Substrate-Binding Sites of UBR1, the Ubiquitin Ligase of the N-End Rule Pathway*

Zanxian Xia‡¶, Ailsa Webster†, Fangyong Piatkov‡, Michel Ghislain¥, and Alexander Varshavsky‡1

‡From Division of Biology, California Institute of Technology, Pasadena, CA 91125, †Celltech R&D, 216 Bath Road, Slough SL1 4EN, U.K., §Department of Microbial Pathogenesis, Yale University, New Haven, CT 06536-0812, ¥Group of Physiological Biochemistry, University of Louvain, Croix du Sud 5/15, B-1348 Louvain-la-Neuve, Belgium

1Address correspondence to: Alexander Varshavsky, Division of Biology, California Institute of Technology, Pasadena, CA 91125. Tel.: 626-395-3785; Fax: 626-440-9821; Email: avarsh@caltech.edu

¶Present address: Center for Stem Cell and Regenerative Medicine, University of Southern California, Los Angeles, CA 90033.

*This research was supported by grants to A.V. from the National Institutes of Health (GM31530, DK39520), the Ellison Medical Foundation, and the Sandler Program in Asthma Research.

Running head: UBR1 ubiquitin ligase

Key words: / proteolysis / ubiquitin / yeast / Saccharomyces cerevisiae / substrate recognition /

Substrates of a ubiquitin-dependent proteolytic system called the N-end rule pathway include proteins with destabilizing N-terminal residues. N-recognins, the pathway’s ubiquitin ligases, contain three substrate-binding sites. The type-1 site is specific for basic N-terminal residues (Arg, Lys, His). The type-2 site is specific for bulky hydrophobic N-terminal residues (Trp, Phe, Tyr, Leu, Ile). We show here that the type-1/2 sites of UBR1, the sole N-recognin of the yeast Saccharomyces cerevisiae, are located in the first ~700 residues of the 1,950-residue UBR1. These sites are distinct in that they can be selectively inactivated by mutations, identified through a genetic screen. Mutations inactivating the type-1 site are in the previously delineated ~70-residue UBR motif characteristic of N-recognins. Fluorescence polarization and surface plasmon resonance were used to determine that UBR1 binds, with $K_d$ of ~1 $\mu$M, to either type-1 or type-2 destabilizing N-terminal residues of reporter peptides, but does not bind to a stabilizing N-terminal residue such as Gly. A third substrate-binding site of UBR1 targets an internal degron of CUP9, a transcriptional repressor of peptide import. We show that the previously demonstrated in vivo dependence of CUP9 ubiquitylation on the binding of cognate
dipeptides to the type-1/2 sites of UBR1 can be reconstituted in a completely defined in vitro system. We also found that purified UBR1 and CUP9 interact nonspecifically, and that specific binding (which involves, in particular, the binding by cognate dipeptides to the UBR1’s type-1/2 sites) can be restored either by a chaperone such as EF1A or through macromolecular crowding.

The N-end rule relates the in vivo half-life of a protein to the identity of its N-terminal residue. The ubiquitin (Ub)-dependent N-end rule pathway recognizes several kinds of degradation signals (degrons), including a set called N-degrons (Fig. 1A) (1-9). Although prokaryotes lack the Ub system, they contain the N-end rule pathway, Ub-independent versions of it (10-18). In eukaryotes, an N-degron consists of three determinants: a destabilizing N-terminal residue of a protein substrate, an internal Lys residue(s) of the substrate (the site of formation of a poly-Ub chain), and a conformationally flexible region in the vicinity of these determinants (1,19-23). The N-end rule has a hierarchic structure (Fig. 1A). In eukaryotes, N-terminal Asn and Gln are tertiary destabilizing residues in that they function through their enzymatic deamidation, to yield the secondary destabilizing N-terminal residues Asp and Glu (24,25). The destabilizing activity of N-terminal Asp and Glu requires their conjugation to Arg, one of the primary destabilizing residues, by Arg-tRNA-protein transferase (arginyl-transferase or R-transferase) (5-7,26,27). In mammals and other eukaryotes that produce nitric oxide (NO), the set of arginylated residues contains not only Asp and Glu but also N-terminal Cys (28), which is arginylated after its oxidation to Cys-sulfinate or Cys-sulfonate (27). The in vivo oxidation of N-terminal Cys requires NO, as well as oxygen (O2) or its derivatives (Fig. 1A) (6,7,29). The N-end rule pathway is thus a sensor of NO, through its ability to destroy proteins with N-terminal Cys, at rates controlled by NO, and by oxygen as well.

E3 Ub ligases of the N-end rule pathway, called N-recognins (1,8,29-37), recognize (bind to) primary destabilizing N-terminal residues, including Arg (Fig. 1A). (The term “Ub ligase” denotes either an E2-E3 holoenzyme or its E3 component.) At least four N-recognins, including UBR1, mediate the N-end rule pathway in mammals (30-33,36). The known N-recognins share a ~70-residue motif called the UBR box. Mouse UBR1 and UBR2 are sequelogous (similar in sequence) 200-kDa RING-type E3 Ub ligases that are 47% identical. Several other UBR-containing N-recognins, either confirmed or putative ones, are HECT-type or SCF-type E3 Ub ligases that share the UBR motif with the RING-type UBR1/UBR2 but are largely nonsequelogous to them otherwise (29-33).

[“Sequelog” and “spalog” denote, respectively, a sequence that is similar, to a specified extent, to another sequence, and a 3D structure that is similar, to a specified extent, to another 3D structure (38). The rigor-conferring advantage of “sequelog” and “spalog” is their evolutionary neutrality, in contrast to interpretation-laden terms such as “homolog”, “ortholog” and “paralog”. The latter terms are compatible with the sequelog/spalog terminology and can be used to convey understanding about functions and common descent, if this (additional) information is available (38).]

The N-end rule pathway of S. cerevisiae is mediated by a single N-recogin, UBR1, a 225-kDa sequelog of mammalian UBR1 and UBR2 (Fig. 1A, C) (8,9). S. cerevisiae UBR1 (Fig. 1C) contains at least three substrate-binding sites. The type-1 site is specific for basic N-terminal residues of protein substrates (Arg, Lys, His), while the type-2 site is specific for bulky hydrophobic N-terminal residues (Trp, Phe, Tyr, Leu, Ile). The third binding site of UBR1 targets proteins through their internal (non-N-terminal) degrons, and is allosterically “activated” through a conformational change that is caused by the binding of short peptides to the UBR1’s other two binding sites, type-1 and type-2. The known substrate of the third binding site of UBR1 is CUP9 (8,39,40), a gene encoding transporter of di- and tripeptides. The reversal of UBR1 autoinhibition by (initially) imported peptides accelerates the UBR1-dependent ubiquitylation of CUP9, results in its faster degradation by the N-end rule pathway, and thereby causes derepression of
PTR2 and increased levels of the PTR2 transporter. This positive-feedback circuit allows *S. cerevisiae* to detect the presence of extracellular peptides and to react by increasing their uptake (8,40). ClpS, a 12-kDa prokaryotic N-recognin, is an “adaptor” protein that mediates the (Ub-independent) targeting of N-end rule substrates in bacteria (13-15). ClpS recognizes type-2 (bulky hydrophobic) N-terminal residues, and contains a region of sequelogy to the 225 kDa yeast UBR1, near its UBR box (2,13,42). This similarity and other common features of eukaryotic and prokaryotic N-end rule pathways (1,2,29) suggest that at least some N-degrons, N-recognins and relevant “downstream” proteases had evolved before the split between eukaryotes and prokaryotes.

The functions of the N-end rule pathway include (i) the sensing of heme, owing to inhibition of the pathway’s ATE1 R-transferase (in both yeast and mammals) by hemin (Fe³⁺-heme), which also inhibits N-recognins, the latter at least in yeast (4) (Fig. 1A); (ii) the sensing of NO and oxygen, and the resulting control of signaling by transmembrane receptors, through the conditional, NO/O₂-mediated degradation of G-protein regulators RGS4, RGS5 and RGS16 (6,7); (iii) regulation of import of short peptides, through the degradation, modulated by peptides, of CUP9, the import’s repressor (8,40); (iv) fidelity of chromosome segregation, through degradation of a separase-produced cohesin fragment (43); (v) regulation of apoptosis, through degradation of a caspase-processed inhibitor of apoptosis (44,45); (vi) a multitude of processes mediated by the transcription factor c-FOS, a conditional substrate of the N-end rule pathway (46); (vii) regulation of the human immunodeficiency virus (HIV) replication cycle, through degradation of HIV integrase (47); (viii) regulation of meiosis, spermatogenesis, neurogenesis and cardiovascular development in mammals (6,7,27,31,33), and the leaf senescence in plants (48). Mutations in human UBR1 (Fig. 1A) (29,32,33) are the cause of Johansson-Blizzard Syndrome (JBS), which comprises mental retardation, physical malformations, and severe pancreatitis (49). The abnormalities of *UBR1*<sup>−/−</sup> mice (30) include pancreatic insufficiency (49), a less severe counterpart of this defect in human JBS (*UBR1<sup>−/−</sup>*) patients.

In the present work, we used a genetic screen to probe the type-1 and type-2 substrate-binding sites of *S. cerevisiae* UBR1. This analysis demonstrated a modular organization of these sites, which could be selectively inactivated by specific mutations. We also characterized physical interactions between UBR1 and CUP9 that involve the internal (non-N-terminal) degron of CUP9, and employed a completely defined *in vitro* system to probe the conditional (modulated by dipeptides) polyubiquitylation of CUP9 by the UBR1-RAD6 Ub ligase. In addition, we used both fluorescence polarization (FP) and surface plasmon resonance (SPR) to determine the affinities of purified UBR1 for specific N-terminal residues of model peptides. The results indicated, among other things, that at least the bulk of the observed *in vivo* specificity of UBR1 for N-end rule substrates (1,2,29) (Fig. 1A, C) stems from differences in the physical affinity of UBR1 for destabilizing versus stabilizing N-terminal residues in the N-end rule.

**EXPERIMENTAL PROCEDURES**

*Yeast Strains, Media and Genetic Techniques* – Synthetic yeast media (50) contained 0.67% yeast nitrogen base without amino acids (Difco) and either 2% glucose (SD medium) or 2% galactose (SG medium). Synthetic media lacking appropriate nutrients were used to select for (and maintain) specific plasmids. Cells were also grown in rich medium (YPD) (51). The *S. cerevisiae* strains were YPH500 (*MATα* *ura3*-52 *lys2*-801 *ade2*-101 *his3*-Δ200 *trp1*-1A *his3*-Δ200 *leu2*-Δ1), JD52 (*MATα* *trp1*-Δ63 *ura3*-52 *his3*-Δ200 *leu2*-3,112 *lys2*-801), JD55 (*MATα* *ubr1Δ::His3 trp1*-Δ63 *ura3*-52 *his3*-Δ200 *leu2*-3,112 *lys2*-801) (52) SC295 (*MATα, GAL4 GAL80 ura3*-52, *leu2*-3,112 *reg1*-501 *gal1 pep4*-3) (8) and PJ69-4A (*MATα* *trp1*-901 *leu2*-3,112 *ura3*-52 *his3*-200 *gal4Δ gal80Δ GAL2-ADE2 LYS2::GAL1-HIS3 met2::GAL7-lacZ) (53). The medium for dipeptide import assays was methionine-lacking SD containing allantoin (1 mg/ml) and the toxic dipeptide L-leucyl-L-ethionine (Leu-Eth; 37 μM) (39). To induce the *P<sub>CUP9</sub>* promoter, CuSO₄ was added to a final concentration of 0.1
mM (52). The XGal colony overlay assay was performed using 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (XGal) at 0.1 mg/ml, as described (53). Transformation of S. cerevisiae was carried out using the lithium acetate method (50).

**Reporter Plasmids** – Five sets of reporter plasmids were used. Details of their construction are available upon request. pUB23-X high copy plasmids expressed Ub-X-eK-β-galactosidase fusions (denoted below as Ub-X-βgal) from the galactose-inducible P\(_{\text{GAL1}}\) promoter (19,54). pBAX plasmids expressed Ub-X-eK-ha-URA3 fusions (X=Arg, Leu), denoted below as Ub-X-URA3. The largely cotranslational deubiquitylation of these Ub fusions (55) yielded the corresponding X-βgal and X-URA3 reporter proteins. pBAX plasmids were constructed from pBAM, a pRS314-based low copy plasmid expressing Ub-Met-eK-ha-URA3 (tagged with the ha epitope (50)) from the P\(_{\text{CUP1}}\) promoter (56). The term e\(^K\) (extension (K)) denotes a ~40-residue, E. coli Lac repressor-derived sequence (21,54,56). pDSRULB was constructed by subcloning the 5.4-kb fragment from pUB23-L (encoding Ub-Leu-URA3) into the Smal site of pBAR, encoding Ub-Arg-URA3. The resulting low copy, TRP1-marked pRS314 plasmid expressed Ub-Arg-URA3 from the P\(_{\text{CUP1}}\) promoter and Ub-Leu-βgal from the P\(_{\text{GAL1}}\) promoter. The similarly constructed pDSLURB expressed Ub-Leu-URA3 and Ub-Arg-βgal. The UPR (Ub/protein/reference) plasmids (21,40,57) for expression in S. cerevisiae, termed pBAXUPR, were based on the low copy pRS314 vector (58), and were constructed using PCR. A pBAXUPR plasmid expressed DHFRha-Ub\(^{R48}\)-X-eK-ha-URA3 (denoted below as DHFR-Ub-X-URA3) from the P\(_{\text{CUP1}}\) promoter (X=Met, Arg, Leu, Glu, Lys, His, Ile, Phe, Tyr, Trp). The analogous UPR-based plasmids, termed pXβgalUPR, expressed DHFRha-Ub\(^{R48}\)-X-eK-βgal (X=Met, Arg, Leu), denoted as DHFR-Ub-X-βgal. pXβgalUPR were constructed by cutting an appropriate pBAXUPR plasmid with BamHI and Smal, recovering the vector-containing fragment, and ligating it to the 3.8 kb βgal-encoding fragment produced from pUB23-M using BamHI and ScaI.

**UBR1 Plasmids** – S. cerevisiae UBR1 plasmids encoding wild-type or mutated UBR1 included pRBUBR1, pRB208, pFlagUBR1, pFlagNT1UBR1, pUBR1NT1140Flag, p209NTUBR1Flag, p454NTUBR1Flag, p710NTUBR1Flag, p866NTUBR1Flag, and p1002NTUBR1Flag. Details of construction are available upon request. UBR1 alleles were expressed from the P\(_{\text{ADHI}}\) promoter of the high copy, LEU2-marked Yeplac181 vector (59), in SC295, a protease-deficient S. cerevisiae strain. UBR1 and its derivatives that were encoded by the above plasmids carried either the ha epitope (pRBUBR1 and pRB208) or the flag epitope, pRB208 and pRBUBR1 (a derivative of pRB208) encoded full-length, C-terminally ha-tagged S. cerevisiae UBR1 (UBR1\(^{HA}\)). pFlagUBR1 encoded \(^\text{i}\)UBR1, the N-terminally flag-tagged full-length UBR1. pFlagNT1UBR1 encoded \(^\text{i}\)UBR1\(^{710-1140}\), the N-terminally flag-tagged C-terminally truncated variants of UBR1. The design and expression of \(^\text{i}\)UBR1 and UBR1\(^{1-717}\) were described (8). The plasmid p209NTUBR1Flag, p454NTUBR1Flag, p710NTUBR1Flag, p866NTUBR1Flag, and p1002NTUBR1Flag, which encoded, respectively, UBR1\(^{209-1140}\), UBR1\(^{454-1140}\), UBR1\(^{710-1140}\), \(^\text{i}\)UBR1\(^{866-1140}\), and \(^\text{i}\)UBR1\(^{1002-1140}\), were derivatives of pUBR1NT1140Flag (8), a high copy plasmid that expressed C-terminally flag-tagged UBR1\(^{1-1140}\) from the yeast P\(_{\text{ADHI}}\) promoter. The flag tag was located at the C-terminus of the UBR1 derivatives except for the full-length \(^\text{i}\)UBR1 and \(^\text{i}\)UBR1\(^{1-717}\). UBR1 test proteins remained intact in yeast extracts that were prepared and incubated as previously described (8).

**Expression and Purification of X-SCC1-GST** – A set of 3 pTYB12-X-SCC1-GST plasmids, which expressed Intein-CBD-X-SCC1-GST fusions (X=Arg, Leu or Met), was employed for glutathione-S-transferase (GST)-pulldown assays. Details of construction are available upon request. A specific pTYB12-X-SCC1-GST plasmid was cotransformed into E. coli BL21(DE3) together with pRI952, which expressed tRNAs for the codons AGG, AGA and AUA that are rare in E. coli (60). A fresh E. coli colony was inoculated into LB medium (600 ml) containing ampicillin (0.1 mg/ml) and chloramphenicol (25 μg/ml), followed by growth
at 37°C until OD_{600} of 0.5-0.8. Isopropyl-β-D-thiogalactopyranoside (IPTG) was then added to the final concentration of 0.4 mM, followed by a 10-hr incubation, with shaking, at 18°C. The cells were centrifuged at 5,000 g for 10 min at 4°C, washed once with phosphate-buffered saline (PBS), and frozen in liquid N₂. For isolation of proteins, the cell pellet was thawed at 0°C, then resuspended in lysis buffer (0.1% Triton X-100, 0.5 M NaCl, 20 mM Tris-HCl, pH 8.5) containing protease inhibitor cocktail (Roche, Indianapolis, IN) (6 ml of buffer per 1 g of pellet), and cells were disrupted by sonication on ice. The suspension was centrifuged at 13,000 g for 30 min. The supernatant was loaded onto a pre-equilibrated chitin column (10-ml bed volume), at a rate not higher than 0.5 ml/min, and washed with 200 ml of column buffer (0.5 M NaCl, 1 mM EDTA, 20 mM Tris-HCl, pH 8.5), at 1 ml/min. The column was then flushed quickly with 30 ml of freshly prepared cleavage buffer (column buffer containing 40 mM dithiothreitol (DTT)), the flow was stopped, and the column was incubated at 23°C for 16 hr. X-SCC1-GST proteins were eluted with 30 ml of column buffer, concentrated with Centriplus (Millipore, Billerica, MA), and dialyzed against column buffer at 4°C. The resulting samples were stored in small aliquots at -80°C. The presence of desired N-terminal residues in the purified X-SCC1-GST proteins was verified using N-terminal sequencing by Edman degradation. GST-CUP9 was expressed and purified as described (8).

**Expression and Purification of fUBR1 and UBR1^{1-1140f}** – To express and purify the C-terminally flag-tagged UBR1^{1-1140f}, *S. cerevisiae* transformed with the pNTFlagUBR1 was inoculated into 20 ml of yeast SD medium, and grown at 30°C to A_{600} of ~1. The 20-ml culture was re-inoculated into 2 l of SD medium and grown to A_{600} of ~1, followed by the addition of equal volume of yeast YPD medium. The cells were grown for ~3 more generations, until A_{600} of ~4, harvested by centrifugation, washed once with cold 1xPBS, and frozen in liquid N₂.

Other procedures, including purification, were identical to those with UBR1^{1-1140f}.

Concentrations of purified fUBR1 and UBR1^{1-1140f} were determined using the Protein Assay reagent (Bio-Rad, Hercules, CA). Samples of purified fUBR1 and UBR1^{1-1140f} were frozen in liquid N₂ and stored at -80°C.

**GST-Pulldown Assays** – An SCC1-based GST-fusion protein (~2 μg) (see above) was diluted to 0.3 ml in 1XPBS buffer (10% glycerol, 1% Triton X-100, 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 1.4 mM KH₂PO₄, pH 7.4), and incubated with 20 μl of glutathione-Sepharose-4B beads (Amersham Pharmacia Biotech) for 30 min at 4°C. The beads were washed once with 0.5 ml of the 1XPBS buffer and twice with 0.5 ml of binding buffer (10% glycerol, 0.05% Nonidet-P40 (NP40), 0.2 M KCl, 50 mM HEPES, pH 7.5). The washed beads, in 0.15 ml of the binding buffer, were incubated at 4°C for 1 hr with 0.1 ml of *S. cerevisiae* extract (~5 mg/ml total protein) containing either fUBR1 (N-terminally flag-tagged UBR1) or its truncated derivatives (see above) at the same concentration. These assays were performed in either the presence or absence of specific amino acids or amino acid derivatives, including dipeptides (Sigma), at concentrations indicated in figure legends. Binding assays also contained ovalbumin (Sigma; at 0.1 mg/ml, unless not stated otherwise), protease inhibitor cocktail (Sigma), and 50 μM bestatin (Sigma). The beads were washed twice with 0.2 ml of the binding buffer either containing or lacking specific dipeptides, depending on whether a binding assay contained them. Washed beads were resuspended in 30 μl of SDS-PAGE loading buffer, heated at 95°C for 4 min, and 15-μl samples were subjected to SDS-10% PAGE, followed by immunoblotting with anti-flag M2 antibody (Sigma).
CUP9-based GST-pulldown assays with 1UBR1 were carried out similarly except for the following: the total volume was 0.26 ml, with 1 μl of purified 1UBR1 (0.4 mg/ml), either in binding buffer (10% glycerol, 0.05% NP40, 50 mM NaCl, 50 mM HEPES, pH 7.5), or in binding buffer supplemented with 60 μl of S. cerevisiae “empty” extract (no flag-tagged UBR1), or in GST-binding buffer containing ovalbumin (at 10 mg/ml); or in 0.2 ml of binding buffer supplemented with 60 μl of S. cerevisiae extract containing 1UBR1. With UBR11-1140f, the assay was performed similarly, except that the reaction volume was 0.23 ml (0.2 ml of binding buffer plus 30 μl of extract from S. cerevisiae that produced UBR11-1140f or, alternatively, the same (30 μl) volume but containing 1 μl of purified UBR11-1140f (at 0.2 mg/ml), and ovalbumin at 50 mg/ml, as described in figure legends.

**X-Peptide Pulldown Assays – Procedures**
were similar to those previously described (8,29,31,32). Residues 2-9 of synthetic 12-mer X-peptides (X-Ile-Phe-Ser-Thr-Asp-Thr-Gly-Pro-Gly-Gly-Cys, where X=Arg, Phe or Gly), were derived from residues 2-9 of Sindbis virus polymerase nsP4, a previously identified N-end rule substrate (1,61). Hence the acronym for a specific X-peptide, “X-nsP4pep”. 20 μl (5 mg/ml total protein) of S. cerevisiae extract containing (overexpressed) S. cerevisiae UBR1 or its fragments (indicated above) was diluted into 0.2 ml of binding buffer (10% glycerol, 0.05% NP40, 0.2 M KCl, 50 mM HEPES, pH 7.5) in the presence of 0.1 mg/ml ovalbumin, protease inhibitor cocktail (Sigma) and 50 μM bestatin. The samples were incubated with 5 μl (packed volume) of a carrier-linked X-nsP4pep for 1 hr at 4°C. The beads were pelleted by a brief centrifugation, then washed with 0.3 ml of binding buffer three times. The beads were then resuspended in 30 μl of SDS-PAGE sample buffer, and heated at 95°C for 4 min, followed by a brief centrifugation in a microcentrifuge, SDS-10% PAGE of 3-μl samples of supernatants, and detection of flag-tagged UBR1 or its fragments by immunoblotting with anti-flag M2 antibody.

**Labeling, Purification and Analysis of X-Peptides** – The above peptides were also employed as probes in the fluorescence polarization (FP) assay. To conjugate Alexa Fluor-488-C5 maleimide (Invitrogen, Carlsbad, CA) to C-terminal Cys of an X-nsP4pep, this dye (0.6 μmole in 90 μl of dimethylsulfoxide (DMSO)) was added to 0.4 μmole of X-nsP4pep in 90 μl of DMSO, followed by incubation at room temperature for 4 hr. Alexa-labeled X-peptides were purified by HPLC, using analytical-scale VYDAC column (C18 reverse phase), followed by verification of purified peptides by mass spectrometry. The concentration of labeled X-peptides was measured in using either BCA reagent (Pierce, Rockford, IL), or N-terminal sequencing by Edman degradation, or amino acid analysis, with similar results.

**Fluorescence Polarization Assay** – The concentration of fluorescently labeled X-peptides was kept at 10 nM. Before carrying out the FP assay, we ascertained that the levels of background fluorescence of the binding buffer (10% glycerol, 0.02% NP40, 0.2 M KCl, 50 mM HEPES, pH 7.1), of the competitor (unlabeled) dipeptides or X-peptides, and of the UBR11-1140f protein (in the same buffer), at their highest concentrations used, were low enough, less than 5% of the fluorescence of a labeled X-nsP4pep in the assay. The FP assay (62,63) was carried out as described in the published protocol (“Fluorescence Polarization: Technical Resource Guide” (PanVera Corporation, now a part of Invitrogen)), with the following modifications. Purified 1UBR11-1140f was serially diluted into ~14 microcentrifuge tubes, covering the range of its concentrations from more than 10-fold below (estimated) dissociation constant (Kd) of UBR11-1140f interactions with a cognate peptide to more than 10-fold above estimated Kd. Identical samples of an Alexa-labeled X-nsP4pep were added to each of the above tubes, gently mixed, and incubated at 4°C for 30 min, in the final volume of 50 μl. FP measurements were carried out using the Analyst AD&HT Assay Detection System (Molecular Devices, Sunnyvale, CA) and a xenon-arc lamp as a light source, in a standard FP configuration, with filter settings optimized for the excitation wavelength of 485 nm and the emission wavelength of 530 nm. Two fluorescence measurements were collected for each well, with either parallel (S and S) or perpendicular (S and P) polarizers. These data were used to calculate polarization, in
millipolarization units (mP), by the Analyst System. The affinities, expressed as K_d's, were calculated from the above data. Their statistical significance was estimated by nonlinear regression analysis, using Prism (Graphpad, San Diego, CA). FP-based competition assays were carried out similarly, except that concentration of UBR1\textsuperscript{1-1140f} was kept constant, and (unlabeled) competitor X-peptides were added to the binding assay. The concentration of UBR1\textsuperscript{1-1140f} in competition FP assays was close to K_d (~1 μM; see Results) of the UBR1\textsuperscript{1-1140f} interaction with Arg-nsP4pep.

Surface Plasmon Resonance BIACORE Assay – The X-peptides Phe-nsP4pep, Arg-nsP4pep or Gly-nsP4pep were linked, through their C-terminal Cys (see above), to biotin, using EZ-Link PEO-Iodoacetyl Biotin (Pierce), to allow immobilization of peptides onto streptavidin-coated SA sensor chips of BIACORE-2000 (Biacore). Biotin-labeled X-peptides were purified by HPLC, using analytical-scale VYDAC column (C18 reverse phase), followed by verification of purified peptides by MALDI-TOF. The concentration of labeled X-peptides was measured using EZTM-Biotin Quantification Kit (Pierce). A biotin-containing X-peptide, at 0.1 μM in the running buffer (0.02% NP-40, 0.2 M KCl, 50 mM HEPES pH 7.5), was injected over the sensor chip surface at 10 μl min\textsuperscript{-1} at 25°C, resulting in X-peptide’s surface density of 30-35 resonance units. Purified UBR1\textsuperscript{1-1140f} was serially diluted in running buffer and injected at 75 µl min\textsuperscript{-1} for 80 sec at 25°C. After 5 min, to allow dissociation of the formed complexes, the sensor chip was washed by two 1-min injections of 2 M MgCl\textsubscript{2}, at 50 µl/min. BIACORE readouts were corrected by subtracting readouts from the mock surface (lacking an X-peptide) as well as readouts from a buffer-only (control) injection. Raw data were processed using Scrubber (BioLogic Software, Australia) and CLAMP programs and the 1:1 Langmuir interaction model, with a correction for mass transport (64,65).

UBR1-Dependent In Vitro Ubiquitylation System – This completely defined system, consisting exclusively of purified recombinant proteins and small compounds, was described in ref. (40), and in detail in ref. (66). Briefly, C-terminally His\textsubscript{6}-tagged \textit{S. cerevisiae} UBA1 (E1, Ub-activating enzyme) was overexpressed in \textit{S. cerevisiae} SC295 and purified using Ni-NTA (His\textsubscript{6}-affinity) chromatography, immobilized Ub (“covalent” affinity) chromatography and Superdex-200 gel filtration (67). \textit{S. cerevisiae} RAD6 (UBC2) was overexpressed in \textit{E. coli} and purified by fractionation over DEAE, Mono-Q and Superdex-75 columns. N-terminally flag-tagged \textit{S. cerevisiae} UBR1 (\text{\textsuperscript{1}}UBR1) was overexpressed in \textit{S. cerevisiae} as described in a section above, and was purified by fractionation over immobilized anti-flag M2 antibody, immobilized RAD6, and Superdex-200 columns (Fig. 7D). N-terminally flag-tagged, C-terminally His\textsubscript{6}-tagged \textit{S. cerevisiae} \text{\textsuperscript{1}}CUP9his\textsubscript{6}, was expressed and labeled with \textsuperscript{35}S-methionine in \textit{E. coli}, followed by purification over Ni-NTA and anti-flag M2 antibody columns. Details of these and related procedures are described in ref. (66). We also employed a similarly produced, labeled and purified \textsuperscript{35}S-CUP9his\textsubscript{6}, with a different N-terminal tag, called “polyoma” (py) that was bound by a monoclonal antibody (68) (AK6967; a gift from Dr. S. Stevens and Dr. J. Abelson (Caltech)). The \textit{in vitro} ubiquitylation reactions contained the following components: 7 μM Ub, 50 nM UBA1, 50 nM RAD6, 50 nM UBR1, 550 nM \textsuperscript{35}S-labeled \text{\textsuperscript{1}}CUP9his\textsubscript{6}, 25 mM KCl, 5 mM MgCl\textsubscript{2}, 2 mM ATP, 0.1 mM dithiothreitol (DTT), 25 mM HEPES (pH 7.5) and ovalbumin (as a carrier protein) at 0.5 mg/ml. Pairs of dipeptides were added to final concentrations specified in the legend to Fig. 8. All components except UBA1 were mixed on ice for 10 min; UBA1 was then added and reactions shifted to 30°C. The reactions were carried out for 10 min, and were terminated by adding an equal volume 2 x SDS-PAGE loading buffer and heating at 95°C for 5 min, followed by SDS-10% PAGE and autoradiography.

Genetic Screen for Type-1 and Type-2 UBR1 Mutants – Fig. 1B describes the screen’s design. Reporter plasmids expressing either Arg-URA3 (a type-1 N-end rule substrate) and Leu-βgal (a type-2 N-end rule substrate) (pDSRULB plasmid), or Leu-URA3 (a type-2 substrate) and Arg-βgal (a type-1 substrate) (pDSLURB plasmid), were transformed into \textit{S. cerevisiae} JD55, a \text{ubr1}\Delta strain. pRBUBR1, a high copy plasmid expressing UBR1 from the P_{ADH1} promoter, was mutagenized by treatment with 1
M hydroxylamine for 1 hr at 75°C. DNA was precipitated with ethanol, redissolved in 1 mM EDTA, 10 mM Tris-HCl (pH 8.0), and hydroxylamine was removed using G25 spin-desalting column (Roche). Hydroxylamine-treated pRBUBR1 was transformed into JD55 (ubr1Δ) carrying one of two reporter plasmids, either pDSRULB or pDSLURB. The transformants were plated onto SD(-Ura) plates, and the resulting Ura+ colonies (cells that were, presumably, impaired in degrading either Arg-URA3 or Leu-URA3) were replica-plated onto galactose-containing SG plates. The colonies on SG plates were stained for βgal using XGal overlay. Most of the colonies in these experiments stained blue. Since the cells of these colonies were selected, in the preceding step, to be impaired in the degradation of one class (type-1 or type-2) N-end rule substrates, the failure to degrade an X-βgal that belonged to the other class of N-end rule substrates (type-2 or type-1, respectively) meant that the corresponding alleles of UBR1 were defective in degrading both classes of N-end rule substrates. However, some white colonies on XGal were produced as well, with each of the two bi-reporter plasmids, pDSRULB and pDSLURB. The corresponding pRBUBR1 plasmids were rescued from these white colonies, and were retransformed into the JD55 (ubr1Δ) cells carrying other reporter plasmids, either pBAR (Arg-URA3) or pBAL (Leu-URA3), and the transformants were tested for their ability to grow in the absence of uracil. This way, we identified the regions of relevant alterations in the UBR1 ORF. The corresponding fragments were then sequenced to identify the actual alteration(s).

**Pulse-Chase Analysis** – *S. cerevisiae* in liquid SD medium containing 0.1 mM CuSO4 (A600 0.5-1) were labeled for 2 min at 30°C with 0.15 mCi of 35S-methionine/cysteine (EXPRESS, New England Nuclear, Boston, MA) as described (21,52). Cells were harvested by centrifugation, resuspended in 0.3 ml of SD medium containing 10 mM L-methionine and 0.5 mg/ml cycloheximide, and incubated further at 30°C. Samples of 0.1 ml were removed during the incubation and added to microcentrifuge tubes containing 0.5 ml of 0.5-mm glass beads and 0.7 ml of lysis buffer (1% Triton X-100, 0.15 M NaCl, 1 mM EDTA, 50 mM Na-HEPES (pH 7.5), and 1x protease inhibitor cocktail (Sigma)). The cells were lysed by vortexing with glass beads, followed by immunoprecipitation with anti-ha antibody (Covance, Princeton, NJ), SDS-PAGE, autoradiography, and quantitation (21,56). Immunoprecipitation of labeled CUP9-flag was carried out as described (39).

**β-Galactosidase and Other Assays** – Cells in 5-ml cultures (A600 0.8-1) were pelleted by centrifugation at 12,000g for 2 min, lysed with glass beads in 20% glycerol, 1 mM dithiothreitol, 0.1 M Tris-HCl (pH 8), and the activity of βgal was measured in the clarified extract using o-nitrophenyl-β-D-galactoside, as described (50,69). The activity was normalized to the total protein concentration, determined using Bradford assay (BioRad). For growth-dilution assays, *S. cerevisiae* strains were grown under selection for plasmids in SD medium to an A600 of 0.8-1. Cells (1.5 x 10^7) from each sample were pelleted by centrifugation and resuspended in 1 ml of water. Four-fold serial dilutions of cell cultures were set up in microtiter plates as described (39), and cell suspensions were transferred to various media using a 48-pin applicator, followed by incubation at 30°C for ~36 hr to assay for growth.

**RESULTS**
UBR1 Mutants That Are Selectively Defective in Targeting Type-1 or Type-2 N-End Rule Substrates – To locate the regions of *S. cerevisiae* UBR1 that encompass its type-1 and type-2 substrate-binding sites (see Introduction), and to determine whether these apparently distinct sites (1,28,69) are sufficiently modular (independent) to be selectively perturbed by mutations, we carried out a screen described in Fig. 1B and Experimental Procedures. A plasmid that expressed Arg-URA3 (a type-1 N-end rule substrate) and Leu-βgal (a type-2 N-end rule substrate), or an otherwise identical plasmid that expressed a “converse” set of reporters, Leu-URA3 (a type-2 substrate) and Arg-βgal (a type-1 substrate), were transformed into *ubr1Δ* double reporter plasmids (Fig. 1B). By replica-plating transformants on uracil-lacking SD plates, we could identify cells in which an X-URA3 reporter remained long-lived in the presence of pRBUBR1 from the mutagenized pool, as those cells were Ura+, owing to higher steady-state levels of URA3. (Both Arg-URA3 and Leu-URA3 were long-lived in *ubr1Δ* cells but short-lived in wild-type cells (1,70.) Ura+ colonies were then plated on galactose-containing SG plates. The levels of galactose-induced Arg-βgal (in the Leu-URA3-expressing plasmid) or Leu-βgal (in the Arg-URA3-expressing plasmid) could then be assessed by staining the colonies with XGal. A blue, *i.e.*, high-βgal colony (which already proved to be Ura+ at the screen’s preceding step) would imply the presence of a *UBR1* allele that was impaired in degradation of both type-1 and type-2 substrates (Fig. 1B). Most colonies in these experiments stained blue. The rare white colonies on XGal suggested the presence of UBR1 variants selectively impaired in one but not the other of two substrate-binding sites. The pRBUBR1 plasmids were rescued from these colonies, and were retransformed into *ubr1Δ* cells carrying either Met-URA3 (not an N-end rule substrate, and a relatively long-lived protein), or Arg-URA3 (type-1 substrate), or Leu-URA3 (type-2 substrate). Plating transformants on either SD (+Ura) or SD(-Ura) plates could verify whether the corresponding variants of UBR1 were selectively impaired in the degradation of type-1 or type-2 N-end rule substrates (Fig. 1B).

The initial screen involved ~2,000 transformants bearing mutagenized *UBR1*. With each of two reporter plasmids, ~10% of these transformants were Ura+. Among these, no *UBR1* isolates were found that were unable to confer instability on a type-1 N-end rule substrate (basic N-terminal residues) but retained the activity against a type-2 substrate (hydrophobic N-terminal residues). However, this screen did yield 5 putative UBR1 variants of the opposite kind (active against type-1 but not type-2 N-end rule substrates). A more extensive screen, with ~1 x10⁶ transformants, was then carried out exclusively for UBR1 variants of the former class (active against type-2 but not type-1 substrates). It yielded ~17,000 Ura+ colonies, *i.e.*, isolates in which Arg-URA3 was at least partially stabilized. 20 of these colonies showed low levels of βgal in the XGal assay (*i.e.*, they retained degradation of Leu-βgal). The pRBUBR1 plasmids encoding 20 putative type-1-impaired and 5 putative type-2-impaired UBR1 variants were rescued, and the specificity of their targeting defects was confirmed by re-testing them in *ubr1Δ* cells expressing either Arg-URA3 and Leu-βgal or Leu-URA3 and Arg-βgal, using, in particular, pulse-chase and immunoblot assays (data not shown; see also below).

Identifying Mutations in UBR1 – The approximate locations of mutations within the 5,850 bp *UBR1* ORF were determined through fragment swapping between wild-type and mutant *UBR1*s, using unique restriction sites to divide *UBR1* into five nonoverlapping fragments (see Experimental Procedures). Once localized to a specific fragment, the actual mutation was determined by DNA sequencing. All of the type-1 and type-2 mutations were located in the N-terminal 30% of *UBR1*. Among ~20 type-1 mutant isolates of *UBR1* (those active with type-2 substrates but inactive with type-1 substrates), only 5 different amino acid changes were present, with 18 of the 20 mutations affecting either Asp-176 or Gly-172 (Fig. 1C). All of recovered type-1 mutations in *UBR1* were missense alterations, one of them a double missense, in two adjacent residues. Specific changes in the type-1 *UBR1* mutants, were (Cys-Val)¹⁴⁵/¹⁴⁶ → Tyr-Met (a
double missense; one isolate); Gly\textsuperscript{172} → Arg (7 isolates); Gly\textsuperscript{172} → Glu (4 isolates); Asp\textsuperscript{176} → Asn (5 isolates); and Asp\textsuperscript{176} → Glu (2 isolates), encompassing a ~30-residue region of UBR1 (positions 145-176) (Fig. 1C). In contrast to type-1 UBR1 mutants, all 5 of the type-2 UBR1 mutants contained different amino acid alterations; collectively, these (type-2) alterations encompassed a different and larger region of UBR1 (positions 313-560). Specific type-2 alterations were Val\textsuperscript{313} → Ile; Asp\textsuperscript{318} → Asn; His\textsuperscript{321} → Tyr; Pro\textsuperscript{406} → Ser; and Glu\textsuperscript{560} → Lys (Fig. 1C). Together, the identified type-1 and type-2 mutations spanned a 416-residue stretch within the first 560 residues of the 1,950-residue UBR1 protein.

Degradation of N-end Rule Substrates in Type-1 or Type-2 UBR1 Mutants – Specific type-1 and type-2 S. cerevisiae UBR1 mutants were verified and further characterized by transforming the corresponding UBR1-expressing plasmids into JD55 (ubr1Δ) S. cerevisiae carrying pRβgalUPR or pLβgalUPR. The latter are low-copy plasmids expressing, respectively, Arg-βgal or Leu-βgal reporter N-end rule substrates. The relative rates of in vivo degradation of X-βgals were assessed using extensively validated (9,24) steady-state measurements of βgal activity in whole cell extracts from ubr1Δ cells that carried plasmids expressing specific UBR1 mutants (Fig. 2A, B). The results confirmed the initial, screen-derived data that Arg-βgal was long-lived in the presence of type-1 UBR1 mutants, but was degraded with type-2 UBR1 mutants. Conversely, Leu-βgal was long-lived in the presence of type-2 UBR1 mutants, but short-lived with type-1 UBR1 mutants (Fig. 2A, B). In addition to steady-state levels of N-end rule substrates (Fig. 2A, B), their metabolic stability was also assessed by pulse-chase analysis, using the Ub-protein-reference (UPR) technique (21,57). This method provides a “built-in” (long-lived) reference protein as a part of reporter fusion, and thereby increases the accuracy of pulse-chase and analogous assays. In agreement with steady-state data (Fig. 2A, B), Arg-URA3 was relatively long-lived in type-1 UBR1 mutants, exemplified by UBR1\textsuperscript{D176E.1} cells (the suffixes “-1” or “-2” in the superscript denote the inactivation of, respectively, the type-1 or type-2 substrate-binding sites of UBR1). In contrast, Leu-URA3 was short-lived in a type-1 mutant (Fig. 2D and data not shown; see also Fig. 1C). “Conversely”, Leu-URA3 was relatively long-lived in type-2 UBR1 mutants, exemplified by UBR1\textsuperscript{D176E.2}, whereas Arg-URA3 was short-lived in the same mutant (Fig. 2D and data not shown; see also Fig. 1C). (Both Arg-URA3 and Leu-URA3 were long-lived in the absence of UBR1 (Fig. 2C).)

UBR1-Dependent Targeting of CUP9 Requires Intact Type-2 Binding Site of UBR1 – Previous work has shown that the N-end rule pathway controls the import of short peptides (8,39-41). Specifically, a third substrate-binding site of S. cerevisiae UBR1, apparently distinct from its two other (type-1 and type-2) binding sites, targets CUP9. The latter is a transcriptional repressor of, among other genes, PTR2, which encodes a plasma-membrane transporter that can import di- and tripeptides. UBR1 targets CUP9 for ubiquitylation and subsequent degradation through a distinct internal (non-N-terminal) degron of CUP9. In wild-type cells, peptides with destabilizing N-terminal residues bind to the type-1 and type-2 binding sites of UBR1 and thereby allosterically activate its otherwise autoinhibited third site for binding to CUP9 (8,40). This reversal of autoinhibition leads to the targeting of CUP9 by UBR1, to accelerated degradation of CUP9, and thus to the induction of PTR2 expression and peptide import. The resulting positive-feedback circuit allows S. cerevisiae to detect the presence of extracellular peptides and to react by increasing their uptake (8,40). In ubr1Δ yeast, CUP9 is a relatively long-lived protein (t\textsubscript{1/2} > 50 min). As a result, the expression of PTR2 transporter is extinguished, making ubr1Δ cells incapable of importing peptides and therefore resistant to a toxic dipeptide such as leucyl-ethionine (Leu-Eth) (39).

We employed the toxic peptide resistance assay to probe in vivo aspects of UBR1 mutants identified in the present work. Plasmids expressing type-1 or type-2 mutant alleles of UBR1 were transformed into ubr1Δ yeast, and the resulting transformants were streaked onto Leu-Eth (toxic peptide) plates. Growth of cells on such plates indicates that the CUP9 repressor is metabolically stabilized in these cells, resulting in
a down-regulation of expression of PTR2 transporter and therefore in hyperresistence of cells to Leu-Eth. The results showed that CUP9 was stabilized in cells transformed with all of type-2 UBR1 mutants (Fig. 3B), but was still degraded in cells transformed with four of the five type-1 UBR1 mutants (Fig. 3A). (With the fifth type-1 mutant, UBR1<sup>CV145SYM-1</sup> (Fig. 1C), cell growth on Leu-Eth plates was partially retarded relative to ubr1Δ strains.) These findings suggest that the third (CUP9-recognizing) site of UBR1 either structurally overlaps with the type-2 binding site or depends on this site for its function.

**Specificity of UBR1 Interactions with N-End Rule Substrates and X-Peptides** — Our previous work described the use of N-recognin ligands containing the glutathione-S-transferase (GST) moiety linked to the C-terminus of a fragment of <i>S. cerevisiae</i> SCC1, a subunit of cohesin. This 33-kDa SCC1 fragment, which normally bears N-terminal Arg, is produced in yeast through the conditional cleavage by separase ESP1, and is degraded by the N-end rule pathway (43). Purified X-SCC1-GSTs (X=Arg, Leu, Met) bearing type-1 (Arg), or type-2 (Leu), or stabilizing (Met) N-terminal residues were linked to glutathione-Sepharose, and the previously characterized (8) GST-pulldown assay was carried out with yeast extracts that contained the overexpressed, flag-tagged, full-length <sup>1</sup>UBR1 or its derivatives (Fig. 4). Similarly to mouse UBR1 and UBR2 (31), <i>S. cerevisiae</i> <sup>1</sup>UBR1 did not bind to Met-SCC1-GST (Fig. 4D, lane 1) but bound to both Leu-SCC1-GST and Arg-SCC1-GST (Fig. 4A, B, lanes 2). This binding was specific for type-1 versus type-2 N-terminal residues, as the Arg-Ala and Leu-Ala dipeptides selectively inhibited the binding of corresponding (“alternative”) X-SCC1-GST reporters (Fig. 4A, B). To probe specificity of this inhibition, we tested other small compounds as well. Whereas increasing concentrations of Arg-Ala inhibited the binding of Arg-SCC1-GST to <sup>1</sup>UBR1, neither Ala-Arg, nor an equimolar mixture of the corresponding free amino acids (Arg+Ala), nor Arg methyl ester (Arg-OMe), nor agmatine ((4-aminobutyl)-guanidine), the decarboxylation product of Arg and a putative neurotransmitter (71), exhibited this effect (Fig. 4A, C).

Similarly, the binding of Leu-SCC1-GST to <sup>4</sup>UBR1 was inhibited by Leu-Ala but neither by Arg-Ala (31), nor by Ala-Leu, nor by Leu+Ala (a mixture of free amino acids) (Fig 4D).

In agreement with earlier findings (8), both full-length <sup>4</sup>UBR1 and UBR1<sup>1-1140f</sup> (N-terminal “half” of UBR1) exhibited qualitatively indistinguishable patterns of binding specificity (Fig. 4E, F), indicating that UBR1<sup>1-1140f</sup> contained the type-1 and type-2 binding sites. It should be noted that while Arg-OMe (in contrast to Arg-Ala) did not inhibit the binding of Arg-SCC1-GST to either UBR1<sup>1-1140f</sup> or full-length <sup>4</sup>UBR1, the analogous methyl ester Leu-OMe inhibited the binding of Leu-SCC1-GST apparently as efficaciously as did Leu-Ala (Fig. 4C-F). Yet another difference between the inhibition patterns of type-1 and type-2 sites was a comparably strong inhibition, by either Leu-Ala or Leu-Asp, of the binding of Leu-SCC1-GST to UBR1 versus the absence of inhibition, by Arg-Asp (in contrast to Arg-Ala), of the binding of Arg-SCC1-GST to UBR1 (Fig. 4C-F). Detailed understanding of these differences requires the knowledge of UBR1’s crystal structure.

The binding of UBR1 to destabilizing N-terminal residues of polypeptides was also examined using the X-peptide assay, which employed a set of otherwise identical 12-mer peptides (termed X-nsP4pep) whose sequence X-Ile-Phe-Ser-Thr-Asp-Thr-Gly-Pro-Gly-Gly-Cys (X=Arg or Phe) was derived, in part, from the N-terminal sequence of nsP4, the RNA polymerase of the Sindbis virus and a physiological N-end rule substrate (27,61). Purified X-nsP4pep peptides were crosslinked to microbeads through their C-terminal Cys residue (8,31,32). Both full-length <sup>4</sup>UBR1 and several of its truncated derivatives were overexpressed in <i>S. cerevisiae</i> and assayed for their binding to either Arg-nsP4pep or Phe-X-nsP4pep. Whereas UBR1<sup>1-717f</sup>, the 717-residue N-terminal fragment of the 1,950-residue UBR1, was capable of specific binding to X-nsP4pep peptides, shorter N-terminal fragments of UBR1 did not bind (Fig. 4G-H). In addition, UBR1<sup>209-1140f</sup>, which differed from the binding-competent UBR1<sup>1-1140f</sup> by deletion of the first 208 residues, did not bind to either Arg-nsP4pep or Phe-X-nsP4pep (Fig. 4G, H and Fig. 4J). All of these findings were consistent with the conclusions from genetic mapping of the type-1 and type-2 sites (Fig. 1C).
Measuresments of Affinity Between UBR1 and Peptide Reporters Using Fluorescence Polarization Assay – To measure the binding of UBR1 to different N-terminal residues of otherwise identical polypeptides we employed both fluorescence polarization (FP) (62,63,65) and surface plasmon resonance (SPR) (72,73). These measurements used purified UBR1 (1-1140) (Fig. 4F). Its specific binding to N-terminal residues of either protein-size substrates or short peptides was indistinguishable from that of full-length UBR1 in GST-pulldown assays (Fig. 4A-F). The otherwise identical Arg-nsP4pep, Phe-nsP4pep, and Gly-nsP4pep peptides, with type-1 destabilizing, type-2 destabilizing, and stabilizing N-terminal residues, respectively (Fig. 1A), were C-terminally conjugated to Alexa Fluor-488, followed by purification of the resulting fluorescent peptides by HPLC and verification of their structures by mass spectrometry (see Experimental Procedures).

The concentration of X-nsP4pep (C-terminal symbol denotes fluorescent tag) in FP assays was set at 10 nM, well below the (anticipated) micromolar-range of affinities between UBR1 and test peptides. As expected (8), both Arg-nsP4pep and Phe-nsP4pep interacted with UBR1 (1-1140) (Fig. 4F). The calculated K_d's (see Experimental Procedures) were ~1.0 µM and ~0.7 µM, respectively. In contrast, no binding to UBR1 (1-1140) could be detected with the otherwise identical Gly-nsP4pep, i.e., the corresponding K_d was at least 0.1 µM or higher (Fig. 5A, B). The affinities of X-nsP4pep peptides for UBR1 (1-1140) were also measured in a competition FP assay, in which, for example, the binding of Phe-nsP4pep to UBR1 (1-1140) was titrated by the addition of Alexa-lacking (unlabeled) PhenP4pep (Fig. 5F). The corresponding K_d was ~2 µM, close to the value deduced from direct-binding FP assay (Fig. 5B, F). In contrast, the competition by unlabeled Arg-nsP4pep (a type-1 destabilizing N-terminal residue) or by unlabeled Gly-nsP4pep (a stabilizing N-terminal residue) for the binding of Phe-nsP4pep to UBR1 (1-1140) was too weak to measure by FP (Fig. 5F). The K_d of Arg-nsP4pep binding to UBR1 (1-1140) that was deduced from a similar competition assay, using unlabeled Arg-nsP4pep as a competitor, was ~1.2 µM, close to the value of ~1.0 µM from direct-binding assay (Fig. 5A, C). Interestingly, however, the unlabeled Phe-nsP4pep could detectably compete with the binding of (labeled) Arg-nsP4pep to UBR1 (1-1140) (Fig. 5C), with the apparent K_d of ~11 µM, i.e., Phe-nsP4pep was a ~10-fold weaker inhibitor of the binding of Arg-nsP4pep to UBR1 (1-1140) than Arg-nsP4pep itself. As to unlabeled Gly-nsP4pep, its apparent binding to the type-1 site of UBR1 (1-1140) was weaker still but could still be detected in this competition assay, with the corresponding (apparent) K_d of ~36 µM (Fig. 5C).

We conclude that the type-1 and type-2 substrate-binding sites of UBR1 bind to the corresponding N-terminal residues of model substrates with approximately equal affinities (K_d's of ~1 µM). We also conclude that the UBR1's type-2 site, which binds to bulky hydrophobic N-terminal residues (Fig. 1A), is significantly more discriminating than the type-1 site. Specifically, we could not detect the binding of noncognate N-terminal residues to the type-1 site of UBR1, in either qualitative or quantitative assays (Figs. 4 and 5B, F), whereas a type-2 N-terminal residue could weakly compete with a type-1 N-terminal residue for the binding to the type-1 site of UBR1, and a stabilizing N-terminal residue could also compete, albeit even more weakly than a type-2 residue (Fig. 5A-F).

We also carried out competition assays with unlabeled dipeptides and other dipeptide-size compounds, versus either Arg-nsP4pep or Phe-nsP4pep (Fig. 5D, E, G). Some of these competitors were also used in qualitative binding assays (Fig. 4D, E). The results of FP-based competition assays indicated that dipeptides (as distinguished from 12-mer X-nsP4pep peptides) with cognate N-terminal residues were significantly (5-10-fold) weaker competitors relative to unlabeled cognate X-nsP4pep peptides (Fig. 5D, G), and that dipeptides (or analogous compounds) with noncognate N-terminal residues were either not competitors at all or much weaker competitors (see the legend to Fig. 5D, G for details). Among the tested dipeptide sequences, Ala at position 2 of an otherwise cognate dipeptide (Arg-X or Leu-X) resulted in best competitor dipeptides, whereas a negatively charged (Asp) residue at position 2 significantly decreased the affinity of a dipeptide for the corresponding (type-1 or type-2) substrate-
binding site of UBR1 (Fig. 5D, G). A basic (Arg or Lys) residue at position 2 was not as detrimental for the interaction of dipeptide with UBR1 (Fig. 5D, G).

While the length of a competitor peptide is less important than whether or not it bears a cognate N-terminal residue, the data of Fig. 5H, I indicate that the length of a peptide (in the tested range) is a significant, apparently independent parameter that determines its affinity for UBR1. One possibility is that the peptide’s C-terminal carboxyl group (e.g., its negative charge) would weaken the binding of peptide to its cognate (type-1 or type-2) binding site of UBR1, unless the peptide is long enough to place its C-terminal carboxyl group away from the binding site. This interpretation would also account for the fact that free “cognate” amino acids (in contrast to dipeptides bearing them as N-terminal residues) do not bind to UBR1 (Fig. 4C, E).

**Measurements of Affinity Between UBR1 and Model Peptides Using Surface Plasmon Resonance** – To measure the affinity between UBR1 and 12-mer X-nsP4pep peptides using a technique other than FP (Fig. 5), we employed surface plasmon resonance (SPR) and BIACORE-2000 (64,65,72,73). Phe-nsP4pep, Arg-nsP4pep or Gly-nsP4pep, modified to bear a C-terminal biotin moiety (see Experimental Procedures) were immobilized on streptavidin-coated sensor chips. Purified UBR1 (see Fig. 4I) was serially diluted in the running buffer and injected over the chip surface for 80 sec. BIACORE readouts were corrected by subtracting readouts from the mock surface (lacking an X-peptide) as well as readouts from buffer-only (control) injections. Raw data were processed using Scrubber (BioLogic Software), and were globally fit using CLAMP (64), yielding the corresponding Kd’s. As shown in Fig. 6, BIACORE detected the binding of UBR1 to either Arg-nsP4pep (bearing a type-1 destabilizing N-terminal residue) or Phe-nsP4pep (bearing a type-2 destabilizing N-terminal residue), with the Kd’s of ~1.2 µM and ~1.0 µM, respectively (Fig. 6), in excellent agreement with the results of FP-based measurements (Fig. 5 and preceding section). In contrast, the otherwise identical Gly-nsP4pep, bearing a stabilizing N-terminal residue, did not exhibit a detectable binding to UBR1 (Fig. 6C). These BIACORE-based findings (Fig. 6) demonstrated yet again (in addition to GST-pulldown and FP-based data) that at least the bulk of the in vivo specificity of UBR1 for N-end rule substrates (Fig. 1A, B) (see Introduction) stems from differences in its physical affinity for destabilizing versus stabilizing N-terminal residues in the N-end rule.

**Nonspecific Binding of Purified UBR1 to CUP9, and Restoration of Specificity Through Chaperones and Macromolecular Crowding** – Through its type-1 and type-2 sites, the 225-kDa UBR1 can bind to primary destabilizing N-terminal residues in proteins or short peptides apparently “nonconditionally”, i.e., under all conditions tested (8) (Figs. 4-6). In contrast, UBR1 can bind, through its third binding site, to its physiological substrate CUP9 only if the type-1 and type-2 sites of UBR1 are occupied by cognate N-terminal residues of ligands such as dipeptides (8) (e.g., Fig. 7E-G, lanes 2, 3). This dependence of the UBR1-CUP9 interaction (via the C-terminus-proximal degron of CUP9 (8) (Fig. 7B, C) on the state of occupancy of UBR1’s other two binding sites yields, in vivo, a positive-feedback circuit through which S. cerevisiae can sense the presence of extracellular peptides and accelerate their uptake (8,40) (see Introduction). The negligible binding of CUP9 by UBR1 in the absence of occupancy of its type-1/2 sites (e.g., Fig. 7E-G, lanes 2, 3) is thus the consequence of a negligible fraction of UBR1 molecules, under these conditions, that exhibit an “open” (active) CUP9-binding site.

The binding of type-1/2 ligands to the corresponding sites of UBR1 increases, in an ensemble of UBR1 molecules, the probability of a conformational transition that “opens up” the CUP9-binding site of UBR1 (8). The type-1/2 sites of UBR1 encompass its N-terminus-proximal UBR domain (Fig. 1C and Introduction). The allosteric influence of UBR1-bound dipeptides on the probability of transition that induces the UBR1-CUP9 interaction was shown to require the evolutionarily conserved C-terminal region of UBR1 that includes two critical cysteines, Cys-1703, and Cys-1706 (8). In particular, UBR1, the N-terminal “half” of UBR1, retains all three of its substrate-binding sites but exhibits a constitutively “open” (active) CUP9-binding site. More recent findings, using...
UBR1 derivatives that contained sequences allowing site-specific cleavages of UBR1 indicated that a conformational transition in UBR1 that “opens up” its CUP9-binding site is “subtle”, in that it does not appear to involve a large-scale repositioning of UBR1 domains.

GST-pulldowns with UBR1 versus GST-CUP9 were carried out in the presence of yeast extract (8) (e.g., Fig. 7E-G, lanes 2, 3). We repeated these assays with purified UBR1 (see Experimental Procedures) and GST-CUP9. Surprisingly, the specificity of UBR1-CUP9 interaction, i.e., its dependence on the presence of dipeptides with destabilizing N-terminal residues, as well as its dependence on the presence of C-terminal degron of CUP9 (66) were abolished with purified UBR1: it bound to GST-CUP9 irrespective of the presence or absence of dipeptides, and irrespective of the presence or absence of the UBR1-specific degron of CUP9 (Fig. 7C, F, and data not shown). These effects of using purified UBR1 (as distinguished from UBR1 in the presence of yeast extract) were also observed with its N-terminal fragment UBR1\textsuperscript{1-1140} (Figs. 4K and 7B). For example, in the presence of yeast extract, UBR1\textsuperscript{1-1140} bound to CUP9 constitutively (i.e., independently of the presence of dipeptides) (8) but specifically (i.e., depending on the presence of CUP9’s degron) (Fig. 7B). However, purified UBR1\textsuperscript{1-1140} bound either to full-length CUP9 (containing its C-terminal degron) or to C-terminally truncated CUP9\textsuperscript{1-221} that lacked the degron(Fig. 4K, lanes 1-6). The same pattern was observed with purified full-length UBR1\textsuperscript{versus CUP9\textsuperscript{1-254P}} that contained a previously characterized degron-inactivating missense mutation (8) (Fig. 7I).

We found that the specificity of these interactions, i.e., their dependence on the presence of CUP9’s degron, could be restored if GST-CUP9 pulldowns with purified UBR1\textsuperscript{1-1140} or purified UBR1 were carried out in the presence of a “carrier” protein such as ovalbumin, at concentrations (10 to 50 mg/ml) that would be expected to produce “macromolecular crowding” (Fig. 4K, lanes 4-6, compare with lanes 7, 8; Fig. 7F, lanes 5, 6, compare with lanes 7, 8; Fig. 7G, lanes 9, 10, compare with lanes 5, 6; Fig. 7K, lanes 5, 6, compare with lanes 7, 8). These terms refer to volume-exclusion and related effects that recapitulate, in part, the properties of intracellular milieu, where the total protein concentration can be as high as 200-400 mg/ml (74-76).

As expected, and similarly to the effect of ovalbumin, the re-addition of \textsuperscript{1}UBR1-lacking yeast extract to purified \textsuperscript{1}UBR1 also restored the specificity of UBR1 binding to GST-CUP9 (Fig. 7E). We noticed that GST-CUP9 interacted not only with \textsuperscript{1}UBR1 in the yeast extract but also with another protein that was present in extracts irrespective of the presence of \textsuperscript{1}UBR1 (data not shown). Mass spectrometry (MS) of that protein identified it as EF1A, which functions as a translation elongation factor, a protein chaperone, and is a highly abundant protein ((77) and refs. therein). We employed two chromatographic steps, DEAE-Sepharose and CM-Sepharose chromatography, to purify S. cerevisiae EF1A to near homogeneity (Fig. 7H). The addition of purified EF1A to GST-CUP9 pulldowns with purified \textsuperscript{1}UBR1, at the EF1A concentration of 1.2 mg/ml (a relatively low level, in comparison to minimally effective concentrations of ovalbumin) was found to restore the specificity of UBR1-CUP9 interactions. Specifically, in the presence of purified EF1A, a significant binding of purified \textsuperscript{1}UBR1 to GST-CUP9 was observed only in the presence of dipeptides with destabilizing N-terminal residues (Fig. 7G, lanes 7, 8; compare with lanes 2-6 and 9, 10).

A parsimonious interpretation of these findings with purified UBR1 versus its supplementation with ovalbumin, or with yeast extract, or with EF1A is that CUP9, a transcriptional repressor, functions in vivo as a complex containing “global” corepressors TUP1 and SSN6 (78). In the absence of these CUP9 ligands, the CUP9 moiety in purified GST-CUP9 may exhibit a “sticky” region(s) that mediates a nonspecific interaction between CUP9 and UBR1. This region(s) of CUP9 may be shielded from UBR1 by interactions with either EF1A, acting as a chaperone, and/or with another chaperone-like proteins in the yeast extract, thus allowing a specific UBR1-CUP9 interaction to be detected. Because “crowding” levels of ovalbumin (10-50 mg/ml) also restore the specificity of UBR1-CUP9 interaction (Fig 4K and Fig. 7F, G), the presumed “sticky” region(s) of CUP9 may be partially unfolded in the absence of EF1A or ovalbumin. In vivo, CUP9 is likely to be targeted by UBR1 for ubiquitylation (and
subsequent degradation) as a part of the TUP1-SSN6-CUP9 complex (79,80). The previously discovered subunit selectivity of the N-end rule pathway (20,23) is expected to make possible selective degradation of the CUP9 moiety in the TUP1-SSN6-CUP9 complex in vivo, leaving the corepressor TUP1-SSN6 intact. This model remains to be addressed experimentally.

Synergistic Effects of Type-1/2 Dipeptides in a UBR1-Dependent Ubiquitylation System – Our studies of the targeting of CUP9 by the UBR1 Ub ligase (8,40) included the development of an in vitro system that consisted exclusively of small compounds and purified recombinant proteins (Ub, UBA1, RAD6, UBR1, and CUP9) (see Experimental Procedures and ref. (66)). This system exhibited the UBR1-dependent, RAD6-dependent polyubiquitylation of CUP9 that was enhanced by single dipeptides such as Trp-Ala or Lys-Ala that bore destabilizing N-terminal residues, but was not enhanced by compositionally identical single dipeptides such as Ala-Trp or Ala-Lys, which bore a stabilizing N-terminal residue (40,66). As with most such systems, the extent of its specificity, in this case the extent of its UBR1 dependence vis-à-vis polyubiquitylation of CUP9, was a function of both UBR1, RAD6 and Ub concentrations (ref. (66) and data not shown.)

In the present study, we asked whether the system’s specificity could also recapitulate, at the level of UBR1-dependent polyubiquitylation of CUP9, the previously demonstrated dependence of the UBR1-CUP9 in vitro interaction (assayed by GST-pulldowns) on the presence of both type-1 and type-2 dipeptides (e.g., both Arg-Ala and Leu-Ala) (8) (see also Fig. 7). Indeed, we observed a clear enhancement of the UBR1-dependent CUP9 polyubiquitylation upon the addition of both Arg-Ala and Leu-Ala (at micromolar levels), as distinguished from either Arg-Ala alone or Leu-Ala alone (Fig. 8). Specifically, little polyubiquitylated CUP9 was formed in reactions that contained no added dipeptides (Fig. 8, lane 2). Some polyubiquitylated CUP9, bearing relatively short polyUb chains, was formed in the presence of a mixture of Leu-Ala (a type-2 dipeptide) and Ala-Arg (bearing a stabilizing N-terminal residue), at 2 µM each (Fig. 8, lane 5). The distribution of polyUb chain sizes shifted slightly upward at 10 µM of Leu-Ala plus Ala-Arg (Fig. 8, lane 6; compare with lanes 2, 5). However, when a mixture of type-1 (Arg-Ala) and type-2 (Leu-Ala) dipeptides was present, the polyubiquitylation of CUP9 was detectable already at 0.8 µM of each dipeptide (Fig. 8, lane 7; compare with lane 2). The average size of CUP9-linked polyUb chains greatly increased at 2 µM of type-1/2 dipeptides, and the average size of such chains was much higher than the one observed at the same concentration of Leu-Ala plus Ala-Arg (Fig. 8, lane 8; compare with lane 6). Finally, at 5 µM of type-1/2 dipeptides, the size of most polyUb chains became too large for the resulting CUP9 conjugates to be fractionated under conditions used, resulting in a diffuse “band” of large conjugates near the origin of the gel (Fig. 8, lane 9; compare lanes 7, 8, and also with lane 6, which illustrates the average size of polyUb chains in the presence a type-2 dipeptide (Leu-Ala plus Ala-Arg) at 10 µM, twice as high as the 5 µM level of type-1/2 dipeptides together).

Thus, in addition to the previously demonstrated dependence of the physical binding of UBR1 to CUP9 on the presence of both type-1 and type-2 dipeptides (8) (see also Fig. 7), our present findings (Fig. 8) strongly suggest that the processivity of polyubiquitylation of CUP9 by the UBR1-RAD6 Ub ligase is increased when both of the type-1 and type-2 binding sites of UBR1 are occupied by cognate dipeptides. One constraint on possible mechanisms of UBR1 processivity stems from the fact that the bulk of either purified UBR1 or UBR1 in yeast extracts is a monomer, as determined by gel filtration (refs. (66,70) and data not shown).

DISCUSSION

Since the 1990 cloning of S. cerevisiae UBR1, the first molecularly characterized E3 Ub ligase (70), it was clear that at least the bulk of specificity of UBR1 toward polypeptides with primary destabilizing N-terminal residues (Fig. 1A) stems from its selective binding to those residues, rather than, e.g., from effects of a bound N-terminal residue of a substrate on the kinetics of substrate ubiquitylation by the UBR1-RAD6 holoenzyme. For example, immunoprecipitation of an epitope-tagged UBR1 from yeast extracts containing Arg-βgal (an N-end rule substrate) versus the otherwise identical Val-βgal (not an N-
end rule substrate) coprecipitated Arg-βgal but not Val-βgal (70). Quantitative measurements, in the present study, of these differential affinities of UBR1 confirmed and extended the earlier inferences. Specifically, using both fluorescence polarization (FP) and surface plasmon resonance (SPR), we found that UBR1 bound to either type-1 (basic) or type-2 (bulky hydrophobic) N-terminal residues of 12-mer X-peptides with \( K_d \) of ~1 \( \mu \)M. In contrast, neither FP nor SPR could detect the binding of UBR1 to an otherwise identical reporter peptide bearing a stabilizing N-terminal residue such as Gly (Figs. 5 and 6).

The above affinity of the 225-kDa yeast UBR1 for cognate (primary destabilizing) N-terminal residues (Fig. 1A) is in agreement with the approximately equal affinity of ClpP, the 12 kDa E. coli N-recognin (see Introduction), for its cognate (bulky hydrophobic) N-terminal residues (2,13). The relatively low \( K_d \) of ~1 \( \mu \)M affinity of UBR1 for N-degrons, i.e., the relatively high rate of UBR1 dissociation from substrate proteins may facilitate the transfer of a targeted, polyubiquitylated substrate to downstream components of the N-end rule pathway, such as Ub-binding chaperones and the 26S proteasome. Although UBR1 is a relatively scarce protein in \textit{S. cerevisiae} (~300 molecules per haploid cell), the UBR1-dependent N-end rule pathway is highly efficacious. For example, wild-type levels of UBR1 sufficed to reduce the steady-state level of an N-end rule reporter such as Arg-βgal to less than 5% of its level in the absence of UBR1, even under conditions where Arg-βgal was expressed from the induced \( P_{\text{GAL1}} \) promoter on a high copy plasmid and would have been an abundant protein in the absence of its degradation by the N-end rule pathway (9,54,69).

UBR1 was also found to interact with specific subunits of the 26S proteasome (81), in apparent contradiction to the notion of rapid transfer of a UBR1-targeted substrate to “downstream” effectors. One possibility is that a significant fraction of UBR1 molecules might be associated, at any given time, with the (much more abundant) 26S proteasome. In other words, the UBR1-dependent targeting of N-end rule substrates may take place in the context of a UBR1-proteasome complex. This model remains to be addressed experimentally.

Genetic screens, in the present work, for \textit{UBR1} mutants that lost the ability to target type-1 N-end rule substrates but retained the targeting of type-2 substrates and \textit{vice versa} have yielded both classes of mutants, thereby indicating a modular organization of the type-1 and type-2 substrate-binding sites of UBR1. The type-1 site (it binds to N-terminal Arg, Lys or His) is located in the ~70-residue UBR domain of UBR1, whereas the type-2 site (it binds to N-terminal Trp, Phe, Tyr, Leu, or Ile) resides at least in part downstream of the type-1 site and moreover, may partially overlap (or functionally interact) with the third substrate-binding site of UBR1, the one that targets CUP9 (Figs. 1C and 3). We also found that the previously characterized UBR1-dependent \textit{in vitro} ubiquitylation system (40,66) exhibited a synergistic effect of two dipeptides with cognate type-1 and type-2 N-terminal residues on the rate of CUP9 polyubiquitylation (Fig. 8). This ubiquitylation-based result extended our earlier observation (8) (see also Fig. 7) that the physical binding of UBR1 to CUP9 \textit{in vitro} required the occupancy of both type-1 and type-2 binding sites of UBR1 by added dipeptides.

Another finding of the present work is that a specific (CUP9 degron-specific) \textit{in vitro} interaction between GST-CUP9 and UBR1 required the presence of either yeast extract, or the translation factor EF1A (which is known to function as a chaperone), or ovalbumin, the latter at concentrations that produce macromolecular crowding (Figs. 4 and 7). As described in Results, this requirement is likely to stem from the fact that CUP9 functions in a complex with the general repressor TUP1-SSN6 (78-80). Thus, in contrast to the \textit{in vitro} setting of purified GST-CUP9 (Fig. 7), UBR1 may not have access, \textit{in vivo}, to (potentially) UBR1-interacting surfaces of CUP9 outside of its C-terminus-proximal degron (8) (Fig. 7). It remains to be determined whether the short \textit{in vivo} half-life of CUP9 (~5 min in the absence of UBR1-activating dipeptides in growth medium and < 2 min in their presence (40)) reflects the UBR1-mediated targeting of CUP9 directly at chromatin, i.e., at transcriptional promoters that comprise the CUP9 regulon.

In the \textit{S. cerevisiae} UBR1 Ub ligase, in its mammalian sequelog UBR1/UBR2, and in some other N-recognins of mammals and plants...
The UBR domain contains at least the type-1 substrate binding site. A UBR domain is also present in UBR1 sequels that do not function as N-recognins. One example is the S. cerevisiae UBR2 Ub ligase, which is distinct from metazoan UBR2. Physiological substrates of S. cerevisiae UBR2 include RPN4, a short-lived transcription factor that regulates the synthesis of proteosomal subunits (82,83). The mammalian UBR3 Ub ligase (its functions include regulation of olfaction) is yet another UBR domain-containing E3 that is not an N-recognin (29). Physiological ligands of the UBR domains in Ub ligases that are not N-recognins remain to be discovered.

Although prokaryotes lack the Ub system, they contain the N-end rule pathway, Ub-independent versions of it (10-15). Remarkably, although the N-end rules of prokaryotes are subsets of eukaryotic N-end rules, the underlying mechanisms differ beyond the absence of Ub from prokaryotes. For example, the N-end rule of Vibrio vulnificus, a human pathogen, is nearly identical to the mammalian N-end rule. Specifically, among the 13 destabilizing N-terminal residues in the mammalian N-end rule (Fig. 1A), only His, Asn and Gln are stabilizing residues in V. vulnificus (12). (N-terminal Asn and Gln are destabilizing in eukaryotes, owing to the presence of N-terminal amidases (Fig. 1A), which are absent from examined prokaryotes.) However, the similarity of eukaryotic and prokaryotic N-end rules is attained through distinct designs. For example, the N-terminal Arg and Lys residues are destabilizing in eukaryotic N-end rules because they are recognized directly (together with His) by the type-1 substrate-binding sites of the pathway’s Ub ligases (see Introduction). In contrast, N-terminal Arg and Lys are secondary destabilizing residues in both E. coli and V. vulnificus, owing to the presence of L/FK,R-transferase, which conjugates Leu (or Phe) to N-terminal Arg or Lys (10,11,16-18). The resulting additional (bulky hydrophobic) residue at the N-terminus of a substrate is recognized by the ClpS N-recognin, which mediates the delivery of bacterial N-end rule substrates to the proteasome-like ClpAP protease (13-15). Furthermore, the N-terminal Asp, Glu and (oxidized) Cys residues in eukaryotes are arginylated (by R-transferase) before their recognition (through the conjugated N-terminal Arg) by Ub ligases (Fig. 1A). In contrast, N-terminal Asp and Glu in prokaryotes are either stabilizing (unrecognized) residues, e.g., in E. coli, or are destabilizing (e.g., in V. vulnificus) but through their leucylation, i.e., the conjugation to Leu by the bpt-encoded L,D,E-transferase, whose specificity differs from that of the aat-encoded L,FK,R-transferase (2,12).

Thus, the “addition” of Arg and Lys to an early (primordial) N-end rule (which presumably comprised a set of bulky hydrophobic residues) during the evolution of eukaryotes involved the emergence of an “extra”, type-1 (Arg/Lys/His-specific) binding site in Ub ligases (N-recognins) of the N-end rule pathway (Fig. 1A). By contrast, an otherwise similar expansion of N-end rule in prokaryotes involved the appearance of Leu/Phe-conjugating aminoacyl-tRNA-protein transferases such as L,FK,R-transferase (Aat) and L,D,E-transferase (Bpt) (2,12). While these prokaryotic transferases are largely nonsequelogous and recognize different N-terminal residues (see above), they have in common their conjugation of a hydrophobic residue, Leu or Phe, to N-termini of their substrates. This conjugation results in the binding of substrates by the ClpS N-recognin.

A plausible and parsimonious scenario is that an N-end rule pathway that targeted substrates through bulky hydrophobic but not (yet) the other N-terminal residues, emerged in a common ancestor of prokaryotes and eukaryotes. Being sensitive to protein folding and recognizing a “topologically” universal determinant such as a substrate’s N-terminal residue, the N-end rule pathway is likely to have appeared early, possibly coevolving with ribosomes during the emergence of translation-endowed organisms. Subsequent expansions of this proteolytic system in prokaryotes and eukaryotes kept the rule books of N-end rule pathways similar in all organisms (presumably owing to a selection for N-end rules of certain compositions), but the rule-implementing machinery has evolved divergently in prokaryotes versus eukaryotes. For example, Ub ligases and ubiquitylation emerged (or were retained) only in eukaryotes (12,84). In addition, the task of recognizing additional (basic and acidic) N-terminal residues in later (more elaborate) N-end rule pathways was solved quite differently in eukaryotes and prokaryotes, as
described above. Moreover and significantly, the evolution of N-end rule’s Ub ligases (N-recogins) in eukaryotes expanded their repertoire of substrate-binding sites beyond the N-end rule. One example is the third site of *S. cerevisiae* UBR1 that recognizes an internal (non-N-terminal) degron in a transcriptional regulator CUP9, and does so in a way that is controlled, allosterically, by the occupancy of the other two binding sites of UBR1 (see Introduction). A landscape of selective pressures, quasi-neutral mutational drifts, and stochastic fixations of changes (85) that gave rise to specific differences between eukaryotic and prokaryotic N-end rule pathways remains to be understood. This problem is made particularly interesting by the fact that while the functions of the N-end rule pathway in eukaryotes are many and partially understood (see Introduction), the functions of prokaryotic N-end rule pathways are still unknown.

**ACKNOWLEDGMENTS**

We thank R. Deshaies and P. Bjorkman (Caltech, Pasadena, CA) for the use of Analyst Assay Detection System and BIACORE in their laboratories, S. W. Stevens and J. Abelson (Caltech) for the AK6967 antibody, and D. G. Myszka (University of Utah, Salt Lake City, UT) for the CLAMP and Scrubber programs. We are grateful to former and current members of the Varshavsky laboratory, especially to C.-S. Hwang, for helpful discussions and comments on the manuscript. This work was supported by grants to A.V. from the NIH (DK39520 and GM31530) and from the Sandler Program for Asthma Research.

**FOOTNOTES**

1Address correspondence to: Alexander Varshavsky, Division of Biology, California Institute of Technology, Pasadena, CA 91125. Tel.: 626-395-3785; Fax: 626-440-9821; Email: avarsh@caltech.edu.

2The abbreviations used are: Ub, ubiquitin; E1, ubiquitin-activating enzyme; E2, ubiquitin-conjugating enzyme (or ubiquitin carrier protein); E3, ubiquitin-protein isopeptide ligase; R-transferase, arginyl-tRNA-protein transferase; L/FK,R-transferase, Leu-tRNA-protein transferase specific for N-terminal Lys or Arg; L,D,E-transferase, Leu-tRNA-protein transferase specific for N-terminal Asp, Glu, and (oxidized) Cys; NO, nitric oxide; FP, fluorescence polarization; SPR, surface plasmon resonance; βgal, β-galactosidase.

3Z. Xia and A. Varshavsky, unpublished data.

4C.-S. Hwang and A. Varshavsky, unpublished data.
REFERENCES


84. Catic, A., Sun, Z. Y., Ratner, D. M., Misaghi, S., E., Samuelson, J.,
*EMBO J.* **26**, 3474-3483


FIGURE LEGENDS

FIGURE 1. The N-end rule pathway in the yeast *S. cerevisiae*. A, The *S. cerevisiae* N-end rule pathway. N-terminal residues are indicated by single-letter abbreviations for amino acids. Yellow ovals denote the rest of a protein substrate. Hemin (Fe^3+-heme) has been found to inhibit the arginylation activity of both the *S. cerevisiae* and mouse ATE1-encoded Arg-tRNA-protein transferase (arginyl-transferase or R-transferase) (4). At least in mammalian cells, hemin also induces the proteasome-dependent degradation of R-transferase (4). MetAPs, methionine aminopeptidases. “Separase” refers to the ESP1-encoded protease that cleaves, in particular, the SCC1 subunit of cohesin and thereby allows mitosis (86). The resulting C-terminal fragment of SCC1 bears a destabilizing N-terminal residue (Arg in *S. cerevisiae*) and is degraded by the N-end rule pathway, a step that is essential for high fidelity of chromosome segregation (43). The reactions in a shaded rectangle are a part of the N-end rule pathway that is active in eukaryotes (e.g., vertebrates) that produce nitric oxide (NO) (6), but is also relevant to organisms such as *S. cerevisiae*, which lack NO synthases but can produce NO by other routes, and in addition can be influenced by NO from extracellular sources. C* denotes oxidized Cys, either Cys-sulfinate or Cys-sulfonate, produced in reactions mediated by nitric oxide (NO), oxygen (O_2) and their derivatives. Oxidized N-terminal Cys is arginylated by ATE1-encoded isoforms of R-transferase (6,7). Type-1 and type-2 primary destabilizing N-terminal residues (see Introduction) are recognized by UBR1, the sole N-recognin of *S. cerevisiae* that functions as a holoenzyme containing UBR1 and the E2 enzyme RAD6. Through its third substrate-binding site, UBR1 recognizes internal (non-N-terminal) degrons in substrates (denoted by a larger oval) such as CUP9, a transcriptional repressor of regulon that includes the PTR2 peptide transporter (8,39,40). The addition of hemin to *S. cerevisiae* UBR1 *in vitro* (in cell extracts) was found to inhibit the interaction of UBR1 (through its third substrate-binding site) with CUP9 (4). This inhibition was selective in that hemin did not interfere with the binding of primary destabilizing N-terminal residues of peptide reporters to the type-1 or type-2 binding sites of UBR1 (4). B, Design of a genetic screen to isolate UBR1 mutants with selectively inactivated type-1 or type-2 substrate-binding sites. C, Locations of mutations in *S. cerevisiae* UBR1 that were found to selectively inactivate either the type-1 site (UM1 mutations) or the type-2 site (UM2 mutations). The mutant residues are in red. For the names of specific UBR1 domains, see refs. (9,29,32), and the main text.

FIGURE 2. Selective inactivation of type-1 versus type-2 substrate-binding sites in UBR1 mutants. A, Wild-type UBR1 and its mutant alleles were expressed from low-copy plasmids in *S. cerevisiae* JD55 (ubr1Δ) cells that also carried a plasmid expressing Arg-βgal (Ub-Arg-βgal) reporter (see Experimental Procedures). Measurements of βgal activity in cell extracts were used to compare the rates of X-βgal reporter degradation *in vivo* (9,21). The values shown are the means of duplicate measurements of three independent transformants. Standard deviations are indicated above the bars. For the locations of indicated mutations in UBR1, see Fig. 1C. B, Same as in A but with cells expressing Leu-βgal (Ub-Leu-βgal). C, Pulse-chase assays, using a UPR-type construct (21,57) expressing either Met-URA3ha, or Arg-URA3ha, or Leu-URA3ha in ubr1Δ cells (see Experimental Procedures). Lane 1, molecular mass marker; their masses, in kDa, are indicated on the left. D, Same as in C, but with ubr1Δ cells that expressed either wild-type UBR1, or its type-1 UBR1^{D176E-1} mutant, or its type-2 UBR1^{P4665-2} mutant. The ^35S^-bands of X-URA3ha reporters and the DHFRh-Ub reference protein are indicated.

FIGURE 3. *In vivo* effects of mutants in the type-1 or type-2 substrate-binding sites

of UBR1. A, A toxic-peptide (Leu-Eth) resistance assay (39,41) (see Experimental Procedures) with ubr1Δ S. cerevisiae versus its derivatives that expressed type-1 UBR1 mutants (Fig. 1C). B, Same as in A, but the assay was carried with type-2 UBR1 mutants, and the result with wild-type UBR1 is shown as well (Fig. 1C). See the main text for details.

FIGURE 4. Specific interactions between UBR1 and its physiological substrates or model ligands. A, GST-pulldown assays with full-length, flag-tagged UBR1 and immobilized Arg-SCC1-GST. Equal amounts of extract from S. cerevisiae that expressed UBR1 were incubated with glutathione-Sepharose beads preloaded with Arg-SCC1-GST, in the presence of indicated concentrations of either the Arg-Ala (RA) or Ala-Arg (AR) dipeptides. The beads-associated UBR1 was eluted from the beads, fractionated by SDS-PAGE, and detected by immunoblotting with anti-flag antibody (see Experimental Procedures). B, Same as in A but with Leu-SCC1-GST and either Leu-Ala (LA) or Ala-Leu (AL). C, The same assay format, with Arg-SCC1-GST. Lane 1, “5%” refers to a directly loaded sample of UBR1-containing yeast extract that corresponded to 5% of the amount of extract in GST-pulldowns of this panel. Lane 2, Arg-SCC1-GST pulldown assay with UBR1 in the presence of 1 mM AR dipeptide (single-letter notations for amino acids). Lanes 3-10, same as lane 2 but with either A+R (1 mM free Ala and 1 mM free Arg), or RA, or R-OMe (O-methyl ester), or agmatine (see the main text), or RD, or RK, or RF, or RL, all of them at 1 mM. D, Same as in C but with Leu-SCC1-GST and different dipeptides or related compounds, as shown. Lane 1, Met-SCC1-GST (instead of Leu-SCC1-GST); note the absence of its interaction with UBR1 in contrast to Leu-SCC1-GST and Arg-SCC1-GST. E, Same as in C, but with UBR11-1140f, the N-terminal “half” of UBR1. F, Same as in D but with UBR11-1140f. G, Input samples of UBR11-1140f and its derivatives. Lane 1, UBR11-1140f. Lanes 2-7, UBR11-717. UBR1209-1140f, UBR1454-1140f, UBR1710-1140f, UBR1866-1140f, and UBR1180*, respectively. These derivatives of UBR1 were overexpressed in S. cerevisiae as described for UBR11-1140f. The asterisk denotes a protein crossreacting with anti-flag antibody. UBR1180* (lane 7) was employed as a (negative) control for antibody specificity. “10% input” refers to a directly loaded sample of yeast extract that corresponded to 10% of the amount of extract in GST-pulldowns of panels H and J. H, The X-peptide assay (see Experimental Procedures) with UBR11-1140f and its derivatives, described in G. UBR11-1140f, and UBR11-717, but not the other UBR1 derivatives of panel G bound to immobilized Arg-nsP4pep peptide. I, Lane 1, molecular mass markers. Lane 2, Coomassie-stained SDS-PAGE pattern of (overexpressed) UBR11-1140f that was purified from yeast extract using anti-flag affinity chromatography and used for GST-CUP9 pulldowns (see panel K and Experimental Procedures). J, same as in H but with immobilized Phe-nsP4pep peptide. K, Lane 1, 2% input. Lane 2, GST-CUP9 pulldown with UBR11-1140f in S. cerevisiae extract. Lane 3, same as lane 2, but with GST-CUP91-221, a C-terminally truncated mutant of the 306-residue CUP9 that lacks its C-terminal degron (66). Note the absence of binding of UBR11-1140f (in yeast extract) to GST-CUP91-221, in contrast to its binding to full-length GST-CUP91-306. Lane 4, 10% input. Lane 5, same as lane 2 but with purified UBR11-1140f, as distinguished from UBR11-1140f in yeast extract. Lane 6, same as lane 5 but with GST-CUP91-221. Note that purified UBR11-1140f, in contrast to the same protein in yeast extract, binds to both degron-containing (lane 5) and degron-lacking CUP9 (lane 6). Lane 7, same as lane 5 but with ovalbumin (at 10 mg/ml) added to UBR11-1140f before GST-pulldown with full-length GST-CUP9. Lane 8, same as lane 7 but with GST-CUP91-221. Note restoration of the binding specificity of UBR11-1140f in the presence of ovalbumin. Lanes 9 and 10, same as lanes 5 and 6 but “empty” yeast extract was added to purified UBR11-1140f before GST-pulldowns with either full-length GST-CUP9 (lane 9) or degron-lacking GST-CUP91-221.

FIGURE 5. Measurements of affinity between UBR11-1140f and peptides with different N-terminal residues using fluorescence polarization (FP) assay. A, Purified UBR11-1140f (Fig. 4I) was incubated, at increasing concentrations, with either Arg-nsP4pep.
peptide (curve 1) or Gly-nsP4pep (curve 2), each of them at 10 nM (C-terminal "diamond" in the text and a "star" in the figure denote a C-terminal fluorescent tag; see Experimental Procedures). Note the binding to Arg-nsP4pep (bearing a type-1 destabilizing N-terminal residue) and the absence of significant binding to Gly-nsP4pep (bearing a stabilizing N-terminal residue). B, Same as in A but with Phe-nsP4pep versus Gly-nsP4pep (curves 1 and 2, respectively). C, Competition assay, with UBR11-1140f at 0.7 μM, Arg-nsP4pep at 10 nM, versus increasing concentrations of unlabeled Arg-nsP4pep, Phe-nsP4pep, or Gly-nsP4pep (curves 1-3, respectively). D, Same as in C but with unlabeled dipeptides as competitors, at indicated concentrations. Curves 1-6, LA, AR, FA, RD, RK, and RA, respectively (single-letter abbreviations for amino acids). E, Same as in D, but with an equimolar mixture of Arg and Ala (curve 1), Arg-OMe (R-OMe) (curve 2), and Arg-Ala (RA) as unlabeled competitors. F, Same as in C, but with fluorescent Phe-nsP4pep, with UBR11-1140f at 0.7 μM, and the competition with unlabeled Gly-nsP4pep, Arg-nsP4pep, or Phe-nsP4pep (curves 1-3, respectively). G, Same as in D but with Phe-nsP4pep as a labeled peptide, with UBR11-1140f at 0.7 μM, and with unlabeled dipeptides (or their derivatives) as competitors. Curves 1-7, LD, L-OMe (Leu-OMe), LG, LL, LR, LA, and WA, respectively. H, Same as in G, but with Arg-nsP4pep as a labeled peptide, with UBR11-1140f at 0.7 μM, and with unlabeled competitors GGGG, RGD, RGDS, RG, or RAGESGSGC (single-letter abbreviations for amino acids) (curves 1-5, respectively). The RAGES sequence of the latter peptide is identical to the N-terminal sequence of the separase-produced C-terminal fragment of SCC1, a physiological N-end rule substrate (see Results and Discussion). I, Same as in H, but with Phe-nsP4pep as a labeled peptide, with UBR11-1140f at 0.7 μM, and with unlabeled GGGG, FG, FGG, FGGF, FA, and Phe-nsP4pep peptide as competitors (curves 1-6, respectively).

**FIGURE 6.** Measurements of affinity between UBR11-1140f and peptides with different N-terminal residues using surface plasmon resonance. A, Biacore sensograms (see Experimental Procedures) illustrating the binding of UBR11-1140f to the immobilized 12-residue Arg-nsP4pep peptide, bearing a type-1 primary destabilizing N-terminal residue. RU, resonance units. The curves labeled 1 to 5 correspond to the passing of pure UBR11-1140f, at either 0.188 μM, 0.375 μM, 0.75 μM, 1.5 μM or 3 μM, over immobilized Arg-nsP4pep. B, Same as in A but with immobilized Phe-nsP4pep, bearing a type-2 primary destabilizing N-terminal residue (Fig. 1A). C, Same as in A but with Gly-nsP4pep bearing a stabilizing N-terminal residue (no detectable binding to UBR11-1140f at any of its tested concentrations). D and E, Binding isotherms (derived from the data in A and B, respectively) for Arg-nsP4pep and Phe-nsP4pep (see Experimental Procedures).
but 2% of total \(^1\)UBR1 in the assay. Lanes 2, 3, same as lanes 3, 2 (respectively) in C, but an independent experiment. Lane 4, same as lane 1, but with purified \(^1\)UBR1. Lanes 5, 6, same as lanes 2, 3 but with purified \(^1\)UBR1 to which was added an “empty” (lacking \(^1\)UBR1) \(S.\) \textit{cerevisiae} extract. Note that purified \(^1\)UBR1 assayed in the presence of “added-back” yeast extract exhibited specific (dipeptides-dependent) binding to GST-CUP9 (lanes 5, 6), indistinguishably from \(^1\)UBR1 that had not been initially purified (lanes 2, 3). F, Lanes 1-4, same as lanes 1-4 in E but an independent experiment. Lane 4, same as lane 1, but 10% of the input sample of the purified \(^1\)UBR1 (see lane 2 in D). Lanes 5, 6, same as lanes 2, 3 but with purified \(^1\)UBR1. Note the lack of dependence of UBR1-CUP9 interaction on the presence of cognate dipeptides. Lanes 7, 8, same as lanes 5, 6 but GST-pulldowns were carried out in the presence of added ovalbumin, at 10 mg/ml. G, Lanes 1-3, same as lanes 1-3 in E or F, but an independent experiment. Lane 4, same as lane 4 in F, but an independent experiment. Lanes 5, 6, same as lanes 7, 8 in F, but with ovalbumin at 50 mg/ml. Lanes 7, 8, same as lanes 5, 6, but with purified EF1\(\alpha\) added to the assay (to 1.2 mg/ml), instead of ovalbumin. Lanes 9, 10, same as lanes 7, 8, but in the absence of added EF1\(\alpha\). H, Coomassie-stained gel of purified EF1\(\alpha\) (see Experimental Procedures). I, Lane 1, same as lane in F, but an independent experiment. Lanes 2, 3, same as lanes 2, 3 in E or F, but with GST-CUP9\(129-304\), a previously characterized missense mutant of CUP9 (in the region of its degron) that did not bind to UBR1 and was metabolically stabilized \textit{in vivo} (8). Note the absence of binding of \(^1\)UBR1 (in the extract) to GST-CUP9\(129-304\), irrespective the absence or presence of dipeptides. Lane 4, same as lane 4 in F, but an independent experiment. Lanes 5, 6, same as lanes 2, 3 but with purified \(^1\)UBR1. Note that purified \(^1\)UBR1 binds to degron-impaired GST-CUP9\(129-304\), and does so irrespective of the presence or absence of dipeptides. J, Lane 1, same as lanes 4 in F or I, but an independent experiment. Lanes 2, 3, same as lanes 5, 6 in F, but in the presence of ovalbumin (at 50 mg/ml). Lanes 4, 5, same as lanes 2, 3 but with GST-CUP9\(129-304\), instead of “wild-type” GST-CUP9. K, Lane 1, same as lane 1 in F, but an independent experiment. Lane 2, 3, same as lanes 2, 3 in F and G, but with GST-CUP9\(^{222-306}\), an N-terminally truncated CUP9. Note the retention of specific (dipeptides-modulated) binding of this CUP9 derivative to \(^1\)UBR1. Lane 4, same as lane 4 in F, but an independent experiment. Lanes 5, 6, same as lanes 2, 3 but with purified \(^1\)UBR1. Lanes 7, 8, same as lanes 5, 6, but with added ovalbumin (ov), at 50 mg/ml. Note the restoration of specific binding. L, An alternative UBR1-CUP9 binding assay in which the interactions were probed using immunoprecipitation of \(^1\)UBR1 and \(^\text{Py}\)CUP9, a GST-lacking, full-length CUP9 bearing the N-terminal “polyoma”(py) epitope tag (see Experimental Procedures). Flag-tagged \(^1\)UBR1 in the yeast extract was incubated with beads-immobilized anti-flag antibody, followed by isolation of \(^1\)UBR1-bearing beads by centrifugation, washes with binding buffer and incubation with either purified \(^\text{Py}\)CUP9 alone (lanes 2, 3, upper panel) or in the presence of (added) UBR1-lacking yeast extract (lanes 4-7). The incubations were carried out in either the absence (lanes 2, 4, 6) or the presence of Arg-Ala (RA) and Leu-Ala (LA) dipeptides, each of them at 1 mM. The beads containing immobilized \(^1\)UBR1 were pelleted by centrifugation, washed, and the relative amounts of retained \(^\text{Py}\)CUP9 were then determined by SDS-PAGE and immunoblotting with anti-py antibody (see Experimental Procedures). The upper panel shows the amounts of \(^\text{Py}\)CUP9 that were bound to \(^1\)UBR1 in the absence of added yeast extract (lanes 2, 3) or in the presence of extract. Lane 1 shows the amount of \(^\text{Py}\)CUP9 that corresponds to 10% of its input in each of the binding assays. The lower panel is the result of SDS-PAGE and immunoblotting (with anti-flag antibody) of \(^1\)UBR1 that was eluted from anti-flag beads by the flag peptide, to verify the approximate equality of “input” levels of \(^1\)UBR1 in each assay. Note that in the absence of added (UBR1-lacking) yeast extract, the interaction between \(^1\)UBR1 and \(^\text{Py}\)CUP9 was nonspecific, in that it was independent of the absence or presence of cognate dipeptides (lanes 2, 3; compare with lanes 4-7). Lanes 6, 7, same as lanes 4, 5, but the input \(^\text{Py}\)CUP9 was 2-fold higher.

**FIGURE 8.** Synergistic Effects of Type-1/2 Dipeptides in a UBR1-Dependent
**Ubiquitylation System.** Reactions consisting of purified *S. cerevisiae* Ub, purified UBA1 (E1, Ub-activating enzyme), RAD6 (the N-end rule pathway’s E2, Ub-conjugating enzyme), UBR1 (the N-end rule pathway’s E3 Ub ligase), $^{35}$S-labeled, purified CUP9, and ATP were supplemented with pairs of purified dipeptides, at indicated concentrations. The pairs of added dipeptides were Arg-Ala plus Ala-Leu (RA+AL), or Leu-Ala plus Ala-Arg (LA+AR), or Arg-Ala plus Leu-Ala (RA+LA), or Ala-Arg plus Ala-Leu (AR+AL). The reactions were incubated at 30°C for 10 min, followed by SDS-10% PAGE and autoradiography (see Experimental Procedures). Unmodified CUP9 and CUP9-linked poly-Ub chains are indicated on the left. Asterisk on the right denotes a protein present in the initial CUP9 sample, possibly an SDS-resistant dimer of CUP9. Lane 1, purified CUP9 (input). Lanes 2-10, same as in B, but a set of independently performed reactions, with different concentrations of added dipeptide pairs. Note a synergistic enhancement of CUP9 polyubiquitylation in the presence of RA+LA (lane 9).
Relative βgal activity (%)

A

UBR1
D176E-1
G173R-1
G173D-1
CV145YM-1
D318N-2
P406S-2
V313I-2
H321Y-2
E560K-2

B

UBR1
D176E-1
G173R-1
G173D-1
CV145YM-1
D318N-2
P406S-2
V313I-2
H321Y-2
E560K-2

C

Met
Arg
Leu

UBr1Δ

D

UBR1
UBR1D176E-1
UBR1P406S-2

Fig. 2, Xia et al.
Fig. 3, Xia et al.
Fig. 5, Xia et al.
Fig. 6, Xia et al.
Fig. 7, Xia et al.
Fig. 8, Xia et al.