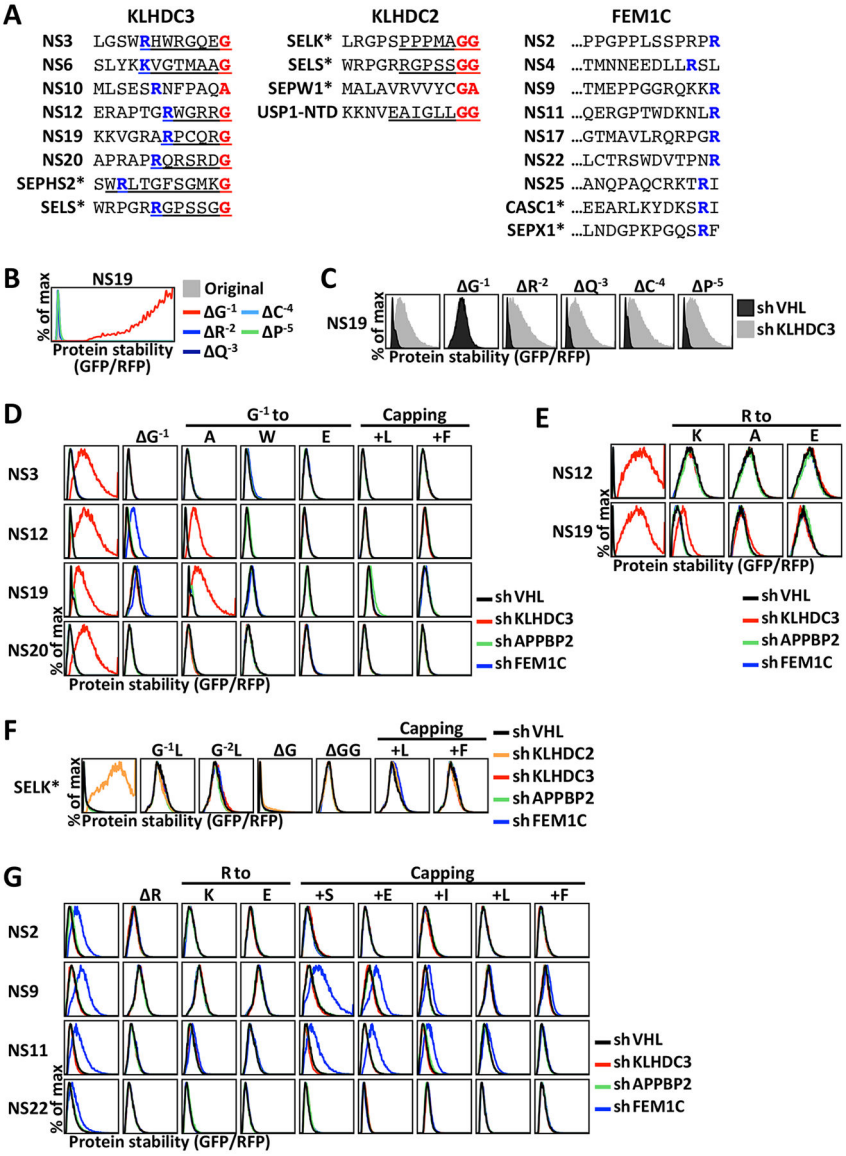

Figure 4.

Characterization of KLHDC3, KLHDC2 and FEM1C degrons. (A) The last 12 amino acids of KLHDC3, KLHDC2 and FEM1C substrates. Critical residues demonstrated by mutagenesis are colored. The minimal lengths of some degrons were mapped and are underlined. (B, C) Stability analysis of NS19 mutants. (D-G) GPS cells expressing substrates indicated at left with mutations labeled at the top were treated with shRNAs against various BC-box proteins and analyzed. To avoid the effect of upstream Arginines, GAPDH fused with 6-amino acid CTTs from KLHDC3 substrates were tested in (E). sh#1 was used for BC-box protein knockdown unless otherwise indicated. See also Figure S3.

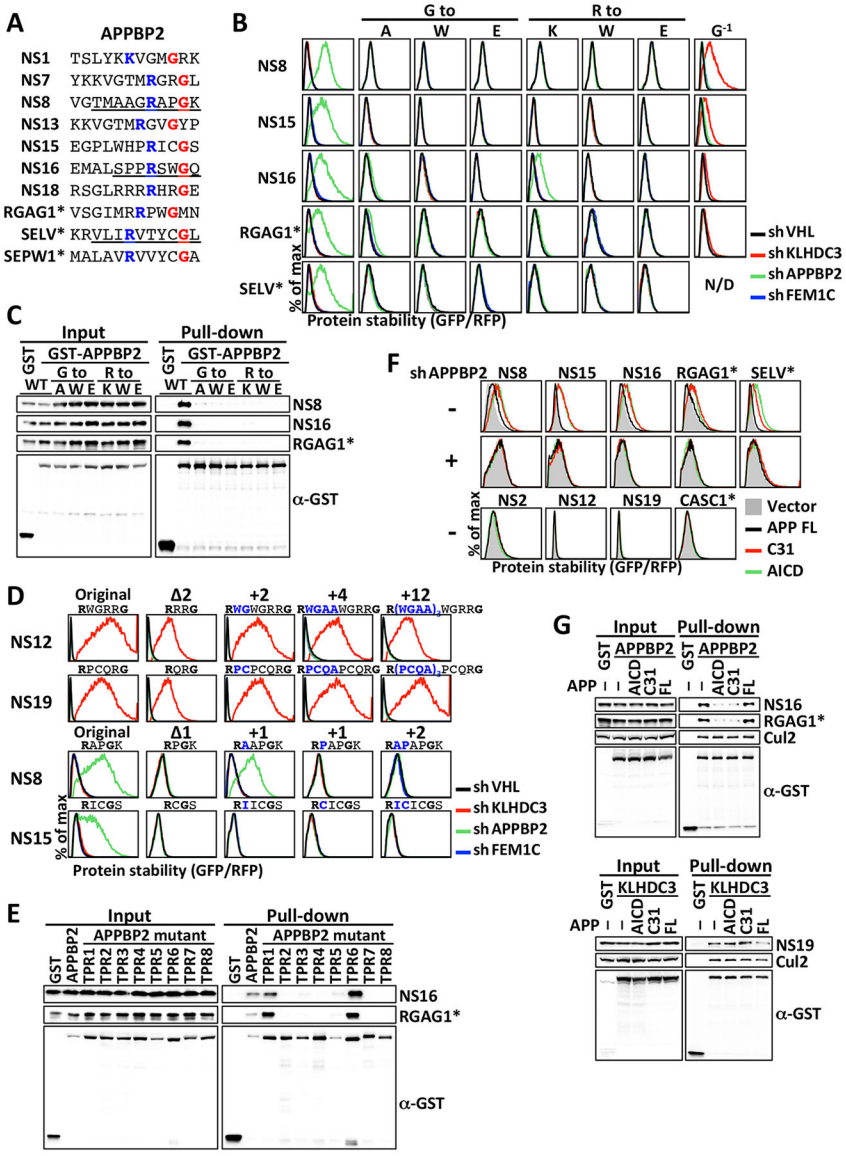

Figure 5.

Characterization of APPBP2-mediated degradation. (A) The C-terminal sequences of APPBP2 substrates. Critical residues are colored and the minimal length of APPBP2 degrons is underlined. (B) Mutagenesis analysis of APPBP2 substrates. G-1 represents mutants that lack amino acids downstream of the critical Glycine (see Fig. S3C for stability comparisons). (C) Physical interaction between GST-tagged APPBP2 and identified (WT) or mutant APPBP2 substrates. (D) Stability analysis of KLHDC3 and APPBP2 substrate mutants with amino acids deleted or added between the critical Arginine and Glycine. Added amino acids are labeled blue. (E) Physical binding between GST-tagged WT or mutant APPBP2 with GFP-tagged APPBP2 substrates (see Fig. S3D for APPBP2 mutation). (F) GPS assay for cells expressing CRL2 substrates treated with or without shRNA against APPBP2 and infected with viruses expressing full-length (FL) or processed APP. (G) GST

pull-down assay using cells expressing GST-tagged BC-box proteins and GFP-tagged CRL2 substrates with or without the presence of full-length or processed APP. See also Figure S3.

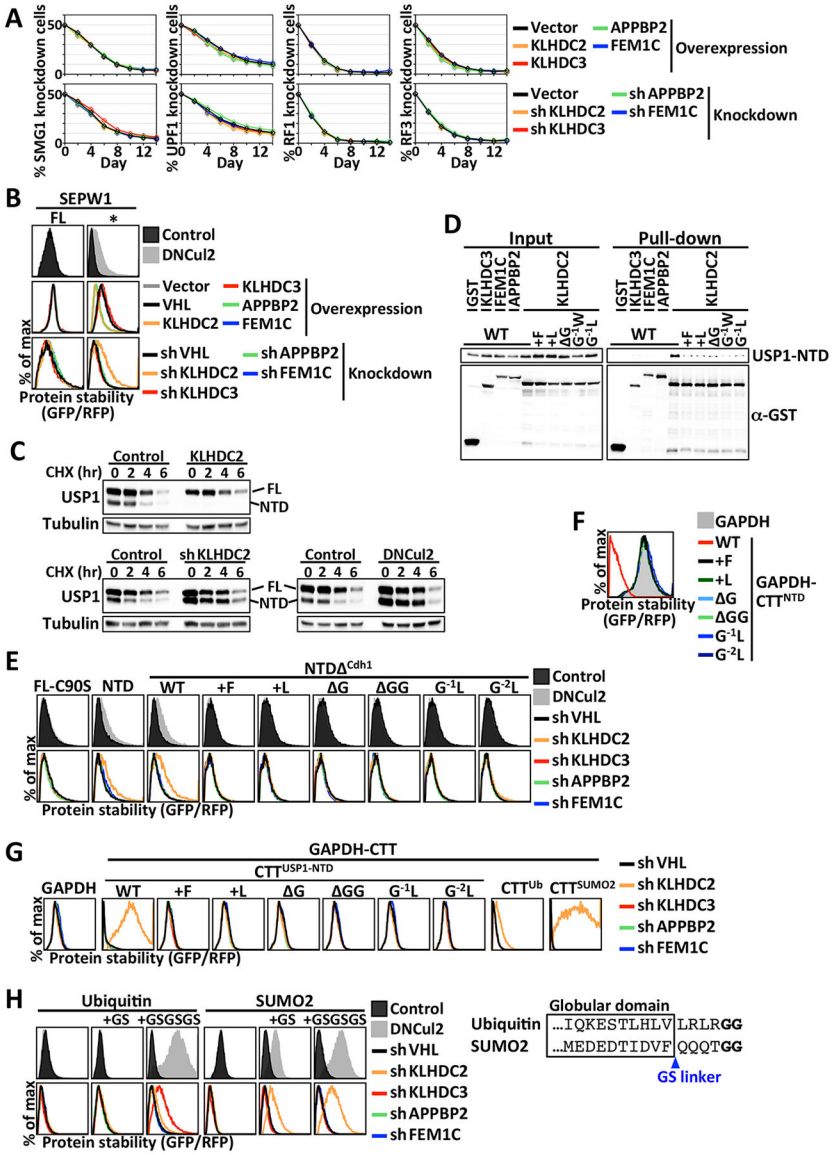

Figure 6.

Physiological functions of CRL2-mediated DesCEND. (A) Competition assay between cells treated and not treated with shRNAs against various proteins involved in NMD (SMG1 and UPF1) or translation termination (RF1 and RF3) under different genetic backgrounds of BC-box proteins (see Fig. S4A for experimental procedures and Fig. S4B for results using a different shRNA). (B) Stability analysis of truncated and full-length SEPW1. (C) CHX-chase analysis of endogenous USP1 using HEK293T cells infected with viruses expressing KLHDC2, shKLHDC2 or DNCul2. Full-length (FL) and the N-terminal domain (NTD) of USP1 are labeled. (D) Physical interaction between GST-tagged BC-box proteins and GFP-tagged wild-type (WT) or mutant USP1-NTD. (E) Stability analysis of various forms of USP1 in HEK293T cells. FL-C90S is a full-length USP1 mutant that is unable to perform autocleavage. NTDAcdh1 lacks the Cdh1 degron (a.a. 295–342). (F, G) Stability analysis of

GAPDH C-terminally-tagged with the CTT of USP1-NTD, ubiquitin or SUMO2. (H) Stability analysis of ubiquitin and SUMO2 with or without GS linkers. See also Figure S4.

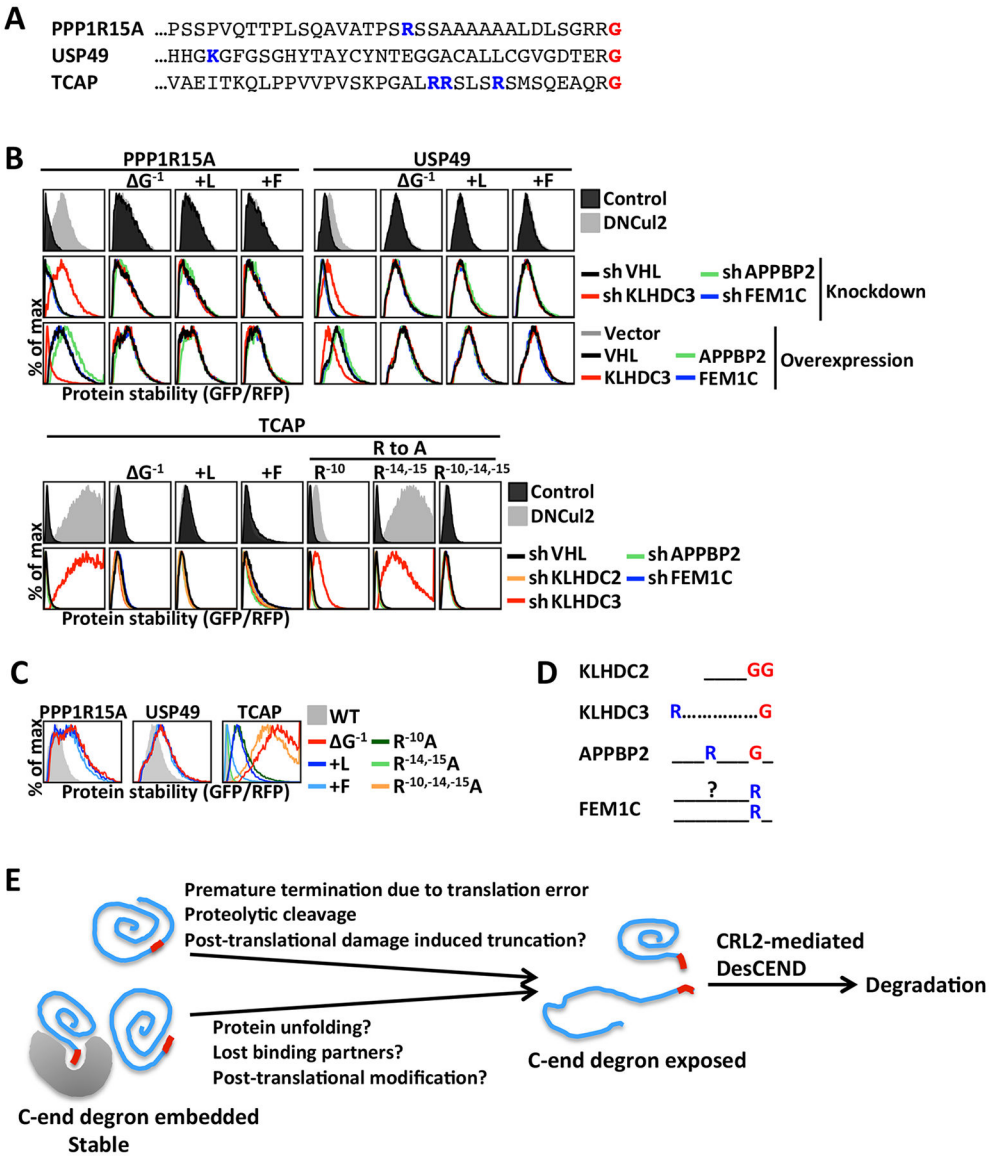

Figure 7.

DesCEND regulates full-length proteins with native C-end degrons. (A) The C-terminal sequences of indicated proteins. Critical residues in KLHDC3 degrons are colored. (B) Stability analysis of various forms of PPP1R15A, USP49 and TCAP. (C) Stability comparison among proteins in (B). (D) Characteristics of C-end degrons and their respective BC-box proteins. (E) Model of the physiological functions of DesCEND.