OXFORD GENETICS

GENETICS, 2021, 217(1), iyaa031
DOI: 10.1093/genetics/iyaa031
Advance Access Publication Date: 22 December 2020
Investigation

Protein quality control degron-containing substrates are differentially targeted in the cytoplasm and nucleus by ubiquitin ligases

Christopher M. Hickey [1,†,‡], Carolyn Breckel [1,‡], Mengwen Zhang,[2] William C. Theune,[3] and Mark Hochstrasser [1,4,*]

[1]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06511, USA
[2]Department of Chemistry, Yale University, New Haven, CT 06511, USA
[3]Department of Biology and Environmental Science, University of New Haven, West Haven, CT 06516, USA
[4]Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06511, USA
[†]Present address: Arvinas, Inc., 5 Science Park, New Haven, CT, USA.
[‡]These authors are co-first authors.

*Corresponding author: Department of Molecular Biophysics and Biochemistry, 266 Whitney Avenue, P.O. Box 208114, New Haven, CT, USA.
mark.hochstrasser@yale.edu

Abstract

Intracellular proteolysis by the ubiquitin–proteasome system regulates numerous processes and contributes to protein quality control (PQC) in all eukaryotes. Covalent attachment of ubiquitin to other proteins is specified by the many ubiquitin ligases (E3s) expressed in cells. Here we determine the E3s in *Saccharomyces cerevisiae* that function in degradation of proteins bearing various PQC degradation signals (degrons). The E3 Ubr1 can function redundantly with several E3s, including nuclear-localized San1, endoplasmic reticulum/nuclear membrane-embedded Doa10, and chromatin-associated Slx5/Slx8. Notably, multiple degrons are targeted by more ubiquitylation pathways if directed to the nucleus. Degrons initially assigned as exclusive substrates of Doa10 were targeted by Doa10, San1, and Ubr1 when directed to the nucleus. By contrast, very short hydrophobic degrons—typical targets of San1—are shown here to be targeted by Ubr1 and/or San1, but not Doa10. Thus, distinct types of PQC substrates are differentially recognized by the ubiquitin system in a compartment-specific manner. In human cells, a representative short hydrophobic degron appended to the C-terminus of GFP-reduced protein levels compared with GFP alone, consistent with a recent study that found numerous natural hydrophobic C-termini of human proteins can act as degrons. We also report results of bioinformatic analyses of potential human C-terminal degrons, which reveal that most peptide substrates of Cullin-RING ligases (CRLs) are of low hydrophobicity, consistent with previous data showing CRLs target degrons with specific sequences. These studies expand our understanding of PQC in yeast and human cells, including the distinct but overlapping PQC E3 substrate specificity of the cytoplasm and nucleus.

Keywords: ubiquitin; proteasome; degron; protein degradation; protein quality control

Introduction

Reversible, covalent attachment of the small protein ubiquitin to other cellular proteins, termed ubiquitylation, is central to eukaryotic biology (Hochstrasser 2009). Ubiquitylation requires three enzymes, called E1 (or ubiquitin-activating enzyme), E2 (or ubiquitin-conjugating enzyme), and E3 (or ubiquitin ligase), that act sequentially. All eukaryotes express multiple E2s and many E3s, each of which binds a set of substrate proteins. Functions of the ubiquitin–proteasome system (UPS) can be divided into two main categories: cellular regulation and protein quality control (PQC). In both cases, the prototypical role for ubiquitin is to reduce the concentrations of specific proteins by targeting them for degradation by the proteasome. In addition to the UPS, many proteasome-independent functions of ubiquitin are appreciated, most notably in signaling, DNA repair, and membrane trafficking and including some types of autophagy, which also contributes to PQC (Pohl and Dikic 2019).

Substrates of the UPS expose degradation elements important for their ubiquitylation and subsequent degradation by the proteasome. We have previously reviewed use of the terms degradation element and degradation signal (or degron; Hickey 2016). Here degron refers to a portion of a protein that is both necessary and sufficient to target an otherwise stable protein for degradation whereas degradation element is used very broadly to include any part of a short-lived protein that contributes to its degradation. Others use the term degron to describe the degradation element specifically recognized by an E3. For many UPS substrates, hydrophobic degradation elements are exposed in proteins that are improperly folded or assembled, and degradation of these proteins is important for PQC. For other UPS substrates, specific

Received: September 13, 2020. **Accepted:** December 7, 2020

degradation elements are created as another layer of cellular regulation. The most notable examples are "phospho-degrons" within proteins, which are often recognized by the F-box subunits of Skp1/Cullin/F-box (SCF) complexes, a type of Cullin-RING-ligase (CRL). Other CRLs recognize purely peptidic degradation elements (Koren *et al.* 2018; Rusnac *et al.* 2018). Some regulatory protein substrates of the UPS are constitutively short-lived, and transcription factors typify this group, often displaying degradation elements similar to those of PQC substrates in the absence of cofactor interaction (Geng *et al.* 2012; Hickey 2016). Constitutive degradation of transcription factors is one mechanism available to ensure rapid changes to specific transcriptional programs upon decreased synthesis of the factor. This is the case for the yeast transcription factor MATalpha2 (α2), the degradation of which ensures rapid yeast mating-type switching (Laney and Hochstrasser 2003; Haber 2012).

Our laboratory has identified and studied two distinct degrons in α2, called degron 1 (*Deg1*) and degron 2 (*Deg2*), depicted in Figure 1A (Hochstrasser and Varshavsky 1990; Hickey and Hochstrasser 2015). *Deg1* represents the N-terminal ~62 amino acids of α2 and proteins bearing *Deg1* at their N-terminus are targeted for degradation by the E3 Doa10, which operates with two E2s, Ubc6 and Ubc7 (Swanson *et al.* 2001). These E2s and Doa10 localize to the endoplasmic reticulum (ER)/nuclear envelope and participate in ER-associated degradation (ERAD) (Mehrtash and Hochstrasser 2019). As part of ERAD, Doa10 targets both integral membrane proteins and non-membrane (soluble) proteins—though these soluble proteins likely have some propensity to interact with membranes. Full-length α2 is also targeted for degradation by the heterodimeric E3 Slx5/Slx8 and the E2 Ubc4 (Chen *et al.* 1993; Xie *et al.* 2010). If the full-length α2 bears mutations in its *Deg2* region, it is almost exclusively targeted by the Doa10 pathway and not the Ubc4/Slx5/Slx8 pathway, consistent with the two degrons being nonoverlapping and functioning independently (Hickey and Hochstrasser 2015; Hickey *et al.* 2018). While cells lacking Slx5 or Slx8 slightly stabilize a *Deg2* fusion protein relative to wild-type (WT) cells, the *Deg2* fusion remains quite unstable in cells lacking Slx5/8 activity, suggesting other pathways still can target the chimeric protein for degradation (Hickey and Hochstrasser 2015).

Doa10 is known to target soluble proteins bearing regions that are both hydrophobic and α-helical (Gilon *et al.* 1998; Johnson *et al.* 1998; Ravid *et al.* 2006; Furth *et al.* 2011). A recent survey of yeast protein N-termini, removed from their normal context, revealed numerous cryptic N-terminal degrons targeted by Doa10 (Kats *et al.* 2018). These degrons are likely normally involved in intra-protein contacts or protein-protein interactions. Similar cryptic Doa10 degrons have been found throughout yeast proteins, including proteins of known structure, solidifying a model that the hydrophobic face of an amphipathic helix is the typical Doa10 substrate (Johnson *et al.* 1998; Geffen *et al.* 2016). Furthermore, several "artificial" degrons for Doa10 have been described, often resulting from out-of-frame translation products (Gilon *et al.* 1998; Geffen *et al.* 2016; Maurer *et al.* 2016). Among the best characterized degrons for Doa10 is an artificial degron called CL1, first described in yeast and later studied in human cells; CL1 was recently found to be targeted by the human ortholog of Doa10—MARCH6—among other E3s (Gilon *et al.* 1998; Bence *et al.* 2001; Stefanovic-Barrett *et al.* 2018; Leto *et al.* 2019). CL1 is an example of a C-terminal degron, which have been studied less extensively than N-terminal degrons (Varshavsky 2012; Kats *et al.* 2018). However, a recent survey of peptides derived from human proteins as C-terminal extensions to GFP revealed numerous examples of C-terminal degrons, many of which are likely cryptic (Koren *et al.* 2018).

In addition to Doa10, the E3s San1 and Ubr1 have been reported to target proteins bearing exposed hydrophobicity in yeast (Heck *et al.* 2010; Prasad *et al.* 2010; Fredrickson *et al.* 2011). While some short-lived proteins appear to be substrates of only one of these two E3s, the full stabilization of several soluble PQC substrates depends on the simultaneous absence of both San1 and Ubr1 (Heck *et al.* 2010; Prasad *et al.* 2010; Khosrow-Khavar *et al.* 2012; Guerriero *et al.* 2013; Summers *et al.* 2013; Samant *et al.* 2018). San1 is a nuclear localized E3 but can function in degrading cytoplasmic PQC substrates if those proteins are transported to the nucleus (Gardner *et al.* 2005; Enam *et al.* 2018). Ubr1 has also been reported to target both cytosolic and nuclear UPS substrates, functioning in both PQC and the N-end rule pathway (Heck *et al.* 2010; Varshavsky 2012; Prasad *et al.* 2018).

Here we have studied multiple degrons, some from naturally short-lived proteins and others artificial—though resembling natural, often cryptic, degrons. We also report our latest collection of yeast gene deletion strains designed to facilitate analysis of the ubiquitin system. Using this collection, we show that (cryptic) degrons derived from both α2 and the similar transcription factor MATa1 are partially targeted by Ubr1. We show that hydrophobic degrons recognized by Doa10 are also targeted by Ubr1, in concert with San1, but only if the substrate is directed to the nucleus. This includes the previously studied CL1. By contrast, several very short and extremely hydrophobic C-end degrons are targeted by Ubr1 and San1 without contributions from Doa10. We show that one of these degrons also functions in human cells, consistent with hydrophobicity being a defining feature for many of the recently reported C-end degrons (Koren *et al.* 2018).

Materials and methods

An updated targeted ubiquitin system collection for ubiquitylation pathway identification

In previous efforts to define the ubiquitylation machinery for specific UPS substrates, we compiled a collection of yeast strains, each deleted for a specific component of the ubiquitin system, in a single 96-well plate and called this collection the targeted ubiquitin system (TUS) library (Ravid and Hochstrasser 2007; Xie *et al.* 2010). The strains of the original TUS library, assembled in 2005, were from the much larger yeast deletion collection, with the genes/proteins chosen based on then known or proposed ubiquitin system components (Giaever *et al.* 2002; Xie *et al.* 2010). Due to the nature of the ubiquitin system, the library is heavily focused on strains deleted for genes encoding E2s and E3s. The original collection consisted largely of genes encoding proteins containing the domains UBC (E2), HECT (E3), RING (E3), or RING-like.

To create an updated TUS library in 2011, the *Saccharomyces Genome Database* (SGD) was again searched for genes coding for proteins involved in the ubiquitin system. To identify genes coding for E3s or E3 components, the terms "ubiquitin," "RING," "HECT," "cullin," and "F-box" were searched. Upon examining this list in relation to the original library, we decided to remove certain strains from the original collection, most notably strains deleted for genes encoding proteins with domains that are RING-like (such as the PHD-finger domain and FYVE domain), which are unlikely to be bona fide RING domains with ubiquitin ligase activity. We then added strains deleted for ubiquitin system components not originally included, most notably genes for non-essential F-box proteins, substrate recognition subunits of the

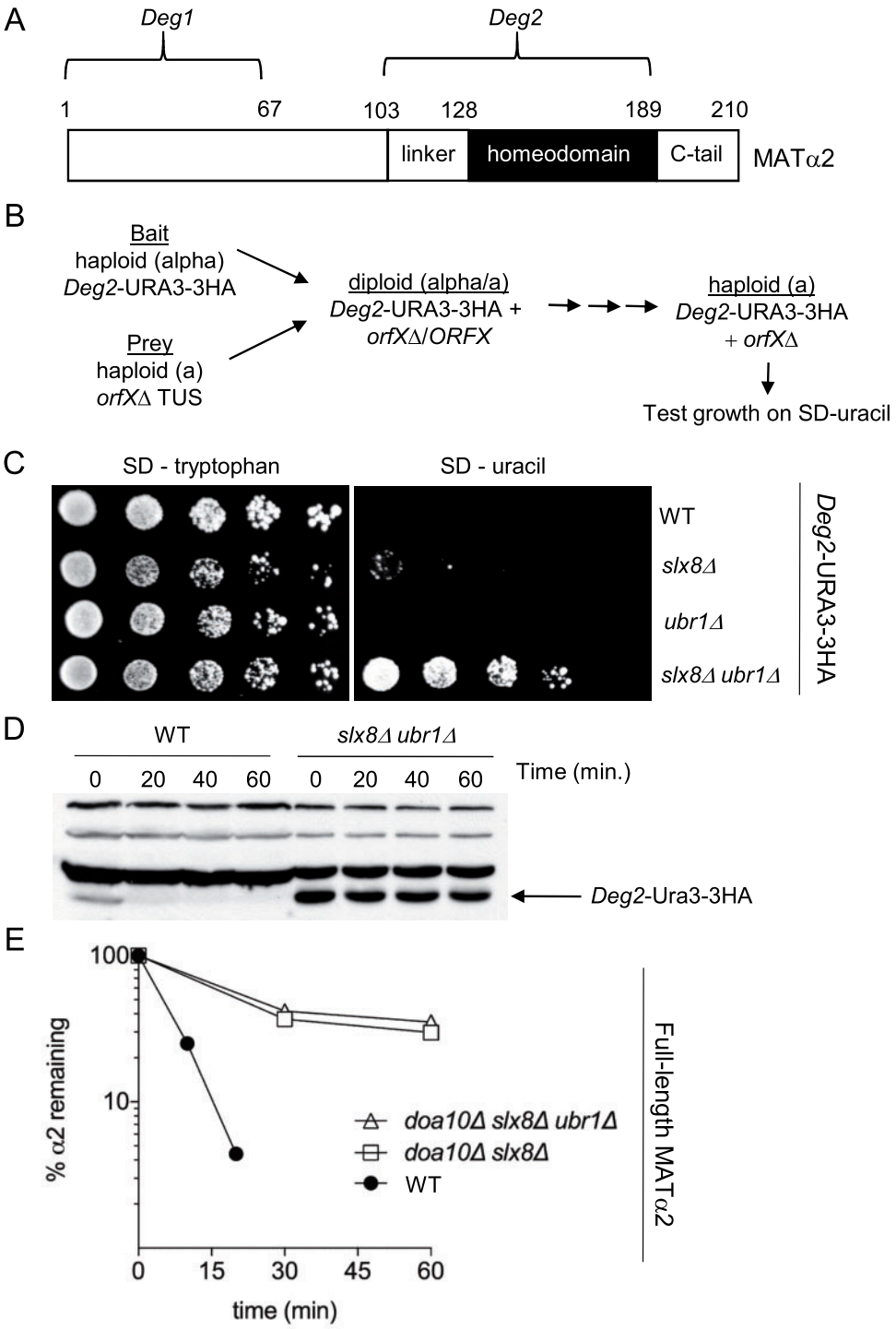

Figure 1 The *Deg2* degron when excised from MATα2 is targeted by two E3s, Slx5/Slx8 and Ubr1. (A) Schematic of the 210 amino acid MATα2 protein, showing the two previously characterized degrons. (B) Schematic for mating array-based yeast ubiquitylation machinery identification (YUMI). (C) Growth of the indicated yeast strains on media lacking either tryptophan or uracil, as indicated. Growth on media lacking uracil depends on the stability of the *Deg2*-Ura3-3HA reporter protein being expressed from a *TRP1*-based plasmid. (D) A cycloheximide-chase immunoblot analysis measuring degradation of the *Deg2*-Ura3-3HA protein in the indicated strains. The arrow indicates the fusion protein of interest. Other bands (anti-HA-reactive yeast proteins) serve as loading controls. To avoid excessive signal of the fusion protein, fivefold less extract was loaded for the *slx8Δ ubr1Δ* series. (E) Quantification of full-length MATα2 degradation rates, as determined by radioactive pulse-chase analysis using an anti-α2 antibody. Data shown are from a single experiment; a biological replicate yielded nearly identical results. Congenic strains used to generate data for this figure were: MHY501 (WT), MHY3716 (*slx8Δ*), MHY474 (*ubr1Δ*), MHY9394 (*slx8Δ ubr1Δ*), MHY3718 (*doa10Δ slx8Δ*), and MHY9396 (*doa10Δ slx8Δ ubr1Δ*).

SCF E3s. We also added three strains we constructed to complete the TUS library: *ubr2Δ*, since the strain present in the original yeast deletion library was dubious; *ubc6Δ*, which was absent from the yeast deletion library used to construct the TUS; and *ubc1Δ*, which we recovered after it had presumably "evolved" to be viable (Liu *et al.* 2015).

The haploid *ubc1Δ* strain of our library was generated by sporulating a *UBC1/ubc1Δ* diploid strain in the BY4743 background

(Dharmacon). To generate haploid *ubc6Δ* and *ubr2Δ* strains, diploid strain BY4743 was transformed with DNA that was PCR amplified from plasmid pFA6a-kanMX6 designed to replace the appropriate gene with the G418-resistance cassette (Longtine *et al.* 1998; Gietz and Woods 2002). The resulting diploids were sporulated and haploids of the desired genotype for the TUS were isolated. All other strains included in the updated TUS library were retrieved from the previous library edition or the mating-type a-haploid yeast deletion library (Dharmacon); the strains were grown in YPD to saturation, mixed with sterile 75% glycerol to 15% glycerol final, and added to the appropriate well of a 96-well plate. The plate was sealed with foil tape and stored at −80°C.

While making subtractions and additions to the library, we also rearranged the strains in the 96-well plate according to protein function (Supplementary Figure S1A and Supplementary Table S3). We call our most recent collection of strains TUS2.1 (Supplementary Table S3), which differs only slightly from TUS2.0, which was also used in the current study and by others (Supplementary Table S4). The list of strains in our updated TUS library was later cross-referenced to a published list of *Saccharomyces cerevisiae* ubiquitin system components (Finley *et al.* 2012). Ubiquitin system components not covered by TUS2.1, which are largely encoded by essential genes, are also listed in Supplementary Table S3.

Plasmids and molecular biology methods

All plasmids used in this study are listed in Supplementary Table S1. To construct plasmid p415MET25-MATa1(cDNA)-URA3, MATa1 cDNA was PCR amplified from a pRS316GAL1-cDNA library and cloned into the SpeI and HindIII sites of p415MET25-X-URA3-3HA (Liu *et al.* 1992; Hickey and Hochstrasser 2015). A similar plasmid with the *TEF1* promoter was constructed by subcloning (Mumberg *et al.* 1995). A screen for nucleotide changes to the *a1* cDNA within this plasmid that led to increased levels of a1-Ura3-3HA revealed several silent mutations, which we presume was due to enhanced mRNA stability (data not shown). Two of these silent mutations (in codons for Y34 and K106) were incorporated into certain MATa1(cDNA)-URA3-3HA constructs to achieve protein levels in an optimal range for growth assays. Constructs expressing fragments of *MATa1* were generated by PCR amplifying the corresponding *a1* DNA from plasmid p415MET25-MATa1(cDNA)-URA3 or p415MET25-MATa1(cDNA; Y34Y and K106K)-URA3. To generate a plasmid for integration of the *Deg2* reporter into the yeast genome, DNA corresponding to (MET25 promoter-Deg2-URA3-3HA-CYC1 transcriptional terminator) was cut from plasmid p415MET25-Deg2-URA3-3HA with SacI and ClaI and cloned into the same sites in pAG32 (which harbors a hygromycin B-resistance cassette; Goldstein and McCusker 1999; Hickey and Hochstrasser 2015).

To generate plasmids for expression of Ura3-HA fusions in yeast, DNA for Ura3-HA was amplified by PCR from plasmid pOC-Ura3-HA-CL1 and cloned into the XbaI and BamHI sites of p415MET25 (Mumberg *et al.* 1994; Gilon *et al.* 1998). Degrons were then cloned into the BamHI and XhoI sites of this plasmid. To make plasmids that express nuclear localization signal (NLS)-GFP-Ura3-HA-degron fusions, first DNA for NLS-GFP was amplified by PCR from the pFA6a-GFP(S65T)-kanMX6 vector, with DNA for the SV40 T-Ag NLS (amino acid sequence PKKKRKV) added to one of the primers (Longtine *et al.* 1998). After cloning this DNA into an intermediate vector, it was subcloned into the SpeI and HindII sites of a p415MET25 vector already containing URA3-HA-CL1, which was amplified from pOC-URA3-HA-CL1 and cloned into the HindIII and XhoI sites of p415MET25 (Gilon *et al.* 1998). Similar plasmids expressing proteins with a nonfunctional mutant NLS (m.nls) were generated by site-directed mutagenesis to generate the amino acid sequence PKTKRKV. To make plasmids that express proteins with degrons other than CL1, the above plasmids were digested with BamHI (just upstream of the DNA for CL1) and XhoI and duplex DNA generated by annealing synthetic oligonucleotides was cloned into these sites. Plasmids expressing reporter proteins with "no degron" were made by cloning in short fragments of synthetic duplex DNA with two tandem STOP codons just after the BamHI site. This plasmid and all degron-expressing constructs were sequence verified.

To make constructs for expression of NLS-GFP-HA-degron or m.nls-GFP-HA-degron in human cells, DNA was amplified from the plasmid pEGFP-C with the aforementioned SV40 NLS or m.nls coded in one of the primers. These DNA fragments were cloned into the HindIII and BamHI sites of pCDNA-FRT/TO (Invitrogen) with no STOP codon added. DNA fragments generated by annealing synthetic oligonucleotides and coding for various degrons were then cloned into these plasmids using the BamHI and XhoI sites. Once again, plasmids expressing reporter proteins with "no degron" were made by cloning in short fragments of synthetic duplex DNA with STOP codons just after the BamHI site. All constructs were sequence verified.

Construction of yeast strains other than those of the TUS

All yeast strains used in this study (other than those of the TUS libraries) are listed in Supplementary Table S2. The mating array-compatible strain MHY2994 was transformed with plasmid pAG32-MET25promoter-DEG2-URA3-3HA-CYC1terminator that had been linearized by digestion with BsrGI, which cleaves the plasmid DNA only once within the CYC1 terminator (Tong *et al.* 2001; Gietz and Woods 2002). The resulting hygromycin B-resistant strain, MHY9368, was tested for slow growth on media lacking uracil. Expression and degradation of *Deg2*-Ura3-3HA in MHY9368 was confirmed by cycloheximide-chase analysis. To generate a strain lacking *SLX8* and carrying the integrated pAG32-MET25promoter-DEG2-URA3-3HA-CYC1terminator plasmid (MHY9369), MHY9368 cells were transformed with a nourseothricin-resistance cassette PCR amplified from pAG25 to delete *SLX8* (Goldstein and McCusker 1999).

A strain with the *UBR1* locus replaced by a nourseothricin-resistance cassette that had been PCR amplified from plasmid pAG25 was generated by replacing the *LEU2*-based cassette in MHY474 (Bartel *et al.* 1990). A *UBC4/ubc4Δ UBC5/ubc5Δ UBR1/ubr1Δ* diploid strain was generated by mating a *ubc5Δ* haploid with a *ubc4Δ ubr1Δ* haploid. This diploid was transformed with linearized versions of p306-UBC4 or p306-ubc4-N78S, as in Stoll *et al.* (2011). Resulting clones were then sporulated and tetrads were dissected, with 2:2 segregation of growth on media lacking uracil to obtain strains MHY9550, MHY9552, and MHY9554. To obtain a strain deleted for *ASI1*, WT diploid yeast MHY606 was transformed with a cassette that had been PCR-amplified from pFa6a-kanMX4 (Longtine *et al.* 1998). Resulting clones were then sporulated, and tetrads were dissected, with 2:2 segregation of growth on G418-containing medium. To obtain a strain deleted for *SAN1*, MHY606 was transformed with a PCR-amplified cassette from pUG6 (Guldener *et al.* 1996). Multiple clones were then sporulated, and tetrads were dissected, with 2:2 segregation of growth on G418 medium. A *HRD1/hrd1Δ DOA10/doa10Δ* diploid strain, MHY6192, was generated by transforming MHY606 with cassettes for deletion of these two genes, with the cassette for

deletion of *DOA10* originally from plasmid pAG32 and the cassette for disruption of *HRD1* as used in the yeast deletion collection (Giaever *et al.* 2002). MHY6192 was then sporulated, and tetrads were dissected to obtain the various single and double deletion strains.

All other strains were made from existing strains by standard yeast mating, sporulation, tetrad dissection, and spore analysis.

Yeast ubiquitin machinery identification analysis

The yeast ubiquitin machinery identification (YUMI) analysis protocol was derived from the synthetic genetic assay (SGA) protocol (Tong and Boone 2006). Chemicals, media, and materials were as in the SGA protocol except all SD/MSG plates containing antibiotics were adjusted to pH 6.3 with sodium hydroxide, which we found necessary for antibiotic solubility. Yeast pinning and streaking were performed manually. The TUS2.0 library was recovered from a 96-well plate stored at −80°C using a pin tool and spotted on YPD agar containing 250 μg/ml G418. This plate was incubated for 2–3 days at 24°C just before mating to the YUMI bait strain. The YUMI bait strain (MHY9368) was grown in YPD broth to saturation, and approximately 5 μl of this culture were added to each of 96 spots on a YPD agar plate using a pin tool. This plate was incubated overnight at 30°C. The strains of the freshly prepared TUS2.0 were then mixed with the arrayed MHY9368 strain on the YPD agar plate using the pin tool and incubated at 24°C for 24 h. The resulting yeast were pinned to a YPD plate containing 250 μg/ml G418 and 300 μg/ml Hygromycin B and plates were incubated at 30°C for 2 days to select for diploids. Diploids were transferred to a sporulation plate using the pin tool and incubated for 5 days at 24°C. Haploid type a cells were selected by pinning yeast on SD-his-lys-arg plates containing canavanine and thialysine and allowing them to grow for 1 day at 24°C. This a-haploid selection step was repeated and the resulting yeast were then transferred to a-haploid selection plates containing 250 μg/ml G418 with incubation at 24°C for 2 days. The resulting yeast were then transferred to a-haploid selection plates containing 250 μg/ml G418 and 300 μg/ml Hygromycin B with incubation at 24°C for 2 days. The resulting yeast, which carry the chromosomally integrated *Deg2-URA3* reporter and a specific gene deletion, were streaked on SD-uracil plates so that 8 strains (corresponding to a single column of the 96-well TUS library) were grown per plate. These plates were incubated at 24°C for several days, with relative growth monitored daily.

For the secondary screen in which the YUMI bait strain was *slx8Δ*, the above procedure was followed except only select strains from TUS2.0 were mated to MHY9369 (see Supplementary Figure S2A). Also different from the above procedure, yeast from a-haploid selection plates containing two antibiotics were struck out twice sequentially on a-haploid selection plates containing three antibiotics (250 μg/ml G418, 100 μg/ml Clonat, and 300 μg/ml Hygromycin B) before analysis on medium lacking uracil (single colonies of yeast were picked from a-haploid selection plates containing the three antibiotics and patched on another plate of the same composition). Two distinct colonies per genotype were grown in liquid media of the same composition as the a-haploid selection plates containing the three antibiotics (except without agar) and tested for growth on SD-uracil after optical density measurement and serial dilution. The two clones yielded nearly identical results, with results from one set shown in Supplementary Figure S2B.

Human cell culture and microscopy

HeLa Flp-In T-REx cells (Thermo) were grown in DMEM (Gibco) supplemented with 10% fetal bovine serum (Atlanta Biologicals or Gibco). Cells were transfected with plasmid DNA (described above) using X-tremeGENE 9 transfection reagent (Sigma). Subsequent cell treatments were begun 24 h after transfection.

To measure levels of GFP protein reporters, one of two methods was used to create cell extracts. For immunoblot experiments shown in Figure 10, B–D, cell monolayers in wells of a 6-well plate were washed with PBS and then treated with Trypsin/EDTA. After cell detachment, trypsin was neutralized by transferring cells to media in 15-ml conical tubes. Cells were washed in 1 ml PBS with transfer to 1.5-ml tubes via centrifugation at 200 × *g* for 5 min. Cells were harvested by centrifugation again and resuspended in 0.15 ml of HeLa lysis buffer: 50 mM Tris-Cl, pH 7.5, 50 mM NaCl, 0.2% Triton X-100, 2 mM EDTA, 1 mM PMSF, and 1x protease inhibitor cocktail (Roche tablet). Extracts were then centrifuged at 21,000 × *g* for 20 min in a tabletop microcentrifuge to pellet insoluble material. Protein concentrations of supernatants were measured using the BCA assay (Thermo) and extracts were separated by SDS-PAGE for immunoblot analysis. For the experiment shown in Figure 10F, cell monolayers in wells of a 6-well plate were washed in PBS and then treated with 0.150 ml of RIPA buffer (Thermo) containing 1× protease inhibitor cocktail (Roche tablet). After a 15 min incubation on ice with occasional tilting of the plates, extracts were transferred to 1.5-ml tubes. Subsequent steps were as described for the method involving cell detachment by trypsin treatment.

For microscopy of HeLa cells, cells were grown at low confluency on coverslips in 12-well plates. Cells were transfected, incubated for 24 h, and then treated with 2 μg/ml doxycycline for 6 h. To process coverslips for microscopy, media was aspirated and cell monolayers were washed three times in Tris-buffered saline (TBS). Cells were fixed for 15 min in 3.7% formaldehyde in TBS, washed 4 times in TBS, permeabilized for 20 min with 0.2% Triton X-100 in TBS, and blocked for 1 h in TBS containing 5% bovine serum albumin (BSA). Coverslips were incubated with antibody to GFP (Roche 1181446000; 1:1:500) and Hoechst 33342 stain (Invitrogen; 1:20,000) in TBS containing 5% BSA and 0.1% Triton X-100. After washing four times in TBS, coverslips were incubated for 1 h with anti-mouse Alexa Fluor 488-conjugated secondary antibody in TBS containing 5% BSA and 0.1% Triton X-100. Cells were washed four times in TBS and then coverslips were removed and placed face down on slides containing a drop of Fluoromount-G (Southern Biotech). Nail polish was added to edges of the coverslips to affix the coverslips to the slides. After the slides dried overnight at room temperature in the dark, images were captured at room temperature on a Zeiss Axioskop microscope (Carl Zeiss Microimaging) fitted with a Zeiss Axiocam camera operated by Axiovision software.

Bioinformatic analysis of human C-end degrons

Peptides listed in Supplementary Table S7A of Koren *et al.* (2018) were uploaded to the "peptide synthesis and proteotypic peptide analyzing tool" (ThermoFisher Scientific) to calculate GRAVY scores, hydrophobicity scores, and isoelectric point (pI). These values were added to an Excel file and various sheets were added to this file based on the criteria indicated in the Excel file (see Supplementary Table S5 of the current study).

Other methods and materials

Microscopy of yeast cells and image processing were performed as previously reported (Hickey *et al.* 2018). Pulse-chase experiments were carried out as previously described (Hickey *et al.* 2018). Cycloheximide-chase methods have also been described previously (Hickey and Hochstrasser 2015). Immunoblot data was either collected on film or with a G: Box system (Syngene, Frederick, MD). For quantification of cycloheximide-chase data, images from scanned films were processed using GeneTools software (Syngene). Antibodies to HA (clone HA-7) and glucose-6-phosphate dehydrogenase (A9521) were from Sigma. Antibody to beta-actin was from Abcam (ab8226) and antibody to GFP was from Clontech (JL-8).

Yeast growth assays were performed as previously described, except the most concentrated spots (far left) were derived from yeast at $OD_{600} = 1.0$, followed by sixfold dilution in series (Watts *et al.* 2015). All growth assays shown represent data obtained from at least two independent transformants (colonies).

Data and reagent availability

Strains and plasmids are available upon request. Supplementary Table S5 contains data on human C-terminal 23-mers. All other supplemental materials are contained within a single PDF file.

Supplemental material is available at figshare DOI: https://doi.org/10.25386/genetics.13273286.

Results

The excised Deg2 degron from MATα2 is targeted by Slx5/8 and Ubr1

To screen the TUS2.0 library against a *Deg2*-Ura3 fusion, we integrated DNA coding for the *Deg2* degron fused to *URA3* (also bearing sequences for triple hemagglutinin and hexahistidine tags; referred to henceforth as 3HA) into a yeast strain designed for SGA analyses (Tong *et al.* 2001). Unlike SGA, our goal was not to analyze synthetic genetic interactions but instead to use this strain for what we call YUMI—for yeast ubiquitylation machinery identification (Figure 1B). Our YUMI bait strain was mated to the TUS library (prey strains), diploids were selected and sporulated, and haploids lacking a specific gene and carrying the reporter were selected. These strains were then struck on solid media lacking uracil to estimate relative levels of the *Deg2*-Ura3-3HA protein.

YUMI analysis with *Deg2*-Ura3-3HA yielded several strains that clearly grew better on media lacking uracil than did the WT strain (Supplementary Figure S1, B and C), which could reflect a decrease in the degradation rate of the *Deg2*-Ura3-3HA reporter; however, their growth was modest compared with what would be expected for a fully stabilized reporter, as observed with certain mutant forms of *Deg2*-Ura3-3HA (Hickey and Hochstrasser 2015). Among hits in the screen were gene deletions known to cause general defects in the UPS (*doa1*Δ, *doa4*Δ, or *ubx1*Δ), including two genes found in the original degradation of alpha2 (*doa*) screen employing a *Deg1* reporter (Hochstrasser and Varshavsky 1990). We chose not to study these hits further. Two of the other hits were strains deleted for *UBC4* or *SLX8*, which was not surprising given the known role of these factors in targeting *Deg2* and α2 (Hickey and Hochstrasser 2015). The remaining hits were *dia2*Δ, *ubc2*Δ, and *ubr1*Δ. Notably, Ubc2 (Rad6) is the E2 that functions with the E3 Ubr1.

Since no single gene deletion led to robust growth based on our reporter, we reasoned that multiple E3s need to be deleted for strong reporter stabilization. We therefore created a bait strain deleted for *SLX8* and carrying the *Deg2*-Ura3-3HA reporter. This strain was mated to all hits from the YUMI analysis (along with some marginal hits, see Supplementary Figure S2), clones carrying the reporter and also lacking the two relevant genes were isolated, and these were evaluated for growth on media lacking uracil. Two clear hits emerged from this analysis: *slx8*Δ *ubr1*Δ and *slx8*Δ *dia2*Δ (Supplementary Figure 2B). Deletion of *DIA2* alone, which codes for an F-box protein, increased growth with multiple Ura3-based short-lived protein reporters (data not shown). Based on published work, we hypothesize that *dia2*Δ yeast have a general alteration in transcriptional regulation compared to WT yeast (in the case of our experimental design, de-repression of the *MET25* promoter) rather than a general UPS defect (Andress *et al.* 2011). Because we were interested specifically in *Deg2* degradation, we focused on Ubr1.

Next, we created an independent *slx8*Δ *ubr1*Δ strain and showed that growth based on a plasmid expressing a *Deg2*-Ura3-3HA reporter protein was indeed much more rapid in *slx8*Δ *ubr1*Δ cells than *slx8*Δ or *ubr1*Δ single mutants (Figure 1C). Increased abundance and metabolic stabilization of the reporter in *slx8*Δ *ubr1*Δ cells were confirmed by cycloheximide-chase analysis (Figure 1D). Thus, Ubr1 has a role in targeting *Deg2*-Ura3-3HA for degradation. However, when degradation of full-length, unmodified α2 was analyzed by radioactive pulse-chase, the effects of deleting *UBR1* were minimal (Figure 1E). We also observed no effect of deleting *UBR1* in growth-based assays with full-length α2-Ura3-3HA, which has the same ubiquitylation machinery requirements as untagged α2 (data not shown). These data, combined with earlier work showing neither Ubc2 nor Ubr1 had a role in degradation of α2 (Chen *et al.* 1993), suggest Ubr1 targets an isolated *Deg2* degron but not *Deg2* within its normal context in α2.

The MATa1 protein is a substrate of multiple ubiquitin ligases, including Ubr1

Among the most short-lived proteins in *S. cerevisiae*, with half-lives less than 15 min, are the three transcription factors regulating yeast mating type: α2, MATa1 (a1), and MATα1 (Hochstrasser and Varshavsky 1990; Johnson *et al.* 1998; Nixon *et al.* 2010). Of these three proteins, the least is known about proteolysis of a1 (depicted in Figure 2A). We therefore set out to identify pathways that target a1 for degradation using the TUS library. We began by creating constructs expressing a1-Ura3-3HA. Rather than screen the entire TUS library with the a1-Ura3-3HA reporter, we decided to initially screen the E2 gene deletion subset of TUS2.0. However, none of the TUS strains deleted for a single E2 showed significant stabilization of the reporter by either growth or cycloheximide-chase assay (data not shown). We therefore decided to test whether a1-Ura3-3HA was a substrate of the UPS in general. In a temperature-sensitive strain with a mutant proteasome subunit (*cim3-1*), a1-Ura3-3HA was very stable at the restrictive temperature (Figure 2B; Ghislain *et al.* 1993). The a1-Ura3-3HA protein was also stabilized in a *uba1-204* (E1) temperature-sensitive mutant when assayed at the restrictive temperature (data not shown; Ghaboosi and Deshaies 2007). Therefore, a1-Ura3-3HA is a UPS substrate.

We next tested whether fragments of a1 could act as independent degron sequences, perhaps each with fewer degradation pathways. Indeed, both the N-terminal half of a1 (amino acids 1–68, lacking the homeodomain) and the C-terminal half of a1 (amino acids 69–126, the a1 homeodomain) yielded short-lived proteins as N-terminal fusions to Ura3-3HA (Figure 2 and data

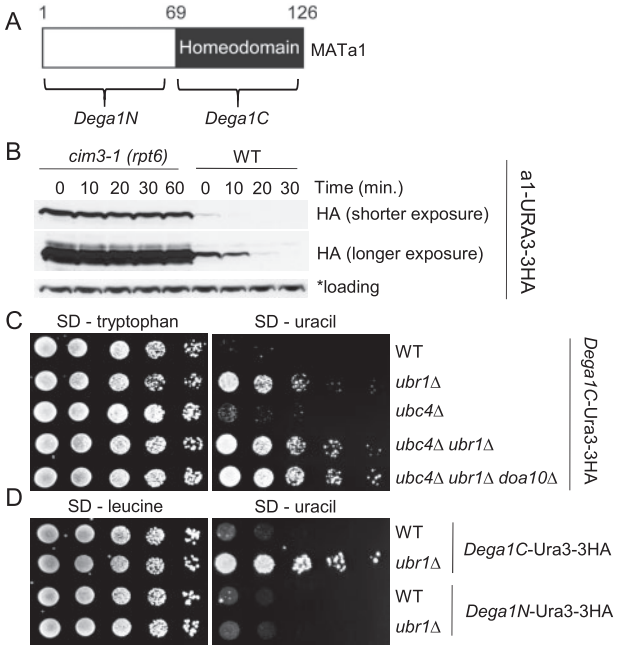

Figure 2 The naturally short-lived MATa1 protein bears multiple degrons and is targeted by several ubiquitin ligases, including Ubr1. (A) Schematic for MATa1 and its degrons. (B) MATa1-URA3-3HA is a proteasome substrate based on cycloheximide-chase analysis in MHY4464 (*cim3-1*) and MHY2836 (congenic WT). Cells were shifted from 24°C to 37°C 60 min prior to addition of cycloheximide. An anti-HA cross-reactive band (marked with an asterisk) served as a loading control. (C) The C-terminal half of MATa1 acts as a degron targeted primarily by Ubr1. Growth assays used media lacking either tryptophan or uracil for strains bearing a *TRP1*-based plasmid expressing *Dega1C*-Ura3-3HA. Strains used were: MHY500 (WT), MHY9380 (*ubr1Δ*), MHY513 (*ubc4Δ*), MHY9496 (*ubc4Δ ubr1Δ*), and MHY9384 (*ubc4Δ ubr1Δ doa10Δ*). (D) Degradation of a reporter with the N-terminal half of MATa1 is not inhibited in *ubr1Δ* cells. Growth assays were performed on media lacking leucine or uracil in strains bearing *LEU2*-based plasmids expressing the indicated proteins. The *ubr1Δ* strain used was from the TUS library and the corresponding WT strain was BY4741.

not shown). We call these degrons *Dega1N* and *Dega1C*, respectively. Given that *Dega1C* contains a homeodomain, as does the *Deg2* sequence from α2, we decided to test whether the pathways that target *Deg2* might also have a role in *Dega1C* degradation. In fact, growth assays suggested that *Dega1C*-Ura3-3HA is mainly a Ubr1 substrate, with additional targeting by a pathway involving the E2 Ubc4 (Figure 2C). By contrast, *Dega1N*-Ura3-3HA was not stabilized by loss of Ubr1 based on growth rates (Figure 2D). Employing growth assays, we also did not observe apparent stabilization of *Dega1N*-Ura3-3HA in any of the individual E2 deletion strains of TUS2.0 (data not shown). Consistent with a lack of *Dega1N*-Ura3-3HA stabilization in *ubc7Δ* cells, the protein was not stabilized in *doa10Δ* cells, as tested via cycloheximide-chase (Supplementary Figure S3A). We also subjected *Dega1N*-Ura3-3HA to YUMI analysis. The two best hits in this analysis were the *ubx1Δ* and *slx8Δ* strains (data not shown). However, in neither strain was growth particularly robust. Therefore, we presume that *Dega1N*-Ura3-3HA is recognized by multiple ubiquitylation pathways, potentially one involving Slx5/Slx8—which is known to operate with the E2 Ubc4/5 (Uzunova *et al.* 2007; Xie *et al.* 2010).

The E2s Ubc4 and Ubc5 are 93% identical and mostly functionally redundant, as strains lacking either protein have mild (*ubc4Δ*) or undetectable (*ubc5Δ*) growth phenotypes but strains lacking both Ubc4 and Ubc5 exhibit severe growth defects or are inviable depending on the strain background (Seufert and Jentsch 1990; Stoll *et al.* 2011). While some UPS substrates are significantly stabilized when yeast lack Ubc4 but retain Ubc5 (an example being α2), other proteins are not stabilized unless both Ubc4 and Ubc5 are missing (Nixon *et al.* 2010). The extremely poor growth of *ubc4Δ ubc5Δ* yeast has been hypothesized to be in large part due to reduced function of the essential HECT E3 Rsp5 as a *ubc4Δ ubc5Δ* strain expressing a form of Ubc4 that can operate with HECT E3s but not RING E3s (the N78S mutant) grows much better than *ubc4Δ ubc5Δ* cells without any Ubc4/5 (Stoll *et al.* 2011). We tested whether *ubc4Δ ubc5Δ* cells expressing Ubc4(N78S) had a noticeable defect in a1-Ura3-3HA degradation, as assayed by cycloheximide-chase. Indeed, a1-Ura3-3HA was more stable in this strain than in WT cells (Supplementary Figure S3B). Given the clear role of Ubr1 in targeting *Dega1C*-Ura3-3HA, we tested whether Ubr1 contributes to the degradation of full-length a1. We found a1-Ura3-3HA was robustly stabilized in a *ubc4Δ ubc5Δ ubr1Δ* strain expressing Ubc4(N78S) (Supplementary Figure S3B). In summary, full-length a1, at least in the context of a Ura3-3HA fusion, appears to be targeted largely by Ubr1 and other unidentified RING-based E3s.

An out-of-frame peptide from MATa1 is a novel degron with distinct targeting in the cytoplasm and nucleus

While analyzing potential degrons derived from the N-terminal domain of a1 (data not shown), we tested whether a short a1-derived sequence could function as a C-terminal degron by creating a construct to express Ura3-HA followed by amino acids 20–40 of a1. Yeast carrying this plasmid as the only form of *URA3* grew very rapidly on media lacking uracil, suggesting the protein is long-lived (Figure 3A). Serendipitously, in cloning the a1-derived sequence into our Ura3-HA fusion plasmid, one clone suffered a frame shift that gave rise to a novel out-of-frame C-terminal sequence we will refer to as *DegOOF*. This clone grew very poorly on media lacking uracil, suggesting rapid degradation of the fusion protein. Protein stability studies with *DegOOF*-bearing substrates confirmed its degron activity (see Figure 5). This 17-residue sequence is very hydrophobic, with the final eight amino acids all bearing hydrophobic side chains (ILIIFIIL; Figure 3A). Previous studies have shown that model proteins bearing an NLS and five or more contiguous hydrophobic amino acids at the very C-terminus are often recognized by San1 in yeast, an E3 mostly localized in the nucleus (Fredrickson *et al.* 2011, 2013). However, Ura3-HA-*DegOOF* had no clear NLS but was short-lived.

Assuming Ura3-HA-*DegOOF* is a short-lived protein based on growth assay results, we transformed the plasmid expressing this reporter into a subset of strains from the TUS2.1 library and tested for growth on medium lacking uracil to help identify the ubiquitylation machinery for *DegOOF*. The strains tested were those lacking the E2s Ubc4, Ubc5, Ubc6, Ubc7, Ubc8, Ubc11, or Ubc13, and those lacking the E3s Ubr1, Ubr2, Slx8, San1, Ufd4, or Ltn1. Growth was rapid only with E2 mutant strains lacking either Ubc6 or Ubc7, the E2s that operate with the E3 Doa10 (data not shown). Indeed, cells lacking Doa10 that expressed Ura3-HA-*DegOOF* also grew rapidly on medium lacking uracil, suggesting stabilization of the protein. Strains that expressed the same reporter and lacked either Hrd1 or Asi1, the other E3s known to operate with Ubc7, grew similarly to WT cells (Figure 3B).

Since proteins with significant regions of exposed hydrophobicity are known to be targeted by the nuclear E3 San1, we wanted to test whether appending the NLS from SV40 T antigen would change the pathways that target *DegOOF* for degradation. For this and subsequent work with other degrons, we created

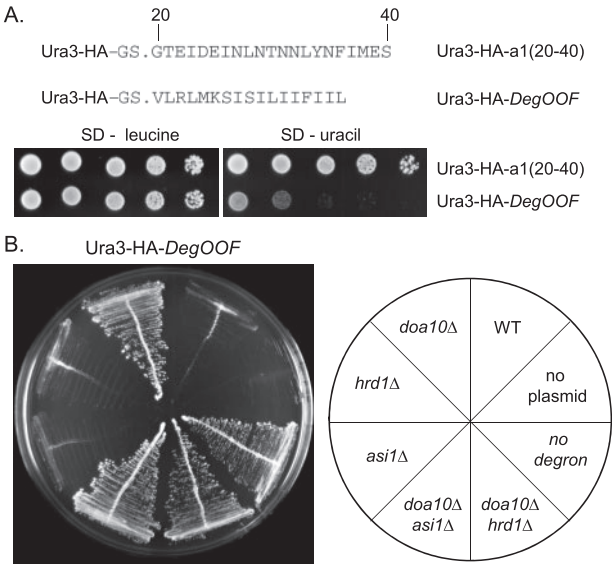

Figure 3 A Ura3-based reporter protein with a novel degron, *DegOOF*, at its C-terminus is targeted for degradation by the ERAD ligase Doa10. (A) Plasmids expressing the indicated Ura3-HAtag-fusion proteins from the *MET25* promoter were transformed into WT yeast (MHY501) and growth on media lacking either leucine or uracil was assayed. The numbered amino acids fused to the C-terminus of Ura3-HA shown above the growth assay data are residues 20–40 of MATa1. The sequence of amino acids now called *DegOOF* are derived from an out-of-frame clone of the DNA for codons 20–40 of MATa1. (B) Growth assay on media lacking uracil in strains carrying a plasmid expressing Ura3-HA-*DegOOF*, except the two streaks noted as "no plasmid" or "no degron." These two sectors show the growth of a WT strain carrying either no plasmid or a plasmid encoding Ura3-HA with no degron. Strains used for this figure were: MHY501 (WT), MHY6197 (*doa10*Δ), MHY6199 (*hrd1*Δ), MHY6194 (*doa10*Δ *hrd1*Δ), MHY8940 (*asi1*Δ), and MHY8949 (*doa10*Δ *asi1*Δ).

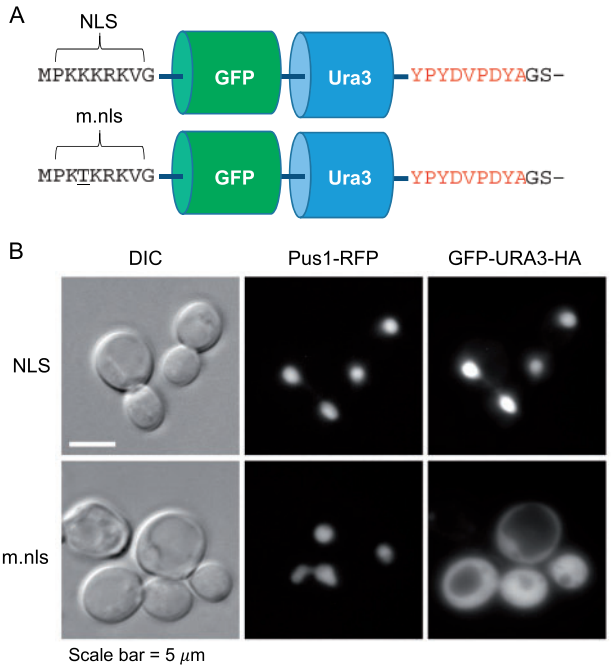

Figure 4 Matched GFP reporter proteins for testing the effects of nuclear localization on degron properties in yeast. (A) Schematic for the reporter proteins, which are expressed from plasmids bearing the *MET25* promoter. The NLS is the strong SV40-NLS and the mutant NLS (m.nls) has a single substitution (underlined) known to perturb its NLS function. DNA for C-terminal degrons can be cloned so the degron follows a Gly-Ser linker after the HA tag (in red). (B) Representative fluorescence images of WT yeast (MHY501) expressing the GFP reporter proteins with no degrons fused. The cells also expressed Pus1-RFP, which is a nuclear marker. DIC, differential interference contrast.

constructs that express the protein NLS-GFP-Ura3-HA with cloning sites to add in-frame C-terminal fusions after the HA tag (Figure 4A). For matched proteins that are not directed to the nucleus, we constructed plasmids that express m.nls-GFP-Ura3-HA proteins, in which the NLS is inactivated by a single missense mutation (see Methods and Figure 4A). Localization of NLS-GFP-Ura3-HA was very distinct from localization of m.nls-GFP-Ura3-HA, with the expected strong nuclear enrichment for the former protein (Figure 4B).

Interestingly, the nuclear NLS-GFP-Ura3-HA-*DegOOF* was targeted for degradation by more E3s than m.nls-GFP-Ura3-HA-*DegOOF* (Figure 5, B–E). Whereas the *DegOOF* fusion with the inactive m.nls was targeted mainly by Doa10, the *DegOOF* fusion directed to the nucleus was targeted by at least three E3s: Doa10, San1, and Ubr1. Ubr1 may also play a minor role in targeting *DegOOF* that is not directed to the nucleus (Figure 5C), but this could involve the fraction of the fusion protein that enters the nucleus without a specific NLS (Figure 4B). For the *DegOOF* reporter directed to the nucleus, levels were strongly increased by simultaneous deletion of Doa10 and Ubr1 (Figure 5, B and D). However, the rate of degradation was not much different between *doa10*Δ and *doa10*Δ *ubr1*Δ strains (Figure 5D). Analysis of NLS-GFP-Ura3-*DegOOF* degradation in a *doa10*Δ *ubr1*Δ *san1*Δ strain suggests that San1 plays a more prominent role when levels of the reporter increase in the absence of Doa10 and Ubr1 (Figure 5E).

Based on its sequence and targeting by Doa10, the *DegOOF* degron is likely to form an amphipathic alpha helix, although the boundaries of the helix are not known (Figure 6, A and B). Interestingly, a truncated form of *DegOOF* lacking the last five residues, which we call *DegOOF*-5, is also a degron with properties very similar to *DegOOF*, as determined by growth assays in various strains expressing NLS- and m.nls-containing reporter proteins (Figure 6C). Truncating just 2 additional residues to make *DegOOF-7* results in apparent loss of degron activity. *DegOOF*-5 terminates with only three consecutive hydrophobic residues. The targeting of nuclear-directed *DegOOF*-5 by Ubr1 and San1, in addition to Doa10, suggests that five or more contiguous hydrophobic residues are not needed at the C-terminus of this degron for San1 or Ubr1 recognition.

As a C-terminal extension of ~20 residues with high hydrophobicity, the *DegOOF* degron resembles the CL1 degron, which has been shown to drive Doa10-dependent degradation of proteins in yeast and MARCH6-dependent degradation in human cells (Gilon *et al.* 1998; Stefanovic-Barrett *et al.* 2018). However, no published studies of CL1 have examined the E3s responsible for degradation of CL1 fusion proteins that concentrate in the nucleus. We therefore tested the degradation dependence of NLS-GFP-Ura3-CL1, in this case by pulse-chase analysis. Strong stabilization of NLS-GFP-Ura3-CL1 required the simultaneous loss of Doa10, San1, and Ubr1 (Figure 6D). Thus, like *DegOOF* and *DegOOF*-5, the CL1 degron becomes a target for San1 and/or Ubr1 in addition to Doa10 if directed to the nucleus.

A short, extremely hydrophobic degron is targeted by San1 and Ubr1 but not ERAD ligases

While the requirements for a Doa10-dependent degron are not fully understood, a wealth of data suggests that an alpha helix with significant hydrophobicity is common to most Doa10 degrons (Geffen *et al.* 2016). By contrast, much literature suggests

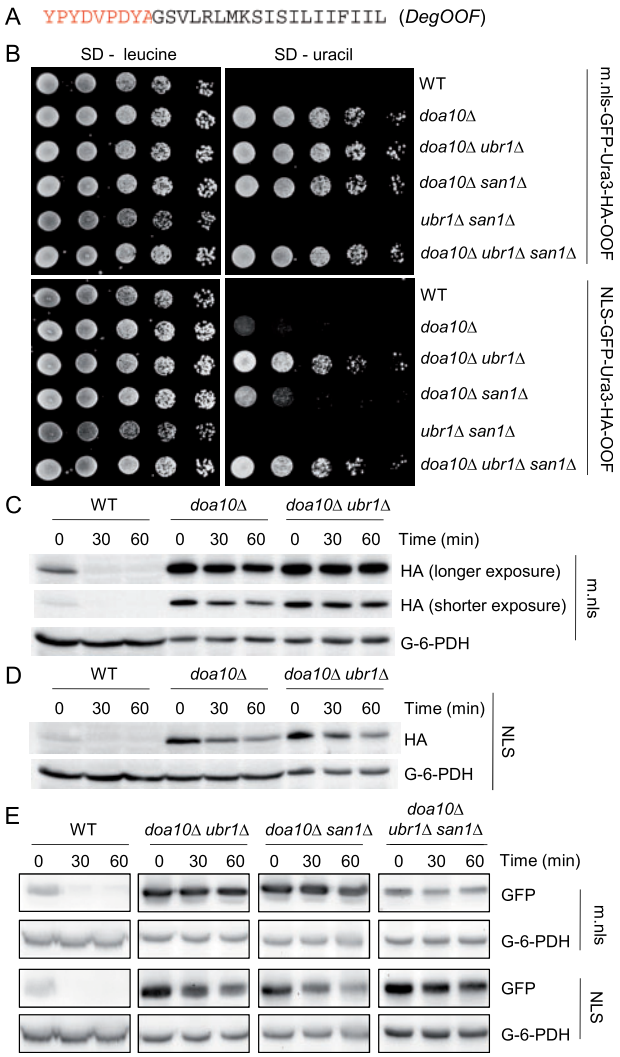

Figure 5 The *DegOOF* degron is targeted mainly by Doa10 in the cytoplasm but by Doa10, San1, and Ubr1 if directed to the nucleus. (A) Sequence of the HA tag (in red) and degron located at the C-termini of the reporter proteins described in Figure 4. (B) Growth assays on media lacking leucine or uracil of strains with plasmids expressing m.nls-GFP-Ura3-HA-*DegOOF* (cytoplasmic) or NLS-GFP-Ura3-HA-*DegOOF* (nuclear). (C) Cycloheximide-chase analysis using anti-HA immunoblotting in strains expressing m.nls-GFP-Ura3-HA-*DegOOF*. To avoid excessive signal of the fusion protein, fivefold less extract was loaded for strains lacking Doa10. (D) Cycloheximide-chase analysis in strains expressing NLS-GFP-Ura3-HA-*DegOOF*. Of important note, fivefold less extract was loaded for the *doa10*Δ *ubr1*Δ series. (E) Cycloheximide-chases as in parts C and D but including the *doa10*Δ *ubr1*Δ *san1*Δ strain and using anti-GFP to detect the fusion proteins. Fivefold more extract was loaded for WT. Strains used to generate data in this figure were: MHY501 (WT), MHY1631 (*doa10*Δ), MHY9564 (*doa10*Δ *ubr1*Δ), MHY3208 (*doa10*Δ *san1*Δ), MHY9563 (*ubr1*Δ *san1*Δ), and MHY9567 (*doa10*Δ *ubr1*Δ *san1*Δ).

that Ubr1 and San1 cooperate to target proteins with exposed hydrophobicity in general (Heck *et al.* 2010; Prasad *et al.* 2018). As noted above, a conspicuous feature of the *DegOOF* degron is that the final eight residues are all hydrophobic. We next made constructs that expressed fusion proteins bearing only the final seven residues of *DegOOF* after the HA tag and a 2-residue linker. We called this *DegF7H*, for degron with final seven hydrophobic amino acids. Both the nuclear-directed and m.nls-bearing *DegF7H* were short-lived proteins, and in both cases the proteins were targeted by San1 and Ubr1 and not Doa10 (Figures 7 and 8).

Accumulation of the reporter proteins bearing *DegF7H* in cells lacking Ubr1 and San1 was evident by growth assay and when these proteins were visualized for GFP fluorescence via microscopy (Figure 7). While the GFP signal was minimal for both the NLS- and m.nls-bearing reporter proteins in WT cells, the reporter proteins accumulated to readily detectable levels in *ubr1*Δ *san1*Δ cells (Figure 7C). The nuclear-directed protein concentrated in the nucleus, with bright puncta observed in many cells. In contrast, a significant fraction of the GFP signal for the m.nls-bearing protein was outside of the nucleus, including some puncta clearly not in the nucleus. Interestingly, however, there was some concentration of the m.nls-GFP-Ura3-*DegF7H* signal in the nucleoplasm and some of the puncta observed for this reporter appeared to be at or near the nuclear membrane (Figure 7C and data not shown). Moderate concentration of m.nls-bearing protein in the nucleoplasm of *ubr1*Δ *san1*Δ cells is consistent with San1 and Ubr1 normally functioning in degradation of this pool of the protein. The various puncta observed in our studies likely constitute previously characterized types of protein inclusions, though further study would be required for confirmation (Kaganovich *et al.* 2008; Johnston and Samant 2020).

A recent report showed that two distinct nonmembrane PQC substrates were fully stabilized in either of two double mutant strains: *ubr1*Δ *san1*Δ or *doa10*Δ *hrd1*Δ (Samant *et al.* 2018). We therefore wanted to test how m.nls-GFP-Ura3-HA-*DegF7H* would behave in each of these strains. Degradation of this protein was unaffected in the *doa10*Δ *hrd1*Δ strain (Figure 8), consistent with our finding that Doa10 does not target *DegF7H* and many studies suggesting that under normal growth conditions, yeast Hrd1 only targets intramembrane or ER luminal degrons, and not degrons exposed to the cytoplasm (Mehrtash and Hochstrasser 2019).

Additional hydrophobic C-end degrons targeted by redundant E3s with more rapid degradation in the nucleus

Many C-terminal degrons of relatively high hydrophobicity have been reported in recent years (Fredrickson *et al.* 2011, 2013; Maurer *et al.* 2016; Koren *et al.* 2018). However, the ubiquitylation pathways that target them are often unclear. With a focus on identifying San1 substrates in yeast, the Gardner group has used constructs expressing GFP-NLS-degron proteins from the strong *GAL1* promoter (Fredrickson *et al.* 2011, 2013). While several degrons were shown to be San1 substrates in this work, other hydrophobic degrons were not stabilized in *san1*Δ cells. One of these degrons, which we call pentaV, terminates with five contiguous valine residues and was part of a series of similar degrons we will refer to as C-terminal contiguous hydrophobic amino acid degrons (C-CHADs). The published GFP-NLS-pentaV protein was only slightly stabilized in cells lacking three PQC E3s: Ubr1, San1, and Doa10 (Fredrickson *et al.* 2013).

In our work, NLS-GFP-Ura3-pentaV was expressed from the relatively weak *MET25* promoter, and it appeared, based on growth on medium lacking uracil, to be very short-lived and only partially stabilized in a *san1*Δ *ubr1*Δ strain (Figure 9A). We did not observe any further apparent stabilization with additional deletion of the E2 gene *UBC4* or the E3 gene *DOA10*. Interestingly, the corresponding m.nls-GFP-URA3-pentaV protein was apparently not nearly as short-lived as the nucleus-directed protein based on growth assays (Figure 9A). We confirmed a drastic difference in degradation rates of these two proteins, which differ only by a single amino acid in the NLS, by cycloheximide-chase analyses (Figure 9, B and C). Further study revealed that Ubr1 plays a primary role in targeting the NLS-GFP-Ura3-pentaV reporter, with

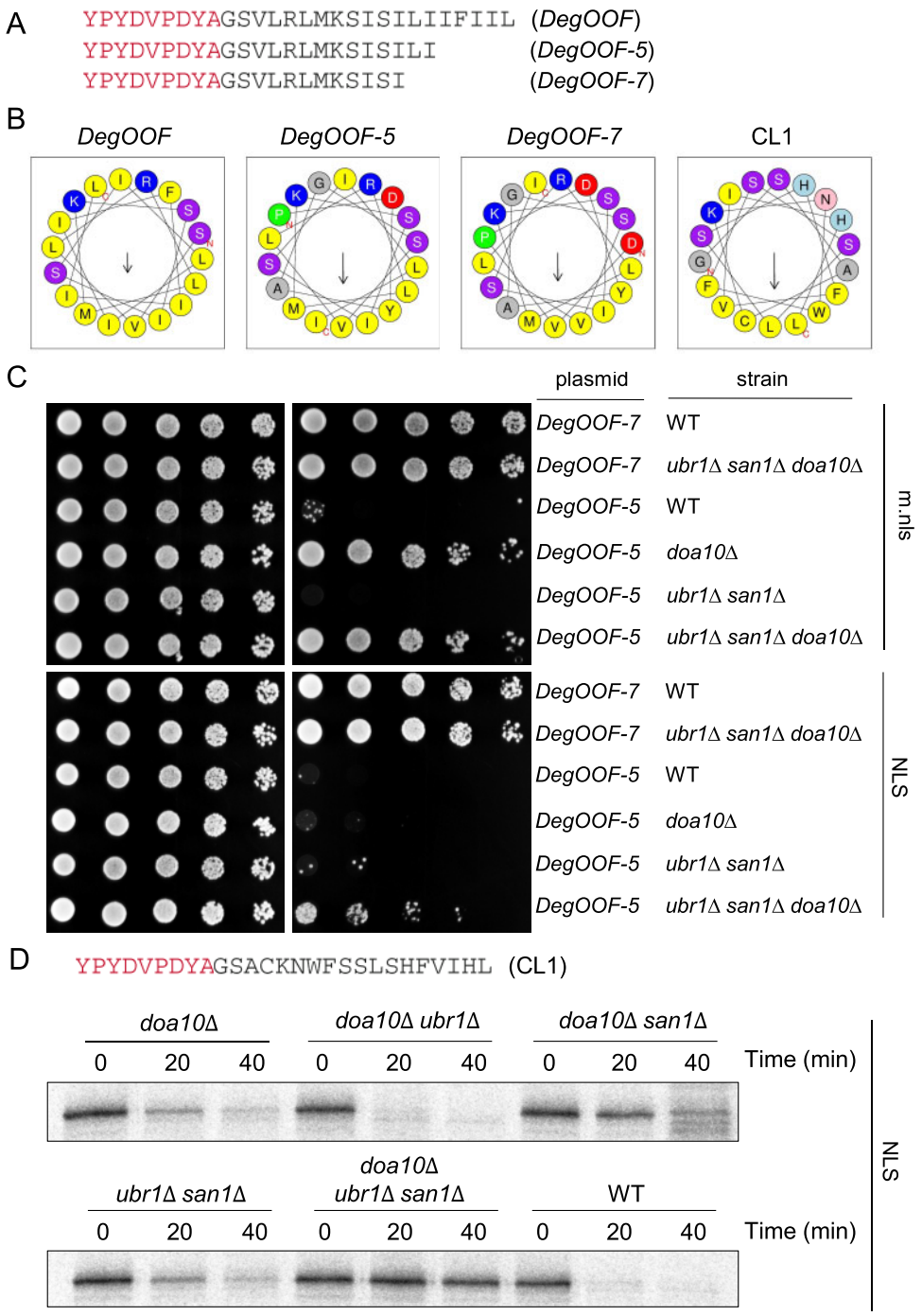

Figure 6 A truncated form of *DegOOF* and the CL1 degron are also targeted by a combination of Doa10, Ubr1, and San1 if directed to the nucleus. (A) Sequences for the HA tag (in red) followed by the *DegOOF* degron or truncated forms of *DegOOF*, as indicated, located at the C-termini of the reporter proteins described in Figure 4. (B) Helical wheel plots using an 18-residue window for potential alpha helices formed by the sequences listed in part A compared to that likely formed by the well-characterized CL1 degron. Plots were generated using Heliquest. Note that some of the predicted helices involve amino acids from the HA tag. (C) Growth assays on media lacking leucine or uracil in strains carrying plasmids expressing either NLS- or m.nls-GFP-Ura3-HA-fusions, as indicated. (D) Pulse-chase analysis for degradation of the NLS-GFP-Ura3-HA-CL1 protein in the indicated strains. These data are representative of data from three independent pulse-chase experiments. Strains used to generate data in this figure are: MHY501 (*WT*), MHY1631 (*doa10Δ*), MHY9596 (*doa10Δ ubr1Δ*), MHY3208 (*doa10Δ san1Δ*), MHY9563 (*ubr1Δ san1Δ*), and MHY9567 (*doa10Δ ubr1Δ san1Δ*).

contributions from San1 upon increased levels of the reporter protein when Ubr1 is missing (Figure 9, D and E). Thus, a degron first studied because of its likely recognition by San1 is targeted by Ubr1, San1, and additional unidentified degradation pathways.

In other work aimed at finding novel C-terminal degrons in yeast, a degron ending in seven contiguous hydrophobic amino acids and called "10–34" was shown to be targeted by the E3 Ltn1 (Maurer *et al.* 2016). Ltn1 is best characterized as an E3 targeting proteins on stalled ribosomes (Doamekpor *et al.* 2016). Importantly, the sequence after Ura3 in the original Ura3-10-34 construct included many residues between Ura3 and the sequence designated as the degron—the seven contiguous hydrophobic residues VVLVVVF (see Supplementary Figure S4B;

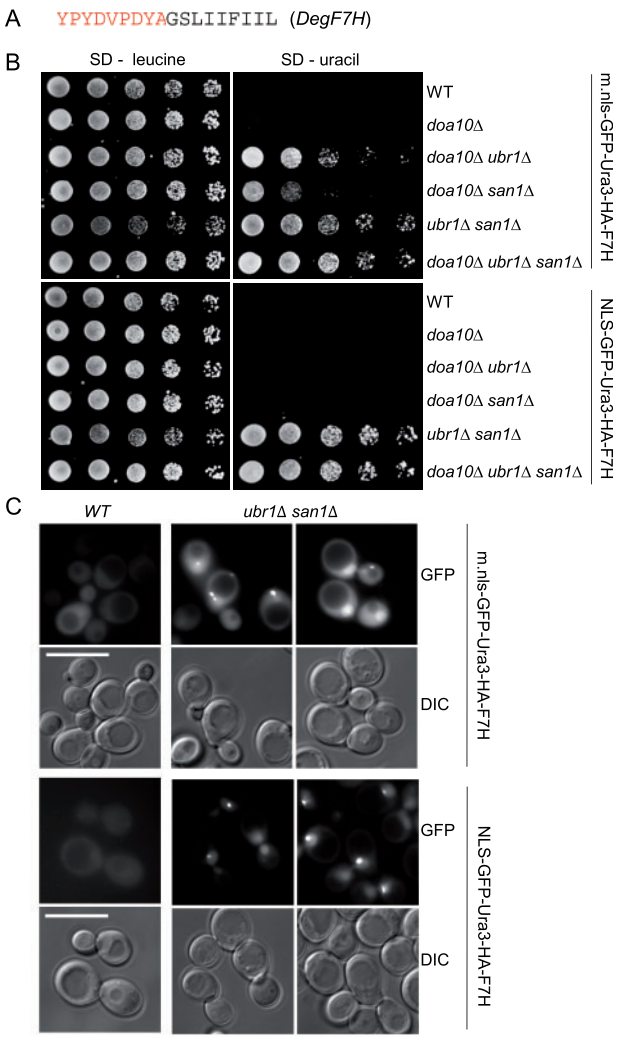

Figure 7 A hydrophobic degron bearing the final seven residues of *DegOOF*, called *DegF7H*, is targeted by Ubr1 and San1 but not Doa10. (A) Sequence of the HA tag (in red) and C-terminal degron. (B) Growth assays on media lacking leucine or uracil in strains bearing plasmids expressing m.nls-GFP-Ura3-HA-*DegF7H* or NLS-GFP-Ura3-HA-*DegF7H*. (C) Strains of the indicted genotype were transformed with plasmids expressing either m.nls-GFP-Ura3-HA-*DegF7H* or NLS-GFP-Ura3-HA-*DegF7H* and cells were imaged via DIC or for GFP fluorescence. Strains used to generate data in this figure were: MHY501 (WT), MHY1631 (*doa10*Δ *ubr1*Δ), MHY9564 (*doa10*Δ *san1*Δ), MHY3208 (*doa10*Δ *ubr1*Δ *san1*Δ), MHY9563 (*ubr1*Δ *san1*Δ), and MHY9567 (*doa10*Δ *ubr1*Δ *san1*Δ).

Maurer *et al.* 2016). We reanalyzed the 10–34 degron in the context of our NLS- or m.nls-GFP-URA3-HA-degron constructs, with only a Gly-Ser linker between the single HA tag and the VVLVVVF sequence. The NLS version was significantly stabilized in a *san1*Δ *ubr1*Δ strain by growth assay (Figure 9C); m.nls-GFP-Ura3-HA-VVLVVVF was not stabilized at all in *ltn1*Δ cells, as measured by growth assay, but was stabilized in *san1*Δ instead (Supplementary Figure S4A). We did verify that the original Ura3-10-34 construct characterized by Maurer *et al.* appeared to be at least partially stabilized in cells lacking Ltn1 (Supplementary Figure S4B). Considering that many of the degrons found in Maurer *et al.* showed some Ltn1 targeting, we hypothesize that an element in the "linker" region between Ura3 and the sequence being called the degron (VVLVVVF) in the original Ura3-10–34 contributes to Ltn1-mediated degradation, likely through ribosome stalling. Like *DegF7H*, the seven-residue sequence previously

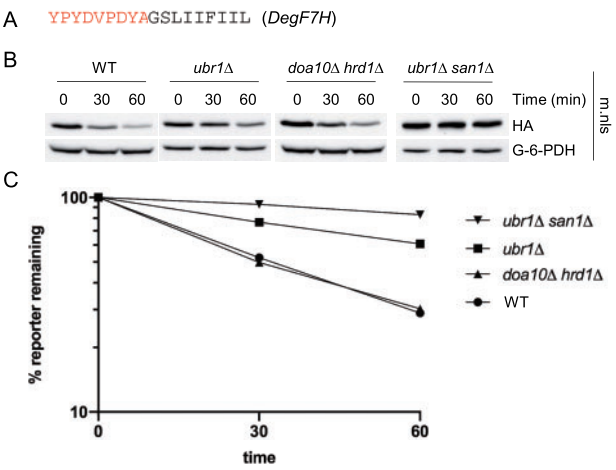

Figure 8 *DegF7H* is not stabilized in ERAD deficient cells. (A) Sequence of the HA tag (in red) and C-terminal degron. (B) Cycloheximide-chase analysis of strains expressing m.nls-GFP-Ura3-HA-*DegF7H*. To avoid excessive signal, less extract was loaded for samples from cells lacking Ubr1. (C) Quantification of anti-HA signal from cycloheximide results shown in part B. Strains used to generate data in this figure were: MHY501 (WT), MHY9532 (*ubr1*Δ), MHY9563 (*ubr1*Δ *san1*Δ), and MHY6194 (*doa10*Δ *hrd1*Δ).

called "10–34" is a C-CHAD and a reporter bearing this C-CHAD appears to be substantially targeted for degradation by the cytoplasmic/nucleoplasmic PQC ligases Ubr1 and San1.

Hydrophobic C-end degrons also function in human cells

A recent report showed that C-terminal extensions to GFP derived from the last 23 residues of all human proteins are often degrons (Koren *et al.* 2018). While that study revealed several degrons with specific sequence signatures, another large group of potential degrons were characterized by their general physicochemical properties, with bulky hydrophobic peptides correlating with metabolic instability. Consistent with hydrophobicity promoting rapid degradation, GFP-HA-*DegF7H* (hereafter shortened to GFP-F7H; Figure 10A) was much less abundant than GFP-HA in HeLa cells upon ectopic expression (Figure 10B). This was true whether we expressed the GFP proteins with the NLS or m.nls, suggesting the GFP-F7H proteins could be degraded in either the cytoplasm or nucleus.

To determine whether the GFP fusion proteins were being degraded by the proteasome, we measured fusion protein levels with and without addition of the MG132 proteasome inhibitor (Figure 10C). Notably, MG132 treatment did not increase GFP fusion protein levels when expressed from the derepressed (by doxycycline) CMV promoter, which causes the fusion proteins to be substantially overexpressed. Previous work with GFP-CL1 showed that the UPS can be overwhelmed by strong overexpression of a degron-containing protein, leading to the accumulation of ubiquitylated proteins (Bence *et al.* 2001). Indeed, overexpression of GFP-F7H led to a greater accumulation of bulk ubiquitylated proteins compared to overexpression of GFP alone (Figure 10D). MG132 treatment led to further accumulation of ubiquitylated proteins, showing the proteasome inhibitor was active in our assays. Based on the very hydrophobic nature of the F7H degron and previous studies of PQC substrate proteins, we reasoned that much of the GFP-F7H protein that potentially escaped UPS-mediated degradation could be in intracellular aggregates. Immunofluorescence analysis revealed this to be the case,

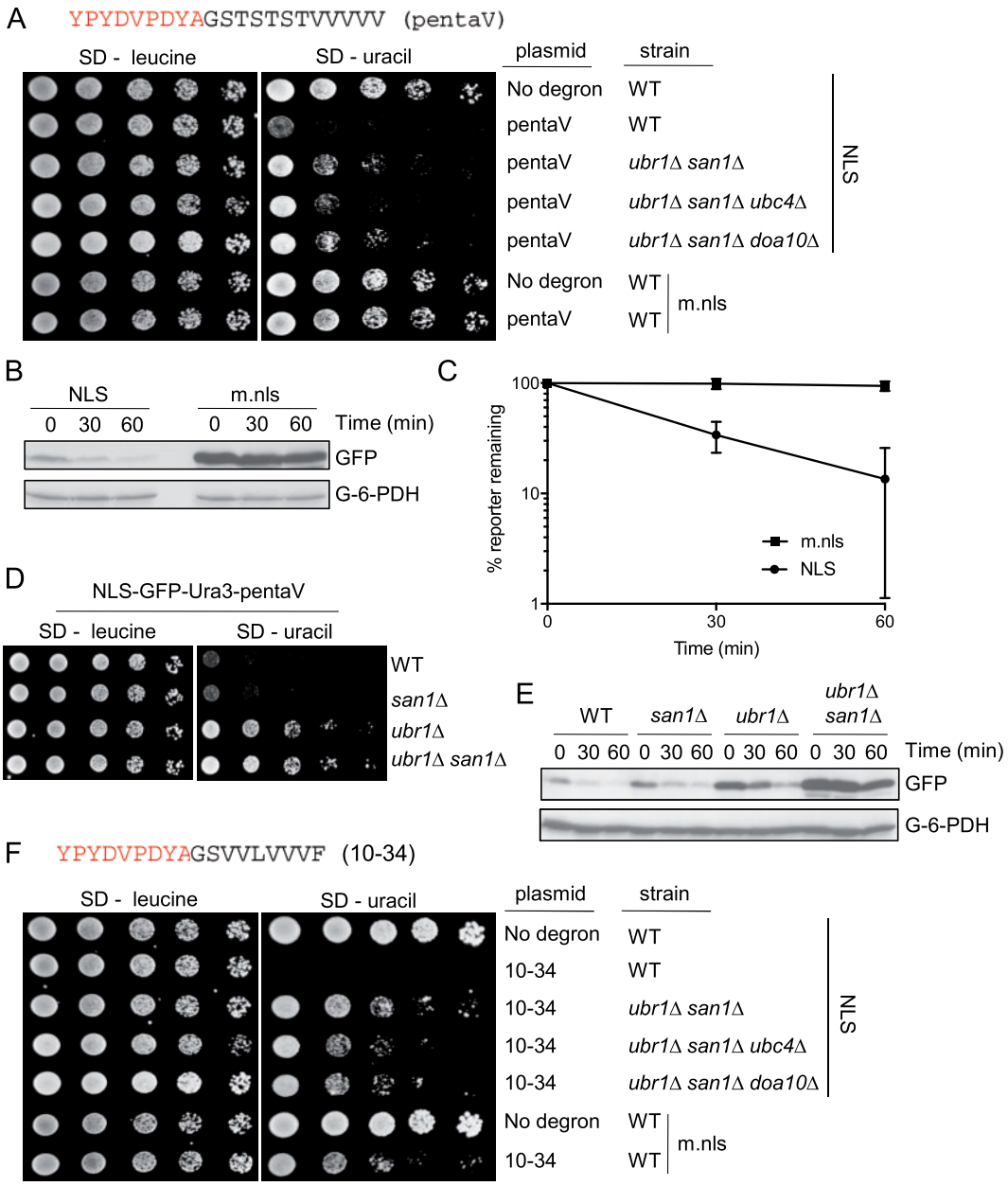

Figure 9 Additional C-end hydrophobic degrons are also partially targeted by Ubr1 and San1. (A) Growth assays on media lacking leucine or uracil in strains carrying plasmids expressing either NLS or m.nls GFP-Ura3-HA-pentaV fusions. The sequence of the C-terminal HA tag (in red) and pentaV degron is shown. (B) Cycloheximide-chase analysis of WT yeast expressing m.nls-GFP-Ura3-HA-pentaV or NLS-GFP-Ura3-HA-pentaV. (C) Quantification of degradation rates for m.nls-GFP-Ura3-HA-pentaV or NLS-GFP-Ura3-HA-pentaV in triplicate from cycloheximide-chase data, as shown in part B. (D) Growth assays of indicated strains carrying plasmids expressing NLS-GFP-Ura3-HA-pentaV fusions. (E) Cycloheximide-chase analysis of strains expressing NLS-GFP-Ura3-HA-pentaV. (F) Growth assays of strains carrying plasmids expressing either NLS or m.nls GFP-Ura3-HA-10-34 fusions. The sequence of the HA tag (in red) and 10-34 degron of the reporter protein is shown. Strains used to generate data in this figure were: MHY501 (WT), MHY9532 (*ubr1*Δ), MHY3178 (*san1*Δ), MHY9563 (*ubr1*Δ *san1*Δ), MHY9586 (*ubr1*Δ *san1*Δ *ubc4*Δ), and MHY9567 (*ubr1*Δ *san1*Δ *doa10*Δ).

with compartment-specific intracellular inclusions depending on the NLS or m.nls appended to the N-terminus of the GFP fusion proteins (Figure 10E).

Since low levels of the GFP reporters were detectable in the absence of induction by doxycycline (Figure 10B and data not shown), we measured reporter levels with and without MG132 under these conditions. In this case GFP fusion protein levels were substantially increased by MG132, showing that *DegF7H* functions as a UPS degron when expressed at moderate levels (Figure 10F). Our results are similar to those observed for CL1, in which cell clones with only moderate expression of a GFP-CL1 fusion protein exhibited MG132-induced stabilization of the reporter (Bence *et al.* 2001, 2005). Together, our studies of GFP reporter proteins bearing a model C-CHAD suggest that extremely hydrophobic degrons are potentially handled by multiple PQC mechanisms in human cells, including UPS-dependent degradation.

Since F7H is an artificial C-end degron, we examined the recently reported proteome-wide C-end degrons for natural—though perhaps cryptic—hydrophobic degrons (Koren *et al.* 2018). We began by using a web-based program to generate two distinct but similar measures of hydrophocity, a GRAVY (Grand Average

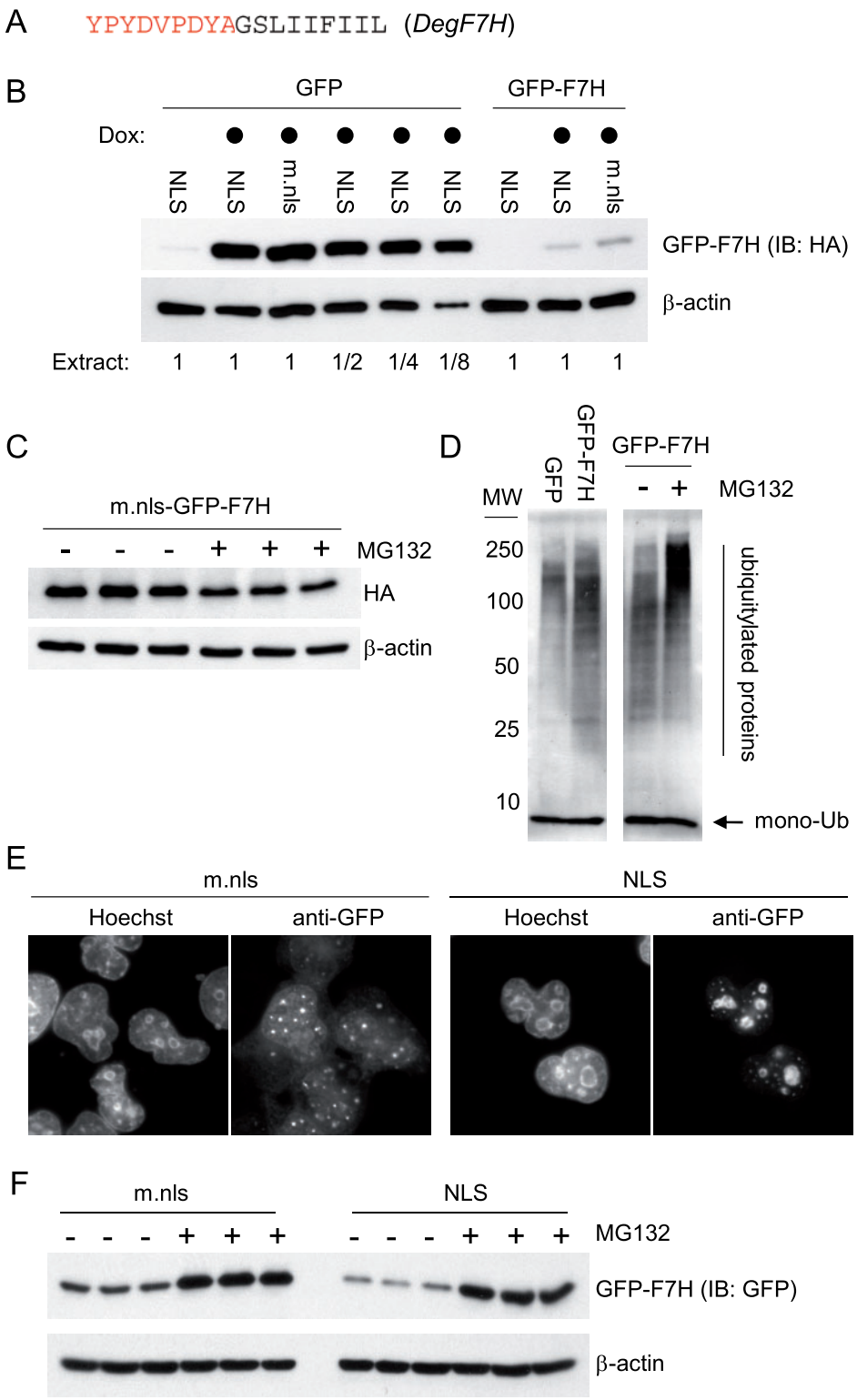

Figure 10 The *DegF7H* sequence functions as a PQC degron in human cells. (A) Sequence of the C-terminal HA tag (in red) and *DegF7H* degron of the GFP-based reporter protein. (B) Immunoblot analysis of extracts from HeLa cells expressing NLS-GFP-HA, m.nls-GFP-HA, NLS-GFP-HA-*DegF7H*, or m.nls-GFP-HA-*DegFH7* from a doxycycline (dox)-inducible CMV promoter. Relative amounts of extract are indicated. Doxycycline was added at 2 μg/ml 6 h before harvest where indicated. (C) Immunoblot analysis of HeLa cell extracts expressing high levels of m.nls-GFP-*DegF7H* (2 μg/ml doxycycline) with and without treatment with MG132 (10 μM) for 6 h. Three independent wells of cells were processed for each condition. Similar results were obtained with cells expressing the NLS-GFP-*DegF7H* protein (data not shown). (D) Immunoblot analysis of select HeLa cell extracts from parts B and C using an antibody to ubiquitin shows that strong expression of GFP-HA-*DegF7H* induces accumulation of ubiquitylated proteins, which is exacerbated by proteasome inhibition via MG132. The extracts analyzed here expressed m.nls-GFP-HA with or without the degron, but similar results were observed with the NLS-GFP-HA versions. (E) Fluorescence images of cells expressing m.nls-GFP-HA-*DegF7H* or NLS-GFP-HA-*DegF7H* stained for DNA (Hoechst) or with anti-GFP antibodies. Images are representative of hundreds of cells observed. (F) Immunoblot analysis of HeLa cell extracts expressing m.nls-GFP-*DegF7H* or NLS-GFP-*DegF7H* with and without treatment with MG132 (10 μM) for 8 h. These cells were not treated with doxycycline. Three independent wells of cells were processed for each condition, as shown. Blotting with anti-HA yielded similar results to those obtained with anti-GFP. Immunoblot for β-actin serves as a loading control throughout this figure.

of Hydropathy) score and the other simply called "hydrophobicity," as well as a theoretical isoelectric point (pI) for the C-terminal 23 amino-acid segments (23-mers) previously assigned a protein stability index (PSI) in the context of a C-terminal extension to GFP (Supplementary Table S5). By plotting hydrophobicity score versus PSI, it is clear that increasing hydrophobicity correlates with a lower PSI; that is, an apparent increased likelihood of inducing degradation (Figure 11A). Conversely, isoelectric point does not correlate with PSI (Supplementary Figure S5A).

Although hydrophobicity is correlated with PSI, the correlation is far from perfect. This is especially evident in two parts of the graph of PSI versus hydrophobicity: (1) a sizable group of peptides of high hydrophobicity that did not yield low PSI values and (2 peptides that yielded a low PSIs but were not hydrophobic. This latter group can be understood since many degrons are not based on hydrophobicity but rather sequence signatures, such as the C-end degrons recently shown to be recognized by various CRL ligases (Koren *et al.* 2018). Indeed, the likelihood of being a CRL substrate markedly decreases with increasing hydrophobicity for the putative C-end degrons (Figure 11B). Interestingly, although a terminal glycine residue is underrepresented in the entire set of natural human C-termini and for peptides with high hydrophobicity that yielded lower than average PSI values, a terminal glycine is overrepresented for peptides that had lower than average PSI values and low hydrophobicity (Figure 11C). This enrichment of glycine-ended peptides is even more striking when focusing on the peptides with the lowest hydrophobicity and PSI values well below average (Supplementary Figure S5B; Koren *et al.* 2018). Three types of glycine-end degrons, differentiated by penultimate residues, have been characterized as substrates for different CRL ligases (Koren *et al.* 2018; Rusnac *et al.* 2018). Other degrons based on specific sequences rather than hydrophobicity are recognized by E3s other than CRLs, which seems to be the case for C-end degrons that end in alanine (Koren *et al.* 2018). Peptides that end in alanine and also cysteine were also overrepresented in the peptides that yielded lower than average PSI values and were of relatively low hydrophobicity overall (Figure 11C, Supplementary Figure S5B, and Supplementary Table S5).

Another likely reason increasing hydrophobicity of the C-end peptides does not perfectly correlate with decreasing PSI is that some hydrophobic sequences have a potential biological rationale for their stability. Among the most hydrophobic peptides (hydrophobicity of 70 or greater), many did not have a low PSI value (Figure 11A). Some of these very hydrophobic peptides terminated with at least five hydrophobic amino acids, somewhat resembling *DegF7H* and other C-CHADs (Table 1). Examination of the proteins with C-terminal peptides that yielded PSI values above average (avg. = 2.76) and hydrophobicity scores above 70 revealed that all of these proteins terminate in a transmembrane domain (TMD)—for example, SNARE proteins (Table 1 and Supplementary Table S5). It is likely that many of the GFP-TMD proteins derived from such "tail-anchored proteins" efficiently inserted into biomembranes and were therefore relatively stable. In contrast, many GFP-TMD proteins had low PSI values, and these proteins likely failed to insert correctly into biological membranes. It is already appreciated that the E3 Doa10/MARCH6 recognizes tail-anchored and GPI-anchored proteins that fail to insert properly into membranes (Ast *et al.* 2014; Stefanovic-Barrett *et al.* 2018; Dederer *et al.* 2019). Our bioinformatic analysis highlights the importance of considering the cellular localization of proteins bearing sequences that may become degrons under other conditions.

Discussion

In this report, we have shown that at least two types of relatively short hydrophobic degrons exist in yeast with distinct E3 ligase specificities that vary depending on the subcellular localization of the protein bearing the degron. One type of degron is recognized by Doa10 at either the cytoplasmic face of the ER or the nucleoplasmic face of the inner-nuclear membrane, presumably because the degron can form an α-helix, but can also be recognized by the E3s San1 and Ubr1 if localized to the nucleus (Figures 5 and 6). Another type of degron, characterized by a contiguous stretch of hydrophobic amino acids, is recognized by San1 and Ubr1 but not Doa10—presumably because it cannot form an α-helix—whether primarily localized to the cytoplasm or the nucleus (Figures 7–9). However, some C-CHADs are only short-lived if directed to the nucleus, which could be related to C-CHAD length/degree of hydrophobicity (Figure 9). The C-CHAD *DegF7H* functions as a degron with or without direction to the nucleus and appeared to be nearly fully stabilized in *ubr1*Δ *san1*Δ cells (Figure 8). These properties and its small size make *DegF7H* an excellent model degron for studies of the often multiplexed function of San1 and Ubr1. Degradation driven by the C-CHAD sequences VVVVV and VVLVVVF is only partially due to the combined activity of San1 and Ubr1 (Figure 9), suggesting other pathways for degradation of certain C-CHADs (Fredrickson *et al.* 2013). The mechanisms determining the division of labor between San1, Ubr1, and additional pathways targeting different C-CHADs are an important topic for future studies.

It was recently shown that fluorescent proteins (FPs) bearing the C-terminal helical and hydrophobic degron CL1 expressed in human cells are targeted for degradation by multiple E3s (Stefanovic-Barrett *et al.* 2018; Leto *et al.* 2019). One study showed mCherry-CL1 is targeted by a combination of two membrane-embedded E3s, MARCH6 and RNF139/TRC8 (Stefanovic-Barrett *et al.* 2018). The other study implicated RNF139/TRC8 and an E3 that is not an integral membrane protein—the proteasome-associated UBE3C—in degradation of GFP-CL1 (also called GFPu; Leto *et al.* 2019). While the FP-CL1 proteins were not directed to the nucleus in these studies, some fraction of the expressed protein likely diffused into the nucleus due to its small size. Interestingly, UBE3C localizes to the nucleus, at least in MCF-7 breast cancer cells (Okada *et al.* 2015). Thus, the reported differences in E3s responsible for degradation of the two different FP-CL1 proteins may reflect differences in nuclear localization for the different FP-CL1 proteins (which may depend on expression level) or the relevant E3s in the different cell types used.

It would be interesting to determine the E3s responsible for targeting a nucleus-directed version of FP-CL1 in human cells since we observed involvement of additional ubiquitylation pathways when CL1 was directed to the nucleus in yeast (Figure 6). Leto *et al.* identified ubiquitylation machinery for multiple ERAD substrates in human cells, with UBR4 having a role in degradation of the integral membrane protein INSIG1-GFP. The human genome encodes at least seven proteins containing a UBR box, a domain first characterized in yeast Ubr1 (Tasaki *et al.* 2005). Yeast Ubr1 has previously been shown to participate in ERAD under certain conditions (Stolz *et al.* 2013). In our work, Ubr1 can target proteins also targeted by Doa10, namely a reporter protein bearing CL1 and a reporter bearing our newly characterized degron *DegOOF* (Figures 5 and 6). Thus, soluble E3s, including those of the UBR box-containing family, can play important roles in targeting proteins that are also recognized by membrane-associated E3s. One role for the involvement of multiple E3s in degradation

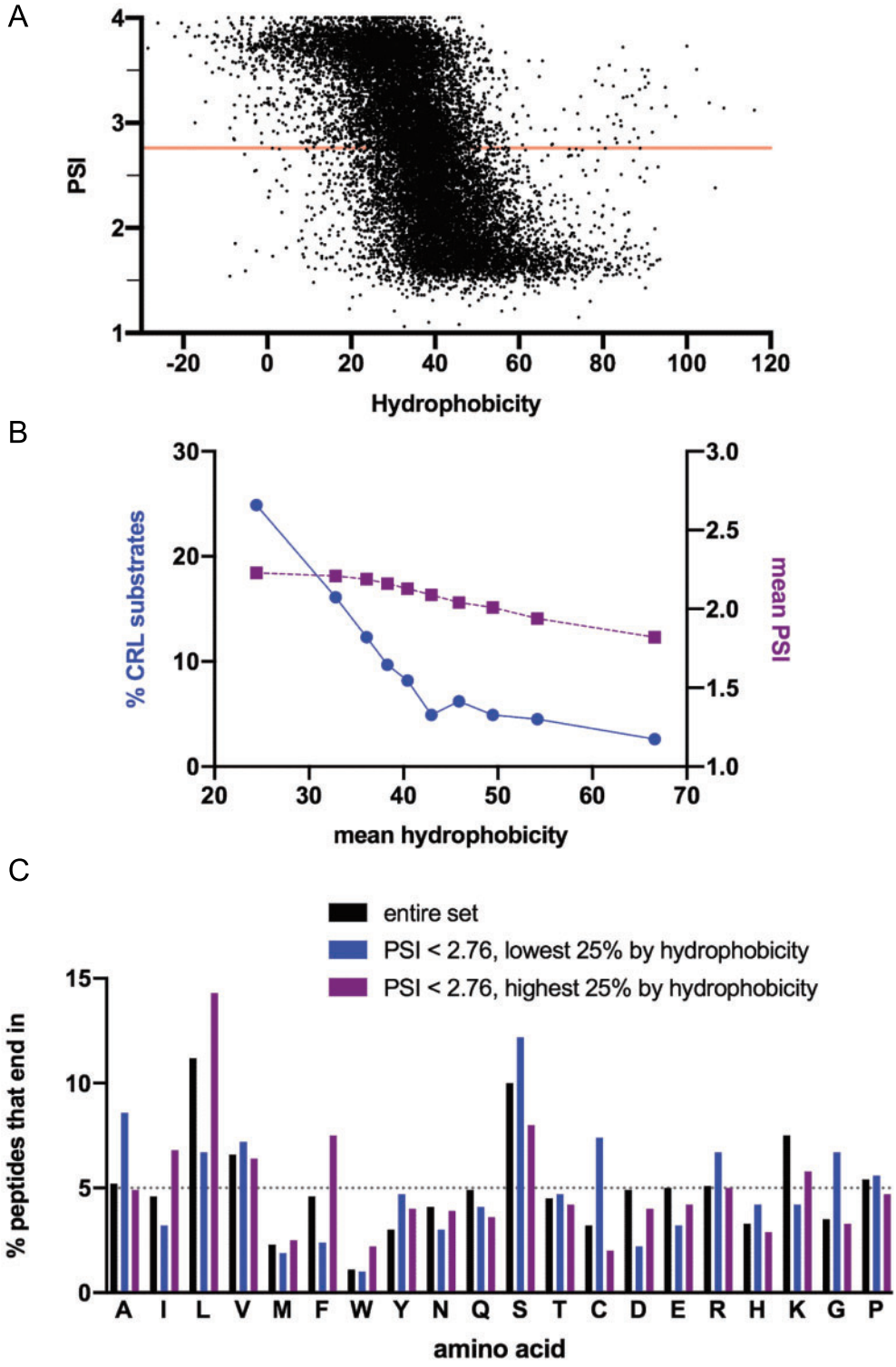

Figure 11 Bioinformatic analysis of human C-terminal segments reveals distinctions between nonhydrophobic and hydrophobic degrons, the latter rarely being targets of CRL-type ubiquitin ligases. (A) Plot of protein stability index (PSI) versus hydrophobicity for 16,345 peptides derived from the C-terminal 23 residues of human proteins. PSI values are from Koren et al. (2018), in which the 23-mer peptides were appended C-terminally to the GFP protein. The red line at PSI = 2.76 shows the mean PSI for the 16,345 GFP-peptide proteins. (B) All 23-mers with PSI values below 2.7 were sorted by increasing hydrophobicity and then separated into 10 groups with an equal number of peptides (770). For each group, mean PSI and mean hydrophobicity was calculated and plotted (purple). Also plotted is the percentage of GFP-23-mer proteins in that group that were assigned as CRL substrates (blue). (C) The percentage of peptides that terminate in the indicated residue (single letter amino acid code) was calculated and plotted for three groups of peptides: (1) the entire set of 16,345 peptides, (2) those that yielded PSI values below average (2.76) and were in the lowest 25th percentile for hydrophobicity, and (3) those that yielded PSI values below average (2.76) and were in the highest 25th percentile for hydrophobicity. A line at 5% represents the percentage expected if all 20 amino acids were used evenly. Part C shows enrichment of glycine, alanine, and cysteine residues at the C-termini of non-hydrophobic putative degrons but not at the termini of hydrophobic putative degrons.

Table 1 Not all hydrophobic C-terminal segments that appear to be C-CHADs led to degradation when appended to GFP expressed in human cells

Gene_ID	Peptide sequence	Hy	PSI	Description
TMEM31	LSFFILLVLLLLLFIIVFILIFF	116.1	3.12	Likely a tail-anchored protein
PLN	NLFINFCLILICLLLICIIVMLL	108.8	3.14	52 amino acid tail-anchored protein
STX5	RWLMVKIFLILIVFFIIFVVFLA	102.3	3.51	SNARE; tail-anchored protein
FMO1	PFESFLKVFSFLALLVAIFLIFL	94.9	3.34	ER enzyme; tail-anchored protein
OR2M2	KFFFLISIFFYDVKILALIMYIA	92.0	1.65	GPCR; multi-transmembrane
STX6	RQWCAIAILFAVLLVVLILFLVL	91.6	3.14	SNARE; tail-anchored protein
OCLM	YLKILYKSGIIWLSWYSFILLVL	90.6	1.85	44 amino acid protein; likely tail-anchored
SOGA3	LFSSLSYTTIFKLVFLFTLFFVL	90.2	1.73	Large protein; predicted tail anchored
RNF185	DEQFLSRLFLFVALVIMFWLLIA	85.7	3.25	Multi-pass membrane protein of ER/mito
ATP5G1/2	GFALSEAMGLFCLMVAFLILFAM	85.4	1.86	ATPase of mitochondria
VRK2	DVYYYRIIIPVLLMLVFLALFFL	84.3	1.92	Tail-anchored S/T-kinase
GP2	PSTAGFLVAWPMVLLTVLLAWLF	82.9	1.88	GPI-anchored protein
TNFRSF10C	SSHYLSCTIVGIIVLIVLLIVFV	81.2	2.97	GPI-anchored protein

After sorting all human protein-derived C-terminal peptides by hydrophobicity, those with hydrophobicity scores above 80 were searched for those ending in at least five consecutive hydrophobic amino acids (defined here as A, L, V, I, M, Y, W, or F). Listed are those peptide sequences plus their hydrophobicity scores (Hy), PSI values from Koren *et al.* (2018), and descriptions of the structure and function of the full-length protein bearing each C-terminus. See Supplementary Table S5 for data showing hydrophobicity and PSI values for all 23-mers.

of a single substrate is ubiquitin chain diversification, including mixed and branched chains (Yau *et al.* 2017; Samant *et al.* 2018; Haakonsen and Rape 2019; Leto *et al.* 2019). The degrons and constructs studied in our current work provide additional tools for studies of degradation and ubiquitin chain diversity, both in yeast and human cells.

Ubr1 has been known for many years to facilitate degradation of N-end rule substrates, proteins with specific N-termini (Bartel *et al.* 1990; Varshavsky 2012). Years later it was appreciated that Ubr1 has a role in PQC of misfolded proteins, often in concert with the nuclear localized San1 (see Introduction). While the hydrophobic C-terminal degrons in our current study are obvious PQC substrates (Figures 5–10), it is unclear which of the two branches of Ubr1 activity is responsible for degradation of fusion proteins bearing *Deg2* and *Dega1C* as their N-termini (Figures 1 and 2). *Deg2* has two important degradation elements, a region within the homeodomain and a set of hydrophobic amino acids within its linker region (Hickey and Hochstrasser 2015). Thus, it is possible that the hydrophobic degradation element in *Deg2* is being targeted by the PQC branch of Ubr1 activity. Our *Deg2* reporter construct is translated to begin with Met-Gln-, which is not a substrate for methionine aminopeptidase activity and is not predicted to be a substrate of Ubr1 (Varshavsky 2012; Wingfield 2017). *Dega1C* is essentially just a homeodomain; however, it has been shown that the homeodomain of a1 does not fully fold or bind DNA in the absence of its binding partner in diploids—α2 (Ke and Wolberger 2003). Thus, one could speculate that the apo form of a1 has some misfolded character that could make *Dega1C* a PQC substrate. Like *Deg2*, the very N-terminus of *Dega1C* (Met-Lys-) is not predicted to be an N-end rule substrate in yeast. We cannot, however, fully rule out proteolytic processing of either of these degron-fusions that might present a N-end rule substrate.

A recent study suggested that the Ubr1/San1 and Doa10/Hrd1 (ERAD) ubiquitin ligase pairs can each function in recognition of the same cytoplasmic soluble PQC substrate (Samant *et al.* 2018). These data are mostly consistent with our findings and those of others studying E3s involved in PQC with the exception of involvement of yeast Hrd1 in PQC of soluble, non-ER luminal proteins. Samant *et al.* used conditions under which the PQC substrates were significantly overexpressed, and direct recognition of these soluble proteins by Hrd1 was not demonstrated. It is possible that by expressing the misfolded proteins at such high levels, the overall capacity of PQC systems was nearly saturated even in WT yeast cells. Eliminating most of the ERAD system by deleting both Doa10 and Hrd1 could then overwhelm the remaining PQC capacity and thus proteins that are not necessarily substrates of ERAD ligases would be stabilized. Our current study and many previous reports have provided evidence that overexpression of PQC substrates can overwhelm the UPS in both yeast and humans (Figure 10). The reasons for and consequences of this phenomenon are complex, as recently reviewed (Johnston and Samant 2020). The PQC degrons studied in our current work (Figures 5–10), particularly *DegF7H*, are much smaller and simpler than an entire protein that is misfolded—as used in Samant *et al.* Simple PQC degrons expressed at moderate levels would be expected to have less chance of indirect impact on E3s not directly involved in their recognition.

DegF7H also functions as a degron in HeLa (human) cells (Figure 10), in which the ubiquitylation machinery for PQC is poorly understood. Future studies could use our matched reporters, m.nls-GFP-*DegF7H* and NLS-GFP-*DegF7H*, to address how human cells handle degrons consisting of consecutive hydrophobic residues. We also have made constructs to allow human cell expression of similar reporters bearing the *DegOOF* degron, which likely forms an alpha helix—like the better studied CL1 (unpublished). Comparing the ubiquitylation pathways that handle these different PQC reporters in human cells will allow modeling of how human cells handle different types of hydrophobic degrons—as we have done in this report for yeast. This is important since the human proteome contains numerous potential hydrophobic degrons, including those naturally at the C-termini of proteins (Figure 11; Table 1; Koren *et al.* 2018).

The TUS collection and the YUMI methods outlined here will serve as a useful resource for those investigating the yeast UPS. Distribution of the TUS library to other laboratories has already resulted in important findings. For example, the original TUS was used in a process analogous to YUMI to identify San1 and Ubr1 as the E3s for "cytoplasmic" PQC substrates (Heck *et al.* 2010). Another study used TUS2.0 to show involvement of the F-box protein Rcy1 in targeting the histone H3 variant Cse4 for degradation (Cheng *et al.* 2016). Our original version of the TUS library did not include F-box protein gene deletions, showing the utility of our updated TUS collection. Cse4 concentrations are tightly controlled in cells, and elevated concentrations of Cse4 cause growth defects—even more so in TUS strains lacking machinery for its

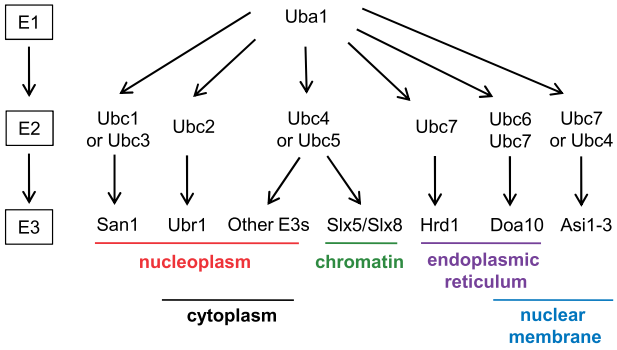

Figure 12 Schematic summary of PQC ubiquitylation pathways in *S. cerevisiae*. The cellular compartment covered by each pathway is indicated. While other pathways have been linked to PQC, the best characterized pathways in yeast are shown. Yeast express 13 Ubc proteins (E2s) total, 10 of which are known to conjugate ubiquitin (the three others conjugate ubiquitin-like proteins). Yeast express over 50 E3s (see Supplementary Table S3), many of which remain uncharacterized or poorly characterized. Based on studies with mammalian cells and expansion in the number of E3s in higher organisms, humans are likely to have additional complexity in PQC ubiquitylation.

ubiquitylation and degradation. Since many proteins are known to be detrimental to yeast growth upon their overexpression, use of the TUS in similar screens with other "toxic" proteins may be fruitful (Liu *et al.* 1992; Sopko *et al.* 2006; Yoshikawa *et al.* 2011). We employed reporter proteins fused to the Ura3 enzyme to estimate protein levels in growth-based assays. Other metabolic enzymes involved in other biosynthetic pathways commonly perturbed for yeast genetics have been employed for these purposes (Kim *et al.* 2013; Watts *et al.* 2015). Growth-based assays using reporter proteins are excellent for genetic screens and also offer facile estimation of protein levels when comparing many constructs or strains. However, measurements of protein degradation rates require other assays, such as the cycloheximide- and pulse-chase experiments we also employed in this study (Figures 1, 2, 5, 6, 8, 9).

Based on our current work and many additional studies, a model for the main yeast ubiquitylation pathways that handle PQC in different cellular compartments can be drawn (Figure 12). Our study employed a set of degrons that collectively were at least partially targeted by each of these pathways—except the membrane-associated E3s Asi1-3 and Hrd1, which have distinct specificities. The degrons we analyzed mark soluble cytoplasmic and nucleoplasmic proteins as UPS substrates. The full repertoire of E3s that function in PQC in human cells remains to be determined, but our work supports the idea that CRL ligases generally do not recognize hydrophobic degrons (Figure 11B). Carboxyl-terminus of Hsc70-interacting protein (CHIP) and the aforementioned E3s that function in degradation of GFP-CL1—including the Doa10 ortholog MARCH6—are the best studied human E3s that recognize hydrophobic degrons (Stefanovic-Barrett *et al.* 2018; Leto *et al.* 2019; Wang *et al.* 2020). Additional studies using a diversity of degrons will be needed to fully delineate PQC degradation pathways in humans.

Acknowledgments

We thank Christian Schlieker (Yale University) for sharing cell culture facilities, Jason Berk (Yale University) for sharing materials for cell culture, and Jianhui Li (Yale University) for assistance with microscopy. We also thank V.J. Rubenstein (Ball State University) for sharing unpublished yeast strains, Susan Michaelis (Johns Hopkins) for sharing plasmids, and Adrian Mehrtash (Yale University) for critical reading of the manuscript.

Funding

This work was supported by NIGMS grants GM046904 and GM136325 to M.H. and a training grant to C.B. (5T32GM007223).

Conflicts of interest

None declared.

Literature cited

Andress EJ, Holic R, Edelmann MJ, Kessler BM, Yu VP. 2011. Dia2 controls transcription by mediating assembly of the rsc complex. PLoS One. 6:e21172.

Ast T, Aviram N, Chuartzman SG, Schuldiner M. 2014. A cytosolic degradation pathway, prerad, monitors pre-inserted secretory pathway proteins. J Cell Sci. 127:3017–3023.

Bartel B, Wunning I, Varshavsky A. 1990. The recognition component of the n-end rule pathway. EMBO J. 9:3179–3189.

Bence NF, Bennett EJ, Kopito RR. 2005. Application and analysis of the gfpu family of ubiquitin-proteasome system reporters. Methods Enzymol. 399:481–490.

Bence NF, Sampat RM, Kopito RR. 2001. Impairment of the ubiquitin-proteasome system by protein aggregation. Science. 292:1552–1555.

Chen P, Johnson P, Sommer T, Jentsch S, Hochstrasser M. 1993. Multiple ubiquitin-conjugating enzymes participate in the in vivo degradation of the yeast mat alpha 2 repressor. Cell. 74:357–369.

Cheng H, Bao X, Rao H. 2016. The f-box protein rcy1 is involved in the degradation of histone h3 variant cse4 and genome maintenance. J Biol Chem. 291:10372–10377.

Dederer V, Khmelinskii A, Huhn AG, Okreglak V, Knop M, *et al.* 2019. Cooperation of mitochondrial and er factors in quality control of tail-anchored proteins. Elife. 8:

Doamekpor SK, Lee JW, Hepowit NL, Wu C, Charenton C, *et al.* 2016. Structure and function of the yeast listerin (ltn1) conserved n-terminal domain in binding to stalled 60s ribosomal subunits. Proc Natl Acad Sci U S A. 113:E4151–4160.

Enam C, Geffen Y, Ravid T, Gardner RG. 2018. Protein quality control degradation in the nucleus. Annu Rev Biochem. 87:725–749.

Finley D, Ulrich HD, Sommer T, Kaiser P. 2012. The ubiquitin-proteasome system of *Saccharomyces cerevisiae*. Genetics. 192:319–360.

Fredrickson EK, Gallagher PS, Clowes Candadai SV, Gardner RG. 2013. Substrate recognition in nuclear protein quality control degradation is governed by exposed hydrophobicity that correlates with aggregation and insolubility. J Biol Chem. 288:6130–6139.

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

Furth N, Gertman O, Shiber A, Alfassy OS, Cohen I, *et al.* 2011. Exposure of bipartite hydrophobic signal triggers nuclear quality control of ndc10 at the endoplasmic reticulum/nuclear envelope. Mol Biol Cell. 22:4726–4739.

Gardner RG, Nelson ZW, Gottschling DE. 2005. Degradation-mediated protein quality control in the nucleus. Cell. 120:803–815.

Geffen Y, Appleboim A, Gardner RG, Friedman N, Sadeh R, *et al.* 2016. Mapping the landscape of a eukaryotic degronome. Mol Cell. 63: 1055–1065.

Geng F, Wenzel S, Tansey WP. 2012. Ubiquitin and proteasomes in transcription. Annu Rev Biochem. 81:177–201.

Ghaboosi N, Deshaies RJ. 2007. A conditional yeast e1 mutant blocks the ubiquitin-proteasome pathway and reveals a role for ubiquitin conjugates in targeting rad23 to the proteasome. Mol Biol Cell. 18:1953–1963.

Ghislain M, Udvardy A, Mann C. 1993. S. Cerevisiae 26s protease mutants arrest cell division in g2/metaphase. Nature. 366: 358–362.

Giaever G, Chu AM, Ni L, Connelly C, Riles L, *et al.* 2002. Functional profiling of the *Saccharomyces cerevisiae* genome. Nature. 418: 387–391.

Gietz RD, Woods RA. 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 350:87–96.

Gilon T, Chomsky O, Kulka RG. 1998. Degradation signals for ubiquitin system proteolysis in *Saccharomyces cerevisiae*. EMBO J. 17: 2759–2766.

Goldstein AL, McCusker JH. 1999. Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. Yeast. 15:1541–1553.

Guerriero CJ, Weiberth KF, Brodsky JL. 2013. Hsp70 targets a cytoplasmic quality control substrate to the san1p ubiquitin ligase. J Biol Chem. 288:18506–18520.

Guldener U, Heck S, Fielder T, Beinhauer J, Hegemann JH. 1996. A new efficient gene disruption cassette for repeated use in budding yeast. Nucleic Acids Res. 24:2519–2524.

Haakonsen DL, Rape M. 2019. Branching out: Improved signaling by heterotypic ubiquitin chains. Trends Cell Biol. 29:704–716.

Haber JE. 2012. Mating-type genes and mat switching in *Saccharomyces cerevisiae*. Genetics. 191:33–64.

Heck JW, Cheung SK, Hampton RY. 2010. Cytoplasmic protein quality control degradation mediated by parallel actions of the e3 ubiquitin ligases ubr1 and san1. Proc Natl Acad Sci U S A. 107: 1106–1111.

Hickey CM. 2016. Degradation elements coincide with cofactor binding sites in a short-lived transcription factor. Cell Logist. 6: e1157664.

Hickey CM, Hochstrasser M. 2015. Stubl-mediated degradation of the transcription factor matalpha2 requires degradation elements that coincide with corepressor binding sites. Mol Biol Cell. 26: 3401–3412.

Hickey CM, Xie Y, Hochstrasser M. 2018. DNA binding by the matalpha2 transcription factor controls its access to alternative ubiquitin-modification pathways. Mol Biol Cell. 29:542–556.

Hochstrasser M. 2009. Origin and function of ubiquitin-like proteins. Nature. 458:422–429.

Hochstrasser M, Varshavsky A. 1990. In vivo degradation of a transcriptional regulator: the yeast alpha 2 repressor. Cell. 61: 697–708.

Johnson PR, Swanson R, Rakhilina L, Hochstrasser M. 1998. Degradation signal masking by heterodimerization of matalpha2 and mata1 blocks their mutual destruction by the ubiquitin-proteasome pathway. Cell. 94:217–227.

Johnston HE, Samant RS. 2020. Alternative systems for misfolded protein clearance: Life beyond the proteasome. FEBS J., doi: 10.1111/febs.15617, in press

Kaganovich D, KR, FJ. 2008. Misfolded proteins partition between two distinct quality control compartments. Nature. 454: 1088–1095.

Kats I, Khmelinskii A, Kschonsak M, Huber F, Knieß RA, *et al.* 2018. Mapping degradation signals and pathways in a eukaryotic n-terminome. Mol Cell. 70:488–501.e5.

Ke A, Wolberger C. 2003. Insights into binding cooperativity of mata1/matalpha2 from the crystal structure of a mata1 homeodomain-maltose binding protein chimera. Protein Sci. 12: 306–312.

Khosrow-Khavar F, Fang NN, Ng AH, Winget JM, Comyn SA, Mayor T. 2012. The yeast ubr1 ubiquitin ligase participates in a prominent pathway that targets cytosolic thermosensitive mutants for degradation. G3 (Bethesda). 2:619–628.

Kim I, Miller CR, Young DL, Fields S. 2013. High-throughput analysis of in vivo protein stability. Mol Cell Proteomics. 12: 3370–3378.

Koren I, Timms RT, Kula T, Xu Q, Li MZ, *et al.* 2018. The eukaryotic proteome is shaped by e3 ubiquitin ligases targeting c-terminal degrons. Cell. 173:1622–1635.e14.

Laney JD, Hochstrasser M. 2003. Ubiquitin-dependent degradation of the yeast mat(alpha)2 repressor enables a switch in developmental state. Genes Dev. 17:2259–2270.

Leto DE, Morgens DW, Zhang L, Walczak CP, Elias JE, *et al.* 2019. Genome-wide crispr analysis identifies substrate-specific conjugation modules in er-associated degradation. Mol Cell. 73: 377–389.e11.

Liu G, Yong MY, Yurieva M, Srinivasan KG, Liu J, *et al.* 2015. Gene essentiality is a quantitative property linked to cellular evolvability. Cell. 163:1388–1399.

Liu H, Krizek J, Bretscher A. 1992. Construction of a gal1-regulated yeast cdna expression library and its application to the identification of genes whose overexpression causes lethality in yeast. Genetics. 132:665–673.

Longtine MS, McKenzie A, 3rd, Demarini DJ, Shah NG, Wach A, *et al.* 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. Yeast. 14:953–961.

Maurer MJ, Spear ED, Yu AT, Lee EJ, Shahzad S, *et al.* 2016. Degradation signals for ubiquitin-proteasome dependent cytosolic protein quality control (cytoqc) in yeast. G3 (Bethesda). 6: 1853–1866.

Mehrtash AB, Hochstrasser M. 2019. Ubiquitin-dependent protein degradation at the endoplasmic reticulum and nuclear envelope. Semin Cell Dev Biol. 93:111–124.

Mumberg D, Muller R, Funk M. 1994. Regulatable promoters of *Saccharomyces cerevisiae*: Comparison of transcriptional activity and their use for heterologous expression. Nucl Acids Res. 22: 5767–5768.

Mumberg D, Muller R, Funk M. 1995. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene. 156:119–122.

Nixon CE, Wilcox AJ, Laney JD. 2010. Degradation of the *Saccharomyces cerevisiae* mating-type regulator alpha1: Genetic dissection of cis-determinants and trans-acting pathways. Genetics. 185:497–511.

Okada M, Ohtake F, Nishikawa H, Wu W, Saeki Y, *et al.* 2015. Liganded eralpha stimulates the e3 ubiquitin ligase activity of ube3c to facilitate cell proliferation. Mol Endocrinol. 29: 1646–1657.

Pohl C, Dikic I. 2019. Cellular quality control by the ubiquitin-proteasome system and autophagy. Science. 366: 818–822.

Prasad R, Kawaguchi S, Ng DT. 2010. A nucleus-based quality control mechanism for cytosolic proteins. Mol Biol Cell. 21:2117–2127.

Prasad R, Xu C, Ng DTW. 2018. Hsp40/70/110 chaperones adapt nuclear protein quality control to serve cytosolic clients. J Cell Biol. 217:2019–2032.

Ravid T, Hochstrasser M. 2007. Autoregulation of an e2 enzyme by ubiquitin-chain assembly on its catalytic residue. Nat Cell Biol. 9: 422–427.

Ravid T, Kreft SG, Hochstrasser M. 2006. Membrane and soluble substrates of the doa10 ubiquitin ligase are degraded by distinct pathways. EMBO J. 25:533–543.

Rusnac DV, Lin HC, Canzani D, Tien KX, Hinds TR, *et al.* 2018. Recognition of the diglycine c-end degron by crl2(klhdc2) ubiquitin ligase. Mol Cell. 72:813–822 e814.

Samant RS, Livingston CM, Sontag EM, Frydman J. 2018. Distinct proteostasis circuits cooperate in nuclear and cytoplasmic protein quality control. Nature. 563:407–411.

Seufert W, Jentsch S. 1990. Ubiquitin-conjugating enzymes ubc4 and ubc5 mediate selective degradation of short-lived and abnormal proteins. EMBO J. 9:543–550.

Sopko R, Huang D, Preston N, Chua G, Papp B, *et al.* 2006. Mapping pathways and phenotypes by systematic gene overexpression. Mol Cell. 21:319–330.

Stefanovic-Barrett S, Dickson AS, Burr SP, Williamson JC, Lobb IT, *et al.* 2018. March6 and trc8 facilitate the quality control of cytosolic and tail-anchored proteins. EMBO Rep. 19:

Stoll KE, Brzovic PS, Davis TN, Klevit RE. 2011. The essential ubc4/ubc5 function in yeast is hect e3-dependent, and ring e3-dependent pathways require only monoubiquitin transfer by ubc4. J Biol Chem. 286:15165–15170.

Stolz A, Besser S, Hottmann H, Wolf DH. 2013. Previously unknown role for the ubiquitin ligase ubr1 in endoplasmic reticulum-associated protein degradation. Proc Natl Acad Sci U S A. 110:15271–15276.

Summers DW, Wolfe KJ, Ren HY, Cyr DM. 2013. The type ii hsp40 sis1 cooperates with hsp70 and the e3 ligase ubr1 to promote degradation of terminally misfolded cytosolic protein. PLoS One. 8: e52099.

Swanson R, Locher M, Hochstrasser M. 2001. A conserved ubiquitin ligase of the nuclear envelope/endoplasmic reticulum that functions in both er-associated and matalpha2 repressor degradation. Genes Dev. 15:2660–2674.

Tasaki T, Mulder LC, Iwamatsu A, Lee MJ, Davydov IV, *et al.* 2005. A family of mammalian e3 ubiquitin ligases that contain the ubr box motif and recognize n-degrons. Mol Cell Biol. 25:7120–7136.

Tong AH, Boone C. 2006. Synthetic genetic array analysis in *Saccharomyces cerevisiae*. Methods Mol Biol. 313:171–192.

Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, *et al.* 2001. Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science. 294:2364–2368.

Uzunova K, Gottsche K, Miteva M, Weisshaar SR, Glanemann C, *et al.* 2007. Ubiquitin-dependent proteolytic control of sumo conjugates. J Biol Chem. 282:34167–34175.

Varshavsky A. 2012. The ubiquitin system, an immense realm. Annu Rev Biochem. 81:167–176.

Wang T, Wang W, Wang Q, Xie R, Landay A, *et al.* 2020. The e3 ubiquitin ligase chip in normal cell function and in disease conditions. Ann N Y Acad Sci. 1460:3–10.

Watts SG, Crowder JJ, Coffey SZ, Rubenstein EM. 2015. Growth-based determination and biochemical confirmation of genetic requirements for protein degradation in *Saccharomyces cerevisiae*. J Vis Exp. 96:52428.

Wingfield PT. 2017. N-terminal methionine processing. Curr Protoc Protein Sci. 88:6 14 11–16 14 13.

Xie Y, Rubenstein EM, Matt T, Hochstrasser M. 2010. Sumo-independent in vivo activity of a sumo-targeted ubiquitin ligase toward a short-lived transcription factor. Genes Dev. 24: 893–903.

Yau RG, Doerner K, Castellanos ER, Haakonsen DL, Werner A, *et al.* 2017. Assembly and function of heterotypic ubiquitin chains in cell-cycle and protein quality control. Cell. 171:918–933 e920.

Yoshikawa K, Tanaka T, Ida Y, Furusawa C, Hirasawa T, *et al.* 2011. Comprehensive phenotypic analysis of single-gene deletion and overexpression strains of *Saccharomyces cerevisiae*. Yeast. 28: 349–361.

Communicating editor: O. Cohen-Fix