Tuning the Strength of a Bacterial N-end Rule Degradation Signal*

Received for publication, March 20, 2008, and in revised form, June 11, 2008. Published, JBC Papers in Press, June 11, 2008, DOI 10.1074/jbc.M802213200

Kevin H. Wang†1, Elizabeth S. C. Oakes§2, Robert T. Sauer‡2, and Tania A. Baker†3

From the †Department of Biology and §Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

The N-end rule is a degradation pathway conserved from bacteria to mammals that links a protein’s stability in vivo to the identity of its N-terminal residue. In *Escherichia coli*, the components of this pathway directly responsible for protein degradation are the ClpAP protease and its adaptor ClpS. We recently demonstrated that ClpAP is able to recognize N-end motifs in the absence of ClpS although with significantly reduced substrate affinity. In this study, a systematic sequence analysis reveals new features of N-end rule degradation signals. To achieve specificity, recognition of an N-end motif by the protease-adaptor complex uses both the identity of the N-terminal residue and a free α-amino group. Acidic residues near the first residue decrease substrate affinity, demonstrating that the identity of adjacent residues can affect recognition although significant flexibility is tolerated. However, shortening the distance between the N-end residue and the stably folded portion of a protein prevents degradation entirely, indicating that an N-end signal alone is not always sufficient for degradation. Together, these data define in vitro the sequence and structural requirements for the function of bacterial N-end signals.

Regulated proteolysis is fundamental for cellular survival because it provides an irreversible control mechanism. For example, progression through the eukaryotic cell cycle requires timely turnover of cyclins to synchronize and order specific events, such as DNA replication and chromosome segregation (1). Proteolysis also initiates the SOS response to DNA damage in bacteria via degradation of the transcriptional repressor LexA (2, 3), thereby allowing expression of repair polymerases and checkpoint proteins (4). Defective protein turnover can initiate events as diverse as loss of competence in Bacillus subtilis (5) and angiogenesis via Hif1 (hypoxia-inducible factor) in mammals (6). The importance of proteolysis as a regulatory mechanism highlights the need to understand the mechanisms by which these proteases select the right substrates and avoid unintended protein destruction.

Energy-dependent proteases are composed of an oligomeric ATP-dependent unfolding enzyme and an enclosed proteolytic chamber (7). The architecture of this chamber requires that substrates pass through an axial entry gate that is too narrow to allow entry of folded proteins. The unfoldase harnesses the energy of ATP hydrolysis to drive mechanical unfolding of protein substrates and to translocate the resultant denatured polypeptide into the proteolytic chamber for peptide bond cleavage (8–10). In *Escherichia coli*, there are several proteolytic complexes; for example, the ClpX and ClpA ATPases unfold substrates and translocate the polypeptide chains into the ClpP proteolytic chamber for degradation (11–13).

Known bacterial degradation signals vary in sequence complexity and in length from a few amino acids to roughly 10 residues (14). Adaptor proteins also confer or enhance recognition by binding both the substrate and the unfoldase. For example, one region of the SspB adaptor binds to the ClpX unfoldase, and another domain recognizes a region of the ssrA degradation tag, facilitating tethering of ssrA-tagged substrates to ClpXP and the probability of productive engagement (15, 16). SspB and ClpX can bind the ssrA tag simultaneously, allowing SspB to hand off substrates to ClpX directly (17). The sequence determinants for the ssrA-SspB and ssrA-ClpX interactions have been characterized both structurally and biochemically (15, 18, 19). In contrast, the mechanism of adaptor-mediated delivery for substrates to ClpAP is not well understood.

ClpS is a multifaceted adaptor, which enhances ClpAP turnover of N-end rule substrates but also prevents ClpAP from degrading other classes of substrates (20–24). Because ClpS possesses both stimulatory and inhibitory activities, it can change the profile of ClpAP degradation dramatically. The evolutionarily conserved N-end rule relates the intracellular half-life of a protein to its N-terminal residue (25–27). In bacteria, proteins beginning with any of the three aromatic amino acids (Phe, Tyr, or Trp) or the aliphatic residue Leu are degraded by ClpAP with assistance from ClpS (21). Side chain hydrophobicity per se is not sufficient for N-end rule recognition, since Ile, Val, and Met do not target substrates for ClpAP degradation. Substrates with the same N-end rule residue but different adjacent sequences are also degraded with varying rates in vivo, indicating that residues beyond the N terminus affect degradation in the bacterial N-end rule (22). It is known that ClpS binds directly to both ClpAP and N-end rule substrates to enhance protein turnover. ClpAP also shows weak affinity for N-end substrates in the absence of ClpS. These observations raise several questions about the mechanism of N-end rule substrate recognition.
recognition by ClpA and ClpS. What is the molecular basis of the sequence signal that determines how efficiently an N-end motif is recognized? What are the individual contributions of ClpA and ClpS in degrading N-end motif substrates?

To address these questions, we mutagenized an N-end pentapeptide (YLFVQ) that efficiently targets substrates for ClpAP degradation (22) and assayed the effects of an in vitro degradation of GFP fusion proteins by ClpAP in the presence and absence of ClpS. We confirmed the importance of N-terminal Leu, Tyr, Trp, or Phe residues for robust ClpAPS degradation (22). Competition experiments also established that modification of the α-amino group substantially diminished ClpAP recognition. The N-end rule thus uses the combination of the N-terminal residue’s side chain and the α-amino group as the principal recognition determinants of the degradation signal. The positive contributions of these two determinants are antagonized by the presence of acidic residues adjacent to the motif, demonstrating that sequence adjacent to the N-terminal residue influences recognition by ClpAPS. Furthermore, N-end signals are not sufficient to promote degradation if the distance between the folded region of the protein and the N-terminal residue is too short, indicating that there is also a structural component to the N-end rule. Examination of individual contributions of ClpS and ClpAP revealed that ClpS bound poorly to acidic N-end motifs but well to short N-end motifs, whereas ClpAP degraded some acidic N-end substrates efficiently but could not degrade short N-end motifs. We conclude that both ClpS and ClpAP are important in determining the efficacy of N-end substrate processing. These results dissect the bacterial N-end rule into components that are important for recognition in vitro and show how the presence of ClpS alters the sequence selectivity of ClpAP.

**EXPERIMENTAL PROCEDURES**

**Plasmids and Proteins**—GFP variants (GFPuv with serine at position 65 changed to threonine) and Y-titin were cloned into a pET23b.smt3 vector using Agel and NotI sites (22). The N-terminal sequences of GFP variants are shown in Fig. 5A; “ASK” initiates the GFP sequence.

For protein expression, substrates were subcloned into the isopropyl β-D-galactopyranoside-inducible vector pET23b.smt3 (pET23b from Novagen) and transformed into E. coli strain BL21 DE3 (28). His10-SUMO-GFP fusions were purified by Ni2⁺-affinity chromatography as described (28) and were >85% pure. Most contaminants were His10-SUMO or full-length His10-SUMO-GFP. Some GFP variants were purified to >95% purity, using a low substitution phenyl-Sepharose column (GE Healthcare) but were degraded at the same rate as GFP proteins not processed with this second purification step. ClpA, ClpP, ClpS, and 35S-YLFVQ-titin were purified as described (28, 29). GFP (100 μM) was acid-denatured by adding hydrochloric acid to 25 mM for 5 min at room temperature (30).

**GFP Degradation Assays**—Loss of GFP fluorescence in degradation assays was monitored using a Photon Technology International fluorimeter (excitation, 488 nm; emission, 511 nm). ClpA6 (50 nM), (ClpP-His6)14 (100 nM), ClpS (450 nM), and GFP substrate (10 nM) were premixed as described at 30 °C (12). For degradation reactions lacking ClpS, 100 nM ClpA6 and 200 nM ClpP14 were used. In Fig. 3, GFP concentrations from 50 nM to 16 μM were used. To initiate degradation, ATP (4 mM) was added at time 0. Initial changes in fluorescence were calculated from the linear portion of the kinetic trace, typically over the first 3 min, and converted to initial rates of GFP protein degradation using a linear standard curve relating fluorescence at 511 nm to GFP concentrations. For determination of steady-state kinetic parameters in Fig. 3, the average initial rates from three independent experiments were plotted as a function of the total substrate concentration. Since [GFP substrate] was not always in excess of [ClpAPS], the data were fitted ($R^2 > 0.95$) by a nonlinear least squares algorithm to a quadratic version of the Michaelis-Menten equation.

$$V_o = k_{cat} \times \left( \frac{[ClpAPS]}{K_M + [GFP]} + \frac{[GFP]}{K_M + [ClpAPS]} \right) = \frac{[GFP]^2 - 4 \times ([ClpAPS] \times [GFP])^{0.5}}{(2 \times [ClpAPS])^{0.5}}$$  \hspace{1cm} (Eq. 1)

Degradation reactions of unfolded GFP were performed using ClpA6 (800 nM), (ClpP-His6)14 (1.6 μM), and ClpS (4.8 μM). ATP regeneration mixture (4 mM ATP, 50 mg/ml creatine kinase, and 5 mM creatine phosphate) was added prior to the addition of unfolded GFP, and the reaction was incubated at 30 °C for 2 min. Unfolded YAC-GFP (1.5 μM) was added at time 0 to initiate the reaction. At each time point, 10 μl of reaction mix was quenched by adding 2.5 μl of SDS loading buffer on ice. Samples were boiled and electrophoresed on a 15% Tris-glycine gel, which was stained with Sypro Orange (Molecular Probes) at a 1:5000 dilution in 7.5% acetic acid and scanned on a Typhoon 9400 imager (excitation, 488 nm; detection, 555 nm). Quantification was performed with ImageQuant 4.0, and intensities were normalized to the ClpP intensity in each lane. Three independent experiments were performed. A representative gel and quantification are shown in Fig. 5C.

**Peptide Competition Assays**—Peptide competition assays were performed by assaying loss of YLFVQ-GFP fluorescence. β-Galactosidase peptides were synthesized by the MIT Biopolymers facility and contained the first 21 residues of β-galactosidase fused to different N-terminal residues. These peptides were added to a final concentration of 50 μM in reactions containing 50/100/450 nM ClpA/P/S and 500 nM YLFVQ-GFP; degradation was started by adding ATP.

The YLFVQ peptide was acetylated by incubating 1 mM peptide in 10 mM Tris (pH 8.9) with 200 mM acetic anhydride overnight at room temperature. Acetyl-YLFVQ was purified by HPLC, lyophilized, and resuspended in H2O. The addition of a single acetyl group was verified by MALDI-TOF mass spectrometry. Acetyl-YLFVQ or unmodified YLFVQ peptide was added to degradation reactions containing 50/100/450 nM ClpA/P/S, 35 nM YLFVQ-GFP, and 3.3 or 6.6 μM peptide. The initial degradation rate in the absence of peptide was normal-

---

The abbreviations used are: GFP, green fluorescent protein; SUMO, small ubiquitin-like modifier; HPLC, high pressure liquid chromatography; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; E3, ubiquitin-protein isopeptide ligase.
Sequence Rules for the N-end Rule Degradation Pathway

![Graphs and images related to the N-end rule degradation pathway.](image)

**FIGURE 1.** The N-end rule depends on the identity of the first residue and the α-amino group. A, N-end GFP constructs are degraded, whereas the non-N-end ILFVQ-GFP is not. 50 nM ClpAPS was incubated with 10 nM GFP, and loss of GFP fluorescence was followed by excitation at 488 and detection at 511 nm. B, peptide competition of YLFVQ-GFP (500 nM) degradation by 50 nM ClpAPS. YLFVQR peptide with a free or acetylated α-amino group (Ac-YLFVQR) was added before initiation of degradation, and rates were normalized to that of a reaction lacking peptide. N.D., the reaction using 6.6 μM YLFVQR peptide competitor was not performed. C, MALDI-TOF spectra of unmodified versus acetylated YLFVQR peptide.

**RESULTS**

Specific Side Chains at the N-terminal Residue Are Critical for Recognition—To allow the N-terminal sequence of a model substrate (GFP) to be modified without constraints imposed by translational initiation or post-translational processing, we constructed and purified variants as His$_6$-SUMO-X$_{7}$-GFP fusions, cleaved these proteins with SUMO protease (28), and repurified the X$_7$-GFP molecule to remove the protease and His$_6$-SUMO fragment. The strong N-end motif YLFVQEL was used as a reference X$_7$ sequence (22); the glutamate-leucine was encoded by a SacI restriction site to facilitate cloning. Variants with the first Tyr replaced by other N-end rule residues (Phe, Leu, or Trp), by aliphatic side chains (Ile or Val), or by Thr were also constructed and purified. At low substrate concentrations where the rate of degradation by ClpAPS (ClpAP plus ClpS) was determined by the second order rate constant ($k_{cat}/K_m$), only the N-end rule substrates were degraded efficiently (Fig. 1A) (data not shown), consistent with the reported selectivity of the N-end rule (26). Among good substrates, the variant with Phe at the N terminus was degraded most rapidly, whereas the variant with Tyr was degraded at the slowest rate. This difference arose from a ~20% reduction in $K_m$ but not $V_{max}$ (not shown), suggesting modest differences in recognition of N-end residues by ClpAPS.

As another probe of the importance of the N-terminal residue, we assayed ClpAPS degradation of YLFVQEL-GFP in the presence of a large excess of peptide competitor consisting of a variable N-terminal residue followed by the first 21 residues of E. coli β-galactosidase. Efficient competition was observed...
when the N-end rule residues Tyr or Leu were at the N terminus but not when Met or Arg occupied this position (Fig. 1B). Therefore, both direct degradation and competition assays can be applied to probe the sequence rules of the N-end signal.

The α-Amino Group Is a Recognition Element of the N-end Rule—The purified precursor His6-SUMO-YLFVQEL-GFP protein was not degraded by ClpAPS (not shown), suggesting that the lack of a free N-terminal Tyr and/or the presence of “upstream” residues prevents recognition. To test the importance of a free α-amino group, we compared inhibition of ClpAPS degradation of YLFVQEL-GFP by the hexapeptide YLFVQR before and after blocking its N terminus by treatment with acetic anhydride. The unmodified peptide was a good substrate degradation by an acidic residue is stronger when placed closer to the N terminus. D, an acidic residue but not Gln at position 2 or multiple Glu residues at positions 3–5 inhibits degradation.

Acidic Residues near the N-end Residue Weaken ClpAPS and ClpS Binding—We previously found that substrates with the same N-end residue but different neighboring sequences were degraded with different $K_m$ values, suggesting that residues beyond the N terminus affect functional interactions with ClpAPS (22). To probe whether these effects are caused by packing or electrostatic interactions, we individually changed residues 2, 3, 4, and 5 of YLFVQEL-GFP to a basic residue (Arg), a small residue (Ala), or an acidic residue (Glu). When low concentrations of these substrates were tested for ClpAPS degradation, the Arg and Ala variants were degraded at rates similar to YLFVQEL-GFP (Fig. 2, A and B), indicating that ClpAPS does not require specific side chains at positions 2–5 for efficient N-end degradation. By contrast, changing residue 2, 3, or 4 to Glu slowed degradation (Fig. 2C), with the largest effect observed when Glu was adjacent to the N-end residue. Indeed, replacing residue 2 with either Glu or Asp slowed degradation more than 10-fold, whereas changing this residue to Gln had only a small effect (Fig. 2D). Thus, the negative charge and not the shape of the position 2 side chain causes poor degradation by ClpAPS. A variant with residues 3–5 replaced by Glu (YLFEEEL-GFP) was degraded very slowly, suggesting that a net negative charge near the N-end residue is poorly tolerated by ClpAPS (Fig. 2D).

To determine if the deleterious effects of acidic residues arose from poor substrate binding or slower turnover by ClpAPS, we determined steady-state kinetic parameters for the parental substrate YLFVQEL-GFP ($K_m = 26 \text{ nM}$; $V_{\text{max}} = 1.2 \text{ min}^{-1}$) and for YEFVQEL-GFP ($K_m = 1400 \text{ nM}$; $V_{\text{max}} = 1.4 \text{ min}^{-1}$) (Fig. 3). These results show that the principal effect of the Leu$^2$→Glu substitution is an approximately 50-fold weakening of apparent affinity of the substrate for ClpAPS. We conclude that acidic side chains at residues 2–4 of N-end degradation signals interfere with ClpAPS binding but not processing. The “N-end receptor sites” in ClpS and/or ClpA may have a negative electrostatic potential that interacts unfavorably with negatively charged residues in the N-end signal.

To examine the relative affinity of ClpS for acidic N-end signals, surface plasmon resonance was used to monitor binding of immobilized ClpS to YEFVQEL-GFP and YLFVEEL-GFP. The YEFVQEL-GFP protein, which was degraded slowly by ClpAPS, also bound poorly to ClpS (Fig. 4A). By contrast, YLFVEEL-GFP has the same net charge but was degraded 6-fold more rapidly by ClpAPS (Fig. 2C) and bound well to ClpS. These results show that acidic residues near the N-end residue influence ClpAPS degradation, at least in part, by weakening ClpS binding and also demonstrate that ClpS binding affinity is correlated with ClpAPS degradation activity of acidic N-end substrates.

Is this ClpS binding defect entirely responsible for the slow degradation of YEFVQEL-GFP by ClpAPS? ClpAP can degrade N-end substrates without ClpS, but with 10–70-fold weaker apparent affinity than ClpAPS, depending on the sequence of
the N-end signal (22). We found that ClpAP degraded both YEFVQEL-GFP and YLFVQEL-GFP at similar rates, indicating that ClpAP itself is not inhibited by an acidic residue at position 2 (Fig. 4B). In contrast, the presence of several acidic residues (YLSEEEL-GFP) near the N-end residue slowed substrate degradation by ClpAP. These results indicate that acidic N-end signals affect ClpAPS and ClpAP recognition differently.

**Length Determinants of N-end Signals**—Erbse et al. (21) found that ClpAPS degraded GFP with an N-end Phe followed by a 10-residue linker but not when the N-terminal Phe was placed adjacent to GFP. In our YLFVQEL-GFP construct, the N-terminal Tyr is seven residues from the alanine that begins the GFP sequence (YLFFVQELSK); the lysine begins the folded region of GFP. To address the role of linker length, we determined rates of ClpAPS degradation of constructs with six alanines between the N-terminal Tyr and the first residue of GFP (YA6-GFP) and variants with the linker reduced by two (YA5-GFP) or three residues (YA4-GFP). The YA6-GFP substrate was degraded by ClpAP with a K_m of 140 nM, a value 6-fold higher than the K_m for YLFVQEL-GFP (Fig. 3). This result is consistent with a modest contribution of residues beyond the N terminus to ClpAPS interactions. The YA4-GFP substrate was degraded about 5-fold more slowly than YA6-GFP (Fig. 5A), showing that linker length influences degradation. No degradation of YA3-GFP was detected even at high substrate concentrations (Fig. 5B). Thus, GFP N-end tags that are too short pose a problem for ClpAPS.

To determine if this defect is due to the proximity of a folded domain adjacent to the YA3 N-end signal, YA3-GFP was acid-denatured prior to the addition into a degradation reaction containing ClpAPS. Unfolded YA3-GFP was degraded rapidly, whereas native YA3-GFP was not turned over even using increased ClpAPS concentrations (Fig. 5C). This result indicates that N-end signals are not effective degradation motifs when located too close to the folded N-terminal region of the substrate. To establish whether ClpS or ClpAP is responsible for this observation, experiments were designed to test the roles of both recognition modules.

**Dissecting the Individual Contributions of ClpS and ClpAP to Substrate Recognition**—In principle, ClpAPS might fail to degrade native YA3-GFP either because ClpS does not bind this protein and/or because ClpA cannot accept this protein from ClpS or cannot unfold it after transfer. In surface plasmon resonance assays, immobilized ClpS bound YLFVQEL-GFP and YA3-GFP to comparable extents but bound very poorly to ILFVQEL-GFP, a non N-end rule protein (Fig. 4A, 6A). To verify that ClpS is selective for N-end residues, fluorescinated peptides with N-terminal Phe, Tyr, Trp, and Ile were incubated with ClpS, and fluorescence anisotropy was measured (Fig. 6B). All peptides except for the Ile variant produced an increase in anisotropy when ClpS was added, indicating that ClpS does not recognize an N-terminal Ile. Together, these data show that the inability of ClpAPS to degrade YA3-GFP does not arise from a ClpS binding defect (Fig. 6C).

We determined rates of ClpAP degradation of YA3-GFP, ILFVQEL-GFP, LLFVQEL-GFP, and YLFVQEL-GFP in the absence of ClpS (Fig. 6D). Under these conditions, the non N-end substrate (ILFVQEL-GFP) was degraded at a rate similar to the two good N-end rule substrates (LLFVQEL-GFP and YLFVQEL-GFP). However, ClpAP did not degrade the short tag variant YA3-GFP (Fig. 6D). Thus, N-end residues located too close to the folded region of GFP do not serve as degradation signals for ClpAP or for ClpAPS.

The preceding experiments suggest that the defect in YA3-GFP degradation arises after binding to ClpS. If this model is correct, then YA3-GFP should act as a competitor and inhibit ClpAPS degradation of another N-end rule substrate by blocking the N-end recognition site on ClpS. In contrast, because ILFVQEL-GFP is not recognized by ClpS but is degraded by ClpAP, this substrate may compete by occupying the degrada-
GFP competes with the $^{35}$S-substrate for binding ClpS but also competes for a binding site in ClpAP, thereby inhibiting both initial recognition and subsequent unfolding and degradation by ClpAP. By contrast, ILFVQEL-GFP is a weaker inhibitor, because it only competes for ClpAP binding, and YA$_3$-GFP is a weaker inhibitor, because it only competes for ClpS binding. Together with the results from Fig. 6, these data suggest that the efficiency of ClpAPS in degrading N-end rule substrates depends on recognition of the substrate by both ClpS and by ClpAP.

DISCUSSION

The original discovery of the bacterial N-end rule identified four N-terminal residues (Leu, Phe, Trp, and Tyr) that target β-galactosidase for degradation (26). It is now known that the ClpA unfoldase and the ClpS adaptor participate in recognition of N-end rule signals (21, 22, 26). The results presented here further define the molecular basis for N-end rule sequence selectivity and the roles of ClpA and ClpS in recognition.

We confirmed that the expected N-end residues mediated ClpAPS degradation of GFP variants with modest differences in efficiency in the order Phe > Leu > Trp > Tyr. By contrast, GFP with an N-terminal Ile showed no detectable ClpAPS degradation at low concentrations, where N-end rule substrates were efficiently degraded. Thus, ClpAPS recognition is highly selective, discriminating between side chains as similar as Leu and Ile. In addition to the N-terminal side chain, we find that a free α-amino group contributes to but is not essential for ClpAPS binding. This finding is consistent with studies showing that blocking the N terminus of an otherwise good N-end rule signal reduced ClpS binding on a peptide blot (21). ClpS is required for high affinity interactions with N-end rule substrates (21, 22), and we find that ClpS alone discriminates between substrates with good N-end rule residues and those with Ile at the N terminus. Thus, ClpS enhances the degradation of N-end substrates by ClpAP by recognizing the α-amino group in combination with a Leu, Phe, Trp, or Tyr side chain at the N terminus (Fig. 8A).

Our results show that residues adjacent to the N-end residue influence the affinity of ClpAPS interactions. Specific side chains at these positions are not required. For example, changing the N-end signal of YLFVQEL-GFP to YAAAAAAAA-GFP increased the $K_m$ for degradation only 6-fold, indicating that residues make small contributions to apparent affinity. Notably, however, an acidic residue at position 2 (YEFVQEL-GFP) increased $K_m$ 50-fold; a variant with acidic residues at positions 3–5 (YLEEEL-GFP) was also a very poor ClpAPS substrate. These effects are probably caused by repulsion between acidic residues in ClpAPS and those in these N-end signals, which is consistent with the slower in vivo degradation of substrates carrying acidic N-end sequences (22). Indeed, we found that ClpS alone bound YEFVQEL-GFP very poorly compared with YLFVEEL-GFP, and mutational studies suggest that Asp$^{35}$ and Asp$^{36}$ of ClpS form part of its binding site for N-end signals (21). Negative electrostatic potential in this binding site would help bind the positively charged α-amino group of N-end signals. Moreover, some endogenous N-end signals contain a basic residue at position 2, because aminoacyl transferase adds
The degradation of ClpAP and ClpS to these degradation motifs are signals with N-terminal residues that have weaker affinities for ClpAPS because of the presence of negatively charged residues; both the number and positions of acidic residues appear to determine precise affinity. At the other extreme are strong N-end signals, such as YLFVQEL, that allow efficient ClpAPS degradation at nanomolar substrate concentrations.

Our results also raise several questions regarding the eukaryotic N-end rule, which recognizes the additional N-end residues Ile (in some organisms), Arg, Lys, and His. The N-end signal receptor Ubr1 is an E3 ubiquitin ligase that possesses a binding site for N-terminal Phe, Leu, Trp, Tyr, and Ile and a separate site for N-terminal Arg, Lys, and His (31). Interestingly, N-terminal Asp and Glu are recognized by the arginyl transferase Ate1p (32), which conjugates an Arg residue to these N termini. Is an N-terminal Arg residue recognized less efficiently when the second residue is acidic? If acidic residues in Ubr1 are important for docking the α-amino group and the Arg for this type of N-end signal, then electrostatic repulsion with acidic residues on the N-end signal may reduce binding affinity just as in the case of ClpS. Additionally, shorter N-end sequences bound by Ubr1, and are these substrates ubiquitinated efficiently? Does the proteasome possess the same steric requirements for N-end signal length as ClpAP, although N-end substrates carry polyubiquitin chains as their proteasome localization determinants?

At present, it is not known how N-end substrates for ClpAPS are generated in the cell. Proteins with good N-end residues do not arise from translation and normal post-translational processing, because the initiator formyl-Met of proteins with second residues Phe, Leu, Trp, or Tyr is not removed by methionine aminopeptidase (33). The next challenge will be to isolate endogenous N-end substrates and to determine the extent and impact of sequence control in the N-end rule degradation.

Acknowledgments—We thank members of the Baker and Sauer laboratories for encouragement and helpful discussions, in particular J. Hou for purified ClpS protein, P. Chien for untagged GFP protein, and I. Levchenko and G. Roman for synthesizing peptides. We are grateful to C. Köhrer and U. Rajbhandary for use of the Biacore 3000 instrument and M. Pruteanu, A. Meyer, and S. Bell for critical reading of the manuscript.

REFERENCES

1. Hershko, A. (1997) Curr. Opin. Cell Biol. 9, 788–799
2. Little, J. W. (1983) J. Mol. Biol., 167, 791–808
3. Neher, S. B., Flynn, J. M., Sauer, R. T., and Baker, T. A. (2003) Genes Dev. 17, 1084–1089
4. Walker, G. C., Marsh, L., and Dodson, L. (1985) Environ. Health Perspect. 62, 115–117
5. Turgay, K., Hahn, J., Burghoorn, J., and Dubnau, D. (1998) EMBO J. 17, 6730–6738
6. Willam, C., Masson, N., Tian, Y. M., Mahmood, S. A., Wilson, M. L, Bicknell, R., Eckardt, K. U., Maxwell, P. H., Ratcliffe, P. J., and Pugh, C. W. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 10423–10428
7. Sauer, R. T., Bolon, D. N., Burton, B. M., Burton, R. E., Flynn, J. M., Grant, R. A., Hersch, G. L., Joshi, S. A., Ken constantly, J. A., Levenchon, I., Neher, S. B., Oakes, E. S., Siddiqui, S. M., Wah, D. A., and Baker, T. A. (2004) Cell 119, 9–18
8. Benaroudj, N., Tarcsa, E., Cascio, P., and Goldberg, A. L. (2001) Biochimie (Paris) 83, 311–318
9. Singh, S. K., Grimaud, R., Hoskins, J. R., Wickner, S., and Maurizi, M. R. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 8898–8903
10. Wang, J., Hartling, J. A., and Flanagan, J. M. (1997) Cell 91, 447–456
11. Kim, Y. I., Burton, R. E., Burton, B. M., Sauer, R. T., and Baker, T. A. (2000) Mol. Cell 5, 639–648
12. Weber-Ban, E. U., Reid, B. G., Miranker, A. D., and Horwich, A. L. (1999) Nature 401, 90–93
13. Gottesman, S., Roche, E., Zhou, Y., and Sauer, R. T. (1998) Genes Dev. 12, 1338–1347
14. Flynn, J. M., Neher, S. B., Kim, Y. I., Sauer, R. T., and Baker, T. A. (2003) Mol. Cell 11, 671–683
15. Flynn, J. M., Levchenko, I., Seidel, M., Wickner, S. H., Sauer, R. T., and Baker, T. A. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 10584–10589
16. Wah, D. A., Levchenko, I., Rieckhof, G. E., Bolon, D. N., Baker, T. A., and Sauer, R. T. (2003) Mol. Cell 12, 355–363
17. Bolon, D. N., Grant, R. A., Baker, T. A., and Sauer, R. T. (2004) Mol. Cell 16, 343–350
18. Levchenko, I., Grant, R. A., Wah, D. A., Sauer, R. T., and Baker, T. A. (2003) Mol. Cell 12, 365–372
19. Song, H. K., and Eck, M. J. (2003) Mol. Cell 12, 75–86
20. Dougan, D. A., Reid, B. G., Horwich, A. L., and Bukau, B. (2002) Mol. Cell 9, 673–683
21. Erbse, A., Schmidt, R., Bornemann, T., Schneider-Mergener, J., Mogk, A., Zahn, R., Dougan, D. A., and Bukau, B. (2006) Nature 439, 753–756
22. Wang, K. H., Sauer, R. T., and Baker, T. A. (2007) Genes Dev. 21, 403–408
23. Katayama, Y., Gottesman, S., Pumphrey, J., Rudikoff, S., Clark, W. P., and Maurizi, M. R. (1988) J. Biol. Chem. 263, 15226–15236
24. Hou, J. Y., Sauer, R. T., and Baker, T. A. (2008) Nat. Struct. Mol. Biol. 15, 288–294
25. Bachmair, A., Finley, D., and Varshavsky, A. (1986) Science 234, 179–186
26. Tobias, J. W., Shrader, T. E., Rocap, G., and Varshavsky, A. (1991) Science 254, 1374–1377
27. Varshavsky, A. (1992) Cell 69, 725–735
28. Malakhov, M. P., Mattern, M. R., Malakhova, O. A., Drinker, M., Weeks, S. D., and Butt, T. R. (2004) J. Struct. Funct. Genomics 5, 75–86
29. Kenniston, J. A., Baker, T. A., Fernandez, J. M., and Sauer, R. T. (2003) Cell 114, 511–520
30. Hoskins, J. R., Singh, S. K., Maurizi, M. R., and Wickner, S. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 8892–8897
31. Varshavsky, A. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 12142–12149
32. Balzi, E., Choder, M., Chen, W. N., Varshavsky, A., and Goffeau, A. (1990) J. Biol. Chem. 265, 7464–7471
33. Flinta, C., Persson, B., Jornvall, H., and von Heijne, G. (1986) Eur. J. Biochem. 154, 193–196