



# HHS Public Access

Author manuscript

*Nat Chem Biol.* Author manuscript; available in PMC 2011 September 01.

Published in final edited form as:

*Nat Chem Biol.* 2011 March ; 7(3): 161–167. doi:10.1038/nchembio.521.

## Defining the geometry of the two-component proteasome degron

Tomonao Inobe<sup>1,2</sup>, Susan Fishbain<sup>1</sup>, Sumit Prakash<sup>1</sup>, and Andreas Matouschek<sup>1</sup>

<sup>1</sup>Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, 2205 Tech Drive, Evanston, IL 60208 and Robert H. Lurie Comprehensive Cancer Center, Northwestern University, 303 East Superior Street L3-125, Chicago, IL 60611, USA

<sup>2</sup>Laboratory for Structural Neuropathology, RIKEN Brain Science Institute, Wako, Saitama 351-0198, Japan

### Abstract

The eukaryotic 26S proteasome controls cellular processes by degrading specific regulatory proteins. Most proteins are targeted for degradation by a signal or degron that consists of two parts: a proteasome-binding tag, typically covalently attached polyubiquitin chains, and an unstructured region that serves as the initiation region for proteasomal proteolysis. Here we have characterized how the arrangement of the two degron parts in a protein affects degradation. We found that a substrate is degraded efficiently only when its initiation region is of a certain minimal length and is appropriately separated in space from the proteasome-binding tag. Regions that are located too close or too far from the proteasome-binding tag cannot access the proteasome and induce degradation. These spacing requirements are different for a polyubiquitin chain and a ubiquitin-like (UbL) domain. Thus, arrangement and location of the proteasome initiation region affect a protein's fate and play a central role in selecting proteins for proteasome-mediated degradation.

### Introduction

The ubiquitin-proteasome system (UPS) plays a central part in cellular regulation and is involved in many diseases<sup>1</sup>. It degrades short-lived regulatory proteins in cellular processes such as signal transduction, cell cycle regulation, and transcription. In addition, it clears the cell of misfolded and damaged proteins and produces some of the peptides displayed at the cell surface as part of the adaptive immune response. Proteolysis occurs in a roughly 2,500 kDa large protein machine known as the proteasome. The proteasome is located in the cytosol and nucleus of cells and thus faces the challenge of having to be able to degrade a vast number of unrelated proteins but to do so with exquisite specificity.

Users may view, print, copy, download and text and data- mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

Correspondence should be addressed to A. M. ([matouschek@northwestern.edu](mailto:matouschek@northwestern.edu)).

**AUTHOR CONTRIBUTIONS and COMPETING FINANCIAL INTERESTS** T.I., S.F., S.P. and A.M. have designed and interpreted the experiments and wrote the manuscript. T.I., S.F., and S.P. performed the experiments.

The authors have no competing financial interests.

Rapidly degraded proteasome substrates usually contain a degradation signal or degron that has two parts: a proteasome-binding tag and a proteasome initiation region<sup>2-4</sup>. The proteasome-binding tag is a polyubiquitin chain attached to the  $\epsilon$ -amino group of a lysine residue in most known proteasome substrates. A polyubiquitin chain of at least four molecules<sup>5,6</sup> allows the proteasome to recognize the substrate through its Rpn10, Rpn13, and perhaps Rpt5 subunits<sup>6-8</sup>. Once recognized, the substrate is unfolded and degraded into small peptides. Some substrates are brought to the proteasome by adaptor proteins such as Rad23, Dsk2, or Ddi1 [9,10]. These adaptors bind polyubiquitin chains through one or two ubiquitin-associated (UBA) domains and the proteasome through a ubiquitin-like (UbL) domain. The UbL domain is recognized by the Rpn1, Rpn13, and human but not yeast Rpn10 subunits on the proteasome<sup>8,11-13</sup>.

Substrate binding to the proteasome is not enough to ensure degradation. In addition to the binding tag, the substrate must contain an unstructured region that can serve as the initiation region where the proteasome engages the substrate and begins proteolysis<sup>2,3</sup>. The two parts of the degron can still function together when separated onto different polypeptide chains that form a complex<sup>14</sup>. The proteasome can then degrade either the subunit with or without the ubiquitin tag, or both. The selection of which subunit to digest appears to depend on properties of the initiation regions.

Many physiological proteasome substrates are part of larger complexes from which the proteasome can extract and degrade individual subunits<sup>15,16</sup>. For example, the complexes formed by cyclins, cyclin-dependent kinases (Cdks) and Cdk inhibitors (Ckis) such as Sic1 and p27<sup>Kip1</sup> are classical examples of structures that are remodeled by the proteasome. During different phases of the cell cycle, Sic1 and cyclin are specifically ubiquitinated and degraded from the complex while other components remain stable<sup>17,18</sup>.

Here we describe a new rule that governs how the proteasome chooses its substrates. We do so by measuring the efficacy of initiation regions in proteasome degrons in an array of model substrates. We find that to be effective initiation regions need to be located at the appropriate distance relative to the proteasome-binding tag. If the substrate binds to the proteasome through a ubiquitin tag, initiation regions immediately adjacent to the ubiquitin function in degradation. In contrast, if the substrate is targeted through a UbL tag, the initiation region must be separated in space from the UbL domain to function. Our findings suggest that substrate binding and degradation initiation occur at separate sites on the proteasome. The spacing rules fit well with the way ubiquitin and UbL tags are used physiologically and help explain how substrates are selected for degradation or manage to escape proteolysis. They also help explain the mechanism by which the proteasome remodels protein complexes by selecting only specific subunits for degradation.

## Results

### Length of initiation regions

To define the rules that determine how well unstructured regions in proteins can serve as proteasome initiation sites, we first measured their minimum length required to support rapid degradation. For this purpose, we constructed a series of proteasome substrates

containing a central dihydrofolate reductase (DHFR) domain. The proteasome targeting part of the degron was located at the N-terminus of DHFR, and the initiation region was located at the C-terminus (Fig. 1a). We then investigated how effectively a series of different degrons targets the DHFR domain for degradation by purified yeast (*Saccharomyces cerevisiae*) proteasome. In one set of constructs, the proteasome-binding tag consisted of four ubiquitin moieties fused in frame N- to C- terminus (Ub<sub>4</sub> tag)<sup>19</sup>. This tag mimics a polyubiquitin chain and probably most closely resembles ubiquitin moieties linked through lysine 63 of ubiquitin<sup>20,21</sup>. In the other set of constructs, the targeting tag was a single UbL domain derived from yeast Rad23<sup>22,23</sup> (UbL tag).

The unstructured tails were derived from either yeast cytochrome *b*<sub>2</sub> or subunit 9 of mold (*Neurospora crassa*) F<sub>0</sub> ATPase because these peptides are known to be unstructured and soluble under standard solution conditions. The tails consisted of 15 to 102 amino acids derived from either protein, and the cytochrome *b*<sub>2</sub> sequences ended in a hexahistidine tag (Fig. 1a).

We measured the effect of the length of the initiation regions in the substrates on degradation by the purified yeast proteasome (Fig. 1b,c). The substrates were synthesized by *in vitro* transcription and translation in *Escherichia coli* or reticulocyte lysate and partially purified. The degradation rates for proteins from both expression systems were equivalent. Thus, it is unlikely that contaminating factors from the reticulocyte lysate expression system, such as UbL-UBA proteins and p97/VCP, affect degradation in these assays. Instead, the proteasome most likely recognized the substrate directly. As expected for proteasomal degradation, the depletion of ATP and the addition of proteasome inhibitor MG132 inhibited proteolysis and proteolysis required the Ub<sub>4</sub> or UbL tags (data not shown). The DHFR domain in these constructs is folded and methotrexate stabilizes it against unfolding and degradation<sup>24</sup> (Supplementary Fig. 1).

Shortening the unstructured region in both Ub<sub>4</sub>- and UbL-tagged substrates decreased degradation rates and finally led to the stabilization of the proteins once a certain critical length of the initiation region was reached (Fig. 1b-d). For Ub<sub>4</sub>-tagged DHFR substrates, the transition occurred between 29 and 34 amino acid long initiation regions: proteins with initiation regions of 34 amino acids or longer were degraded, whereas substrates with initiation regions of 29 amino acids or less escaped degradation (Fig. 1b, d), in agreement with earlier studies<sup>25</sup>. For UbL-tagged substrates, the transition occurred between 34 and 44 amino acids: constructs with initiation regions of 44 amino acids or more were degraded rapidly, whereas constructs with initiation regions of 34 amino acids or less were not (Fig. 1c,d). The difference in critical length does not appear to be due to any specific and strongly pronounced sequence preferences of the proteasome because experiments with initiation regions with different amino acid sequences yield equivalent results (Supplementary Fig. 2). Therefore, degradation initiation in these experiments requires an initiation region of a certain minimum length and this critical length is greater for substrates with the UbL tag than for substrates with the Ub<sub>4</sub> tag.

## Moving the initiation region by inserting spacer domains

The requirement for a 34 or 44 amino acid long initiation region suggests one of two scenarios: 1) the receptor on the proteasome recognizes a long peptide region in the substrate on the order of some 30 to 40 amino acids; or 2) the receptor for the initiation region is a certain distance from the ubiquitin or UbL binding site on the proteasome and requires part of the 34 / 44 amino acid long sequence to bridge the distance between the two locations on the proteasome. Therefore, we tested how degradation rates changed when we moved the initiation region away from the proteasome-binding tag (Fig. 2a).

The N- and C-termini of DHFR are close to each other on the same surface of the protein (Fig. 2b) so that a proteasome-binding tag at DHFR's N-terminus and an initiation region at the C-terminus are separate by only about 13 Å. To move the two components of the degron farther apart, we inserted up to four copies of the 27<sup>th</sup> immunoglobulin-like (I27) domain of the muscle protein titin as spacers between the DHFR domain and the initiation region (Fig. 2a). The structure and stability of the titin I27 domain have been investigated in many biochemical and biophysical studies (*e.g.*, [26,27]). Connecting these I27 domains with short linkers should create an elongated structure because the N- and C-termini are on opposite sides of the domain separated by 44 Å (Fig. 2c; [28,29]). Thus, each I27 insertion into our substrate proteins should spread proteasome-binding tag and initiation region farther apart (Fig. 2). The actual length increase per insert depends on the stiffness of the connection between the I27 domains<sup>29</sup>. For flexible connections, a simple estimate assuming a random walk of freely-jointed rigid I27 domains suggests that the insertion of one domain increases the distance by roughly 44 Å, two domains by roughly 71 Å, and three domains by roughly 87 Å<sup>29</sup>. For rigid and straight connections each insertion can add roughly 44 Å distance<sup>29</sup>.

To test whether the I27 domain insertions did indeed separate the two parts of the degron, we measured fluorescence resonance energy transfer (FRET) between fluorescent probes at both ends of the molecule. We replaced the proteasome binding tag with an enhanced green fluorescent protein (eGFP<sup>30</sup>) domain, inserted a binding motif for the fluorescent dye ReAsH<sup>31</sup> at the C-terminus of the initiation region, and measured FRET between eGFP and ReAsH<sup>31</sup> (Fig. 3a). Excitation of eGFP in results in the emission of green light. In double-labeled proteins, the presence of ReAsH resulted in FRET between the two fluorophores, which could be observed by strong quenching of the green eGFP fluorescence and weaker red light emission by ReAsH (Fig. 3b). Increasing the number of inserted I27 domains in the substrates decreased the FRET efficiency (Fig. 3c), which indicates that the insertions do indeed increase the distance between the two ends of the molecule.

## Optimizing degron spacing for degradation

Moving proteasome binding tag and initiation region apart affected degradation rates differently for the two tags (Fig. 4a-c). In substrates without I27 domain insertions, the proteasome-binding tag and initiation region were located next to each other (Fig. 2b) and proteins with a Ub<sub>4</sub> tag and a 44 amino acid long initiation region were degraded efficiently (Fig. 4a,c). Moving the initiation region away from the Ub<sub>4</sub> tag by inserting I27 domains inhibited and finally abolished degradation (Fig. 4a,c).

The situation was different for the UbL tag. The UbL tag did not target the DHFR substrates for degradation when placed immediately adjacent to a 44 amino acid long initiation region (Fig. 4b,c). Degradation improved with the insertion of one I27 domain and reached a maximum with two I27 domains. Separating the UbL tag and initiation region further with three or more I27 domains, inhibited degradation again (Fig. 4b,c). Thus, the optimal spacing between the proteasome-binding tag and initiation region was different for ubiquitin- and UbL-tagged substrates (Fig. 4c). For the ubiquitin tag, the initiation region had to be close to the proteasome-binding group to function. For the UbL tag, the initiation region and the proteasome-binding tag had to be separated by a certain distance to function together effectively. Both shorter and longer distances inhibited degradation.

One possible explanation for the shorter spacing requirement for Ub<sub>4</sub>-tagged proteins could be that Ub<sub>4</sub> tag binds to the same site on the proteasome as the UbL tag and does so through the first (most N-terminal) ubiquitin domain in the tag. In this case, the other ubiquitin moieties in the Ub<sub>4</sub> tag then function as spacers, effectively fulfilling the same role as the titin insertions in the UbL constructs. However, two results make this explanation unlikely. First, replacing the Ub<sub>4</sub> tag with a single ubiquitin moiety abolished degradation, suggesting that a single ubiquitin was not able to bind the proteasome in the same mode as the UbL domain (Supplementary Fig. 3a), and second, substrates with a Ub<sub>2</sub> tag, which consists of only two ubiquitin moieties and is only half as long as a Ub<sub>4</sub> tag, showed the same spacing requirements as substrates with a Ub<sub>4</sub> tag (Supplementary Fig. 3b).

Another concern is that cells contain the UbL-UBA adaptor proteins Rad23, Dsk2, and Ddi1 as substoichiometric proteasome subunits and these may affect substrates with Ub<sub>4</sub> and UbL tags differently. However, experiments with proteasome that lacked UbL-UBA adaptors because it was prepared from a yeast strain in which the genes for Rad23, Dsk2, and Ddi1 were disrupted (provided by S. Elsassner and D. Finley, Harvard Medical School) show the same dependence of degradation rates on spacing for ubiquitin (Supplementary Fig. 4a) and UbL degrons (Supplementary Fig. 4b) as experiments with proteasome purified from a Rad23, Dsk2, and Ddi1 wildtype strain. Thus, we conclude that the different spacing requirements for the two degrons are caused by different receptors for the Ub<sub>4</sub> and UbL tags on the proteasome.

### Degron spacing and initiation region length

The results so far suggest that the proteasome recognizes the two parts of the degron, the proteasome-binding tag and the initiation region, with distinct binding sites at two separate locations on the proteasome particle. If this model were correct, it should be possible to compensate for incorrect spacing by increasing the length of the initiation region. Similarly, decreasing the length of the initiation region should make degradation more sensitive to the correct spacing of the degron parts in the substrate. To test this prediction, we combined initiation regions of different lengths with proteasome-binding tags at different spacings (*i.e.*, with insertions of different numbers of I27 domains; Fig. 5a).

Ubiquitin degrons supported degradation best when the Ub<sub>4</sub> tag and initiation region were close to each other so that separating them by three titin domains inhibits degradation (Fig. 4a). Increasing the length of the initiation region from 44 amino acids to 102 overcame the

inhibition (Fig. 5b). This pattern held when the full array of substrates was analyzed and longer initiation regions were required for degradation for proteins with unfavorable degron spacings, as predicted (Fig. 5c-d). Similarly, degradation of substrates with longer initiation regions was less sensitive to titin insertions than degradation of substrates with short initiation regions (Fig. 5e-f). For example, substrates with the UbL tag showed a peak of optimal degradation around two I27 insertions when the substrates contained a 34 or 44 amino acid initiation region. A longer initiation region of 102 amino acids caused the peak to broaden so that even constructs without a titin insertion became degraded efficiently (Fig. 5e). Interestingly, for constructs in which I27 insertions appear to have moved the initiation region past the optimal spacing degradation was not completely restored by longer tails. The reason for this observation could be that the position of the tails at the C-terminus of titin domains biases the random search of the tails in the wrong direction or that the pathway back to the proteasome receptor is blocked sterically. Substrates with a Ub<sub>4</sub> tag and shorter initiation regions (34 amino acids) showed a sharp drop in degradation rates when the initiation region was displaced from the tag by one inserted I27 domain and almost a complete block of degradation with three inserted domains. When these proteins contained a longer 102 amino acid initiation regions the decline of degradation rates was slower with longer spacings and even proteins with three or four inserted I27 domains were degraded (Fig. 5f).

These findings befit a model in which the proteasome recognizes the proteasome-binding tag and the initiation region with distinct binding sites at two separate locations on the protease particle. Interestingly, the two proteasome-binding tags bear different spacing requirements. It seems likely that only one recognition site for the initiation region of degradation exists (see discussion). Therefore, these results suggest that the locations of the recognition sites on the proteasome for the two tags differ.

## Discussion

Efficient degradation requires that the proteasome-binding tag and the initiation region bind to their recognition sites on the proteasome simultaneously. Thus, the arrangement or spacing of the tag and the initiation region on the substrate must match the arrangement of their recognition sites on the proteasome. When the unstructured region and ubiquitin tag are located either too close or too distant from each other, the proteasome does not degrade the substrate.

Even with the best spacing of the proteasome-binding tag and initiation region achievable in this set of experiments, effective degradation involved initiation regions exceeding 29 amino acids. This length requirement complements recent *in vitro* and *in vivo* studies that found efficient proteasomal degradation required unstructured regions between 20 to 30 amino acids<sup>2,3,25,32</sup>. The relatively small variations in length requirement between the studies may arise from sequence preferences of the initiation region receptor in the proteasome. For example, the last six residues of cytochrome *b*<sub>2</sub> initiation regions are histidines and it is possible that the proteasome will not recognize the hexahistidine stretch, thus reducing the effective length of these regions by six amino acids<sup>33</sup>. In addition, the distance between the two degron parts is increased in fairly large steps because of the size of the titin I27 domain

inserts<sup>28,29</sup> (about 44 Å N to C-terminus, Fig. 2c). The distance between the proteasome-binding tag and initiation region is possibly shorter than optimal in substrates without a titin insertion and longer than optimal in constructs with one titin insertion. If this is so, part of the unstructured region in the substrate must bridge the missing distance and is unavailable for proteasome binding. Finally, the size of the polyubiquitin chain and the participation of cellular factors in the degradation reaction may affect initiation region length requirements<sup>25</sup>.

The physical size of initiation regions of some 30 amino acids is roughly on the appropriate length scale to fit the proteasome structure. Currently, high-resolution structures of the 26S eukaryotic proteasome are not available. However, in bacteria, a series of AAA+ proteases fulfill the functions of the proteasome and structures of several of these proteases are available. These proteases are simpler in their composition than the proteasome, but their overall structure is analogous<sup>34,35</sup>. Substrates of the bacterial proteases are targeted for degradation by linear targeting sequences (*e.g.*, [36-39]), and these sequences are recognized by internal pore loops in the AAA+ subunit<sup>40-42</sup>. The Rpt subunits of the proteasome also contain AAA domains that form a ring structure so that it seems likely that they too interact with substrates through internal pore loops<sup>43</sup>. Consequently, these loops could serve as the receptor for the initiation region. The structures of the ATPase rings of the archaeal and actinobacterial proteasomes<sup>44</sup> suggest that the distance between the pore loops and the entrance to the Rpt ring is in the range of 60 - 70 Å. 24 to 34 amino acid long sequences in a random coil span some 50 to 70 Å [45], which is roughly within this range.

Our results suggest that the optimal spacing of the two parts of the degron is different for proteins with Ub<sub>4</sub> and UbL tags. Presumably, these distance requirements reflect the locations of the receptors for tag and initiation region on the proteasome: the two parts of the degron on the substrate must fit into the receptors on the proteasome. The Rpn10, Rpn13, and Rpt5 subunits of the yeast proteasome<sup>6-8</sup> recognize the ubiquitin chains and we speculate above that the initiation region is recognized by loops lining the channel at the center of a ring of the proteasome's ATPase subunits, called Rtps in yeast. Since substrates with ubiquitin tags are degraded efficiently when the initiation region is adjacent to the binding tag, at least one of the ubiquitin receptors should be located relatively close to the entrance of the degradation channel in the proteasome structure (Fig. 6a). If the initiation region in the substrate is too far from the ubiquitin moieties, the proteasome can no longer engage the degron effectively (Fig. 6b)

UbL domains are recognized by Rpn1, Rpn13 and human, but not yeast, Rpn10 subunits<sup>8,11-13</sup>. For substrates with a UbL tag, the degron was most effective when the initiation region and the proteasome-binding tag were separated by two I27 domains, which corresponds to approximately 70 - 80 Å [29]. Some conformations of the protein will either shorten or lengthen the distances, and only part of the unstructured region may be required for proteasome binding of the initiation region, further reducing the distance restriction. Nevertheless, in sharp contrast to degrons with a ubiquitin tag, UbL tag and initiation region must be separated by a substantial distance to function together effectively. Therefore, we predict that the UbL receptors on the proteasome will be located further away from the entrance to the degradation channel than the ubiquitin receptors (Fig. 6c).

The spacing requirement for UbL degrons is consistent with our understanding of the physiological function of the UbL domain. UbL domains are not attached directly to proteasome substrates post-translationally but found in targeting adaptors such as Rad23, Dsk2 and Ddi1<sup>9,10</sup>. These adaptors bind the proteasome through their UbL domain and substrates through UBA domains and seem to shuttle proteins for degradation. The adaptors themselves escape degradation and are recycled<sup>22,23,46</sup>. The reason for the escape is presumably that the potential degradation initiation regions on the adaptors are not effective, or at least not as effective as the initiation regions on the bound substrates, perhaps because initiation regions in the adaptors are too close to the UbL domain to engage the proteasome, while the bound proteins are better positioned for degradation (Fig. 6d).

Indeed, one role of the UbL-UBA adaptors could be to compensate for unfavorable spacing of ubiquitin degrons in proteasome substrates. The UPS handles a vast range of substrates. It is possible or even likely that in some of these proteins the ubiquitin modification and initiation regions are spaced suboptimally for rapid proteasome degradation. The UbL and UBA domains in Rad23, Dsk2 and Ddi1 are connected by unstructured loops of some 10 to 100 amino acids in length. These flexible linkers presumably allow the UbL and UBA domains to take up a wide range of different orientations relative to each other. Thus, the UbL-UBA adaptors could facilitate degradation by positioning substrates to allow their initiation region to feed into the degradation channel (Fig. 6d). A mechanism of this type has been described for the bacterial degradation adaptor SspB, which shuttles substrates to the AAA+ protease ClpXP<sup>47</sup>. Substrates bind to the head domain of SspB, which in turn is tethered to the protease through a C-terminal linker<sup>48</sup>. The flexibility of the linker allows SspB to deliver substrates bound in different geometries<sup>47,48</sup>.

Where substrates are recognized directly by the proteasome without adaptors, the spacing requirements for ubiquitin tags may also play a role in substrate selection. The proteasome can remodel protein complexes by degrading only specific subunits of a complex while leaving others intact (*e.g.*, [15-18]). The degraded subunit can be the one carrying the ubiquitin tag, or it can be an unmodified subunit<sup>14</sup>. Thus, if a multi-protein complex is brought to the proteasome, the proteasome chooses which subunit to degrade by selecting the best initiation region. Since the spacing of initiation regions affects initiation, it may also contribute to substrate selection.

*In summary*, we find that the two parts of degrons are recognized by receptors at separate locations on the proteasome. Effective degradation requires that the degron parts are positioned in an arrangements that fits their receptors on the proteasome. For ubiquitin tags, initiation region and proteasome binding tags must be adjacent to each other to function efficiently; in contrast, UbL domains have to be located at a certain distance from the initiation region to target for degradation. These steric requirements for degrons may contribute to substrate selection by the proteasome.



## Methods

### Protein constructs

Proteins were constructed from two different domains: the 27th Ig (I27) domain of the human giant muscle protein titin<sup>26</sup> and *E. coli* DHFR<sup>49</sup>. The I27 domain contained the mutation Cys47/63→Ala to remove potentially reactive sulfhydryl groups. Substrates with the mutant and wildtype I27 domains degraded with similar kinetics (Supplementary Fig. 5). The proteasome targeting signal was either a Ub<sub>4</sub> tag, which consisted of four ubiquitin moieties containing the mutation Gly76→Val connected by their N and C-termini through a six residue linker<sup>19</sup>, or a UbL tag, which consisted of residues 1-77 of *S. cerevisiae* Rad23<sup>22,23</sup>. Either tag was connected to the N-terminus of DHFR through a short linker. The initiation regions were attached to the C-termini of DHFR and were derived from *S. cerevisiae* cytochrome *b*<sub>2</sub> or *N. Crassa* subunit 9 of the F<sub>o</sub> ATP synthase. I27 domains were inserted between DHFR and initiation region and were connected to each other through two-residue linkers. Constructs are described in more detail in the Supplementary Methods.

For the FRET experiments, the sequence coding for eGFP<sup>30</sup> was fused to the N-terminus of DHFR through an eight residue linker, and the linker Cys-Cys-Gly-Pro-Cys-Cys was inserted into the 44 amino acid long cytochrome *b*<sub>2</sub> initiation region adjacent to its C-terminus.

### Protein expression and purification

For degradation experiments, substrates were cloned into the plasmid pGEM-3Zf(+) and expressed as radioactive proteins by *in vitro* transcription and translation supplemented with [<sup>35</sup>S]methionine. Substrates were synthesized using the *E. coli* T7 S30 Extract system (Promega Corp.) or the TNT coupled reticulocyte lysate system (Promega Corp.). Protein expressed using the two different systems behaved indistinguishably. After synthesis, the substrates were partially purified by high-speed centrifugation and ammonium sulfate precipitation.

For the FRET experiments, the substrate were cloned into the plasmid pET3a and expressed from a T7 promoter in *E. coli* strain Rosetta(DE3)pLysS (Novagen). Bacteria were grown at 37 °C, and expression was induced with 0.2 mM IPTG for 12-16 hrs at 22 °C after the culture reached an optical density at 600nm of 0.6. Proteins were purified using a talon metal affinity column (Clontech) equilibrated in buffer containing 50 mM sodium phosphate buffer [pH 7.0] and 300 mM NaCl. Proteins were eluted from the resin with buffer containing 50 mM sodium phosphate [pH 7.0] and 150 mM imidazole and further purified by MonoQ ion exchange chromatography (Pharmacia).

Yeast proteasome was purified from *S. cerevisiae* strain YYS40 (MATa *rpn11::RPN11<sup>3xFLAG</sup>-HIS3 leu2 his3 ura3 trp1 ade2 can1 ssd1*), in which the lid subunit Rpn11 was tagged with three FLAG tags at its C-terminus<sup>50</sup>. The 26S proteasome was purified following a modified version of a previously published protocol<sup>50</sup>. Harvested cells were resuspended in lysis buffer (50 mM Tris-HCl [pH 7.5], 10 mM MgCl<sub>2</sub>, 10% glycerol, 1 mM DTT, 4 mM ATP, 20 mM creatine phosphate, 0.2 mg/ml creatine phosphokinase) and lysed by a pressure homogenizer. After clarification of the lysate by centrifugation and

filtration, FLAG-tagged 26S proteasome was recovered by incubating the lysate with M2-agarose FLAG affinity beads (Sigma) for 2 hrs at 4 °C. The beads were washed with 50 bed volumes wash buffer (50 mM Tris-HCl [pH 7.5], 5 mM MgCl<sub>2</sub>, 10% glycerol, 1mM DTT, 2 mM ATP) and proteasome by eluted with wash buffer supplemented with 100 µg/ml 3xFlag peptide. The eluate was concentrated by ultrafiltration (Amicon) and stored at –80 °C in 15% glycerol.

### Proteasome degradation assays

Degradation assays were performed with 50 nM purified yeast proteasome<sup>50</sup> at 30 °C in 5% [v/v] glycerol, 5 mM MgCl<sub>2</sub>, 50 mM Tris/Cl [pH 7.4], 1 mM DTT, 1mg/ml BSA, 1 mM ATP, 10 mM creatine phosphate, 0.1 mg/ml creatine phosphokinase. Substrate protein was added to purified proteasome in reaction buffer to start the degradation. Samples were withdrawn at the indicated times, added to SDS-PAGE sample buffer, and analyzed by SDS-PAGE. The amounts of radioactive proteins were determined by electronic autoradiography (Instant Imager; Packard) or an imaging plate system (BAS-5000, Fuji).

Degradation experiments were repeated at least three times; data points are shown as the mean of independent measurements and error bars represent standard errors of the mean. Experiments were performed under first order single-turnover conditions (*i.e.*, with much smaller amounts of substrate than proteasome) so that degradation data were described by single exponential decays. The initial rate of these decays is the slope the decay at time zero and is given by the product of the amplitude and the rate constant of the decay. Initial degradation rates shown in graphs were usually calculated from the values of these parameters obtained by curve fitting. For cases where the change in the amount of substrate with time was small, the data were fitted to a straight line as an approximation of the initial phase of an exponential decay and the slope of the line was taken as the initial degradation rate. Data fitting was carried out using Kaleidagraph (version 4.0, Synergy Software). The error bars for degradation rates are standard errors obtained from the least square-fitting algorithm.

### FRET measurements

Purified protein at 10 µM in 50 mM Tris/Cl [pH 7.5], 1mM TCEP was mixed with 5-fold molar excess of ReAsH-EDT<sub>2</sub> in DMSO (Invitrogen). Samples were kept at room temperature for 30 min and then at 4 °C for 12-16 hours. Excess unreacted dye was removed by gel-filtration (NAP5 Sephadex G25 column; GE healthcare).

Steady-state fluorescence was measured with a ISS PC1 spectrofluorimeter in a solution comprised of 300 nM protein, 5% [v/v] glycerol, 5 mM MgCl<sub>2</sub>, 50 mM Tris/Cl [pH 7.4], at 25 °C. eGFP in the ReAsH-labeled protein was excited at 470 nm with a 4 nm bandwidth, and resulting fluorescence spectra between 475 nm and 680 nm (8 nm band width, 2 nm steps) were measured. FRET efficiencies (*E*) were calculated from the FRET-induced quenching of eGFP emission at 507 nm (8 nm band width):  $E = 1 - I_{DA}/I_D$ , where  $I_{DA}$  and  $I_D$  are donor (eGFP) intensities at 507 nm before and after removal of the fluorescence acceptor (ReAsH), respectively. Acceptor was removed by digestion of the protein linker by 30 µg proteinase K added to 500 µL of sample in the optical cell.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

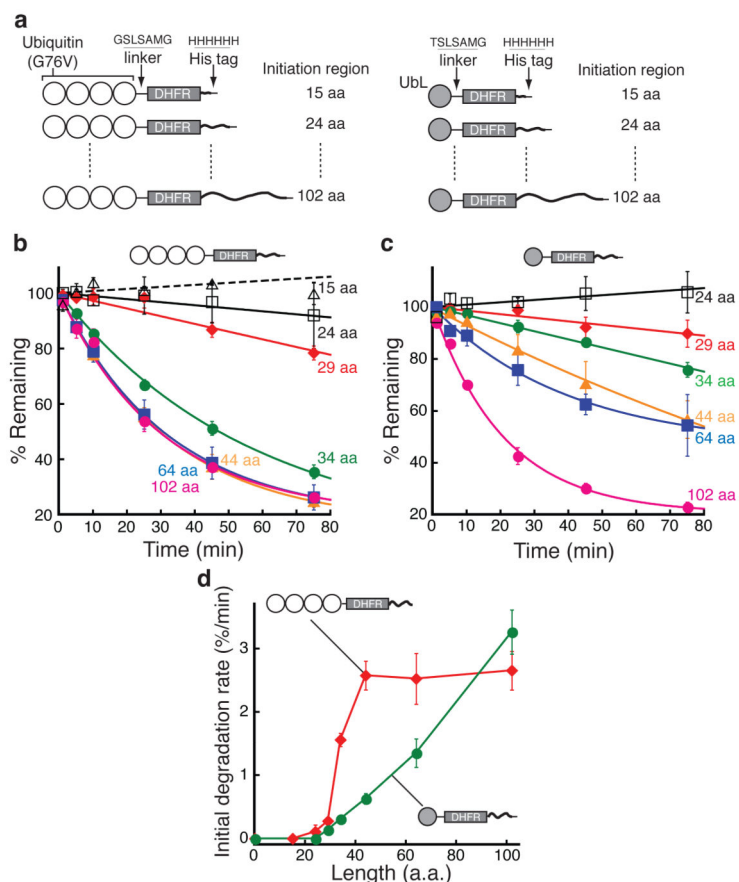
We thank Drs. S. Elsassner (Harvard University Medical School), D. Finley (Harvard University Medical School), B.S. Glick (University of Chicago) and Y. Saeki (Tokyo Metropolitan Institute of Medical Science) for providing plasmids and yeast strains and members of the Matouschek lab for advice and comments. The authors greatly appreciate and gratefully acknowledge Dr. Nobuyuki Nukina (RIKEN) for his support and the use of his laboratory equipment. We also thank Ginger Leigh for editing the manuscript. The work was supported by grants R01GM64003 and U54CA143869 from the US National Institutes of Health, by Ministry of Education, Culture, Sports, Science, and Technology of Japan Grant-in-Aid 22770137 (T.I.) and by the Robert H. Lurie Comprehensive Cancer Center at Northwestern University. T.I. also gratefully acknowledges a Japan Society for the Promotion of Science Postdoctoral Fellowship for Research Abroad.

## References

1. Pickart CM. Back to the future with ubiquitin. *Cell*. 2004; 116:181–190. [PubMed: 14744430]
2. Prakash S, Tian L, Ratliff KS, Lehotzky RE, Matouschek A. An unstructured initiation site is required for efficient proteasome-mediated degradation. *Nat. Struct. Mol. Biol.* 2004; 11:830–837. [PubMed: 15311270]
3. Takeuchi J, Chen H, Coffino P. Proteasome substrate degradation requires association plus extended peptide. *EMBO J.* 2007; 26:123–131. [PubMed: 17170706]
4. Schrader EK, Harstad KG, Matouschek A. Targeting proteins for degradation. *Nat. Chem. Biol.* 2009; 5:815–822. [PubMed: 19841631]
5. Thrower JS, Hoffman L, Rechsteiner M, Pickart CM. Recognition of the polyubiquitin proteolytic signal. *EMBO J.* 2000; 19:94–102. [PubMed: 10619848]
6. Deveraux Q, Ustrell V, Pickart C, Rechsteiner M. A 26 S protease subunit that binds ubiquitin conjugates. *J. Biol. Chem.* 1994; 269:7059–7061. [PubMed: 8125911]
7. Lam YA, Lawson TG, Velayutham M, Zweier JL, Pickart CM. A proteasomal ATPase subunit recognizes the polyubiquitin degradation signal. *Nature*. 2002; 416:763–767. [PubMed: 11961560]
8. Husnjak K, et al. Proteasome subunit Rpn13 is a novel ubiquitin receptor. *Nature*. 2008; 453:481–488. [PubMed: 18497817]
9. Elsassner S, Finley D. Delivery of ubiquitinated substrates to protein-unfolding machines. *Nat. Cell Biol.* 2005; 7:742–749. [PubMed: 16056265]
10. Finley D. Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu. Rev. Biochem.* 2009; 78:477–513. [PubMed: 19489727]
11. Hiyama H, et al. Interaction of hHR23 with S5a. The ubiquitin-like domain of hHR23 mediates interaction with S5a subunit of 26 S proteasome. *J. Biol. Chem.* 1999; 274:28019–28025. [PubMed: 10488153]
12. Elsassner S, et al. Proteasome subunit Rpn1 binds ubiquitin-like protein domains. *Nat. Cell Biol.* 2002; 4:725–730. [PubMed: 12198498]
13. Saeki Y, Sone T, Toh-e A, Yokosawa H. Identification of ubiquitin-like protein-binding subunits of the 26S proteasome. *Biochem. Biophys. Res. Commun.* 2002; 296:813–819. [PubMed: 12200120]
14. Prakash S, Inobe T, Hatch AJ, Matouschek A. Substrate selection by the proteasome during degradation of protein complexes. *Nat. Chem. Biol.* 2009; 5:29–36. [PubMed: 19029916]
15. Johnson ES, Gonda DK, Varshavsky A. cis-trans recognition and subunit-specific degradation of short-lived proteins. *Nature*. 1990; 346:287–291. [PubMed: 2165217]
16. Hochstrasser M, Varshavsky A. In vivo degradation of a transcriptional regulator: the yeast alpha 2 repressor. *Cell*. 1990; 61:697–708. [PubMed: 2111732]

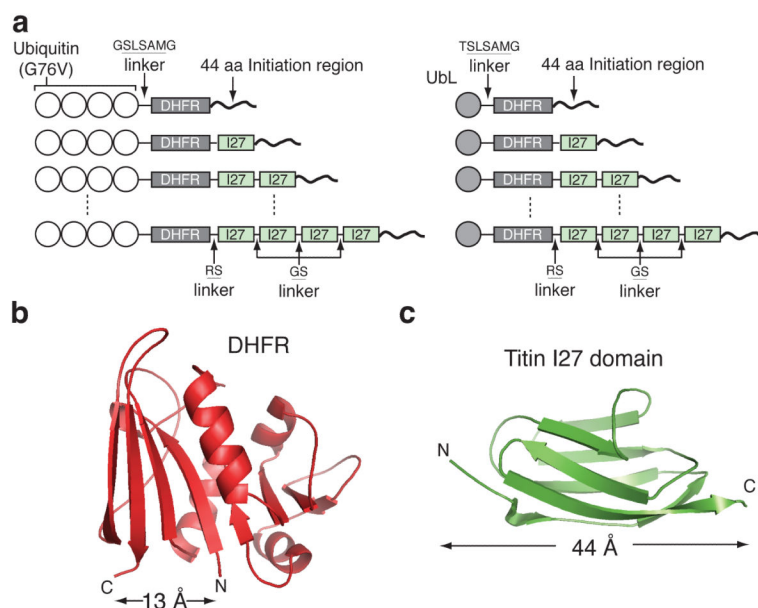
17. Klotzbücher A, Stewart E, Harrison D, Hunt T. The 'destruction box' of cyclin A allows B-type cyclins to be ubiquitinated, but not efficiently destroyed. *EMBO J.* 1996; 15:3053–3064. [PubMed: 8670806]
18. Verma R, McDonald H, Yates JR, Deshaies RJ. Selective degradation of ubiquitinated Sic1 by purified 26S proteasome yields active S phase cyclin-Cdk. *Mol. Cell.* 2001; 8:439–448. [PubMed: 11545745]
19. Stack JH, Whitney M, Rodems SM, Pollok BA. A ubiquitin-based tagging system for controlled modulation of protein stability. *Nat. Biotechnol.* 2000; 18:1298–1302. [PubMed: 11101811]
20. Komander D, et al. Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. *EMBO Rep.* 2009; 10:466–473. [PubMed: 19373254]
21. Saeki Y, et al. Lysine 63-linked polyubiquitin chain may serve as a targeting signal for the 26S proteasome. *EMBO J.* 2009; 28:359–371. [PubMed: 19153599]
22. Watkins JF, Sung P, Prakash L, Prakash S. The *Saccharomyces cerevisiae* DNA repair gene RAD23 encodes a nuclear protein containing a ubiquitin-like domain required for biological function. *Mol. Cell. Biol.* 1993; 13:7757–7765. [PubMed: 8246991]
23. Schaubert C, et al. Rad23 links DNA repair to the ubiquitin/proteasome pathway. *Nature.* 1998; 391:715–718. [PubMed: 9490418]
24. Johnston JA, Johnson ES, Waller PR, Varshavsky A. Methotrexate inhibits proteolysis of dihydrofolate reductase by the N-end rule pathway. *J. Biol. Chem.* 1995; 270:8172–8178. [PubMed: 7713922]
25. Verhoef L, et al. Minimal length requirement for proteasomal degradation of ubiquitin-dependent substrates. *FASEB J.* 2008; 23:123–133. [PubMed: 18796559]
26. Politou AS, Gautel M, Pfuhl M, Labeit S, Pastore A. Immunoglobulin-type domains of titin: same fold, different stability? *Biochemistry.* 1994; 33:4730–4737. [PubMed: 8161531]
27. Rief M, Gautel M, Oesterhelt F, Fernandez JM, Gaub HE. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science.* 1997; 276:1109–1112. [PubMed: 9148804]
28. Improta S, Politou AS, Pastore A. Immunoglobulin-like modules from titin I-band: extensible components of muscle elasticity. *Structure.* 1996; 4:323–337. [PubMed: 8805538]
29. von Castelmur E, et al. A regular pattern of Ig super-motifs defines segmental flexibility as the elastic mechanism of the titin chain. *Proc. Natl. Acad. Sci. USA.* 2008; 105:1186–1191. [PubMed: 18212128]
30. Yang TT, Cheng L, Kain SR. Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. *Nucleic Acids Res.* 1996; 24:4592–4593. [PubMed: 8948654]
31. Adams SR, et al. New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. *J. Am. Chem. Soc.* 2002; 124:6063–6076. [PubMed: 12022841]
32. Beskow A, et al. A conserved unfoldase activity for the p97 AAA-ATPase in proteasomal degradation. *J. Mol. Biol.* 2009; 394:732–746. [PubMed: 19782090]
33. Tian L, Holmgren RA, Matouschek A. A conserved processing mechanism regulates the activity of transcription factors Cubitus interruptus and NF- $\kappa$ B. *Nat. Struct. Mol. Biol.* 2005; 12:1045–1053. [PubMed: 16299518]
34. Larsen CN, Finley D. Protein translocation channels in the proteasome and other proteases. *Cell.* 1997; 91:431–434. [PubMed: 9390550]
35. Baumeister W, Walz J, Zühl F, Seemüller E. The proteasome: paradigm of a self-compartmentalizing protease. *Cell.* 1998; 92:367–380. [PubMed: 9476896]
36. Keiler KC, Waller PR, Sauer RT. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science.* 1996; 271:990–993. [PubMed: 8584937]
37. Gottesman S, Roche E, Zhou Y, Sauer RT. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev.* 1998; 12:1338–1347. [PubMed: 9573050]
38. Flynn JM, Neher SB, Kim YI, Sauer RT, Baker TA. Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol. Cell.* 2003; 11:671–683. [PubMed: 12667450]

39. Gonzalez M, Frank EG, Levine AS, Woodgate R. Lon-mediated proteolysis of the Escherichia coli UmuD mutagenesis protein: in vitro degradation and identification of residues required for proteolysis. *Genes Dev.* 1998; 12:3889–3899. [PubMed: 9869642]
40. Yamada-Inagawa T, Okuno T, Karata K, Yamanaka K, Ogura T. Conserved pore residues in the AAA protease FtsH are important for proteolysis and its coupling to ATP hydrolysis. *J. Biol. Chem.* 2003; 278:50182–50187. [PubMed: 14514680]
41. Hinnerwisch J, Fenton WA, Furtak KJ, Farr GW, Horwich AL. Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. *Cell.* 2005; 121:1029–1041. [PubMed: 15989953]
42. Martin A, Baker TA, Sauer RT. Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. *Mol. Cell.* 2008; 29:441–450. [PubMed: 18313382]
43. Zhang F, et al. Mechanism of substrate unfolding and translocation by the regulatory particle of the proteasome from *Methanocaldococcus jannaschii*. *Mol. Cell.* 2009; 34:485–496. [PubMed: 19481528]
44. Djuranovic S, et al. Structure and activity of the N-terminal substrate recognition domains in proteasomal ATPases. *Mol. Cell.* 2009; 34:580–590. [PubMed: 19481487]
45. Miller WG, Goebel CV. Dimensions of protein random coils. *Biochemistry.* 1968; 7:3925–3935. [PubMed: 5722263]
46. Heessen S, Masucci MG, Dantuma NP. The UBA2 domain functions as an intrinsic stabilization signal that protects Rad23 from proteasomal degradation. *Mol. Cell.* 2005; 18:225–235. [PubMed: 15837425]
47. Levchenko I, Grant RA, Flynn JM, Sauer RT, Baker TA. Versatile modes of peptide recognition by the AAA+ adaptor protein SspB. *Nat. Struct. Mol. Biol.* 2005; 12:520–525. [PubMed: 15880122]
48. McGinness KE, Bolon DN, Kaganovich M, Baker TA, Sauer RT. Altered tethering of the SspB adaptor to the ClpXP protease causes changes in substrate delivery. *J. Biol. Chem.* 2007; 282:11465–11473. [PubMed: 17317664]
49. Rood JI, Laird AJ, Williams JW. Cloning of the Escherichia coli K-12 dihydrofolate reductase gene following mu-mediated transposition. *Gene.* 1980; 8:255–265. [PubMed: 6444603]
50. Saeki Y, Isono E, Toh-E A. Preparation of ubiquitinated substrates by the PY motif-insertion method for monitoring 26S proteasome activity. *Meth. Enzymol.* 2005; 399:215–227. [PubMed: 16338358]

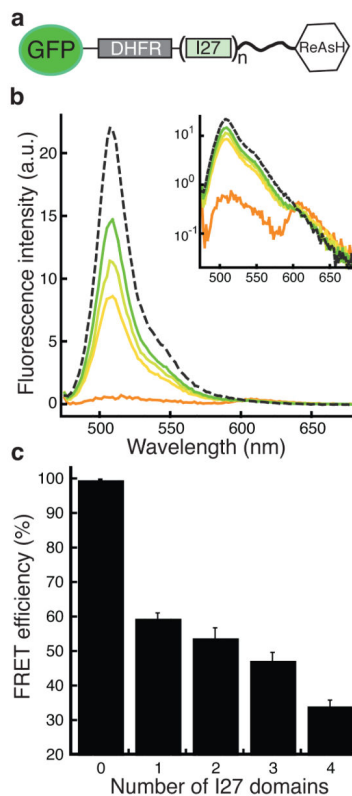


**Figure 1.**

Proteasome-mediated degradation depends on initiation region length. **(a)** Linear representation of substrate proteins with initiation regions of different lengths. The substrates contained an *E. coli* DHFR domain and were targeted to the proteasome by an N-terminal Ub<sub>4</sub> tag (left) or a UbL tag (right). Unstructured tails of different lengths were placed at the C-terminus of DHFR to serve as initiation regions. Degradation kinetics of substrates with ubiquitin **(b)** or UbL **(c)** tags and initiation regions of 15 (open triangles), 24 (black open squares), 29 (red solid diamonds), 34 (green solid circles), 44 (orange solid triangles), 64 (blue solid squares), and 102 (purple solid circles) amino acids derived from *S. cerevisiae* cytochrome *b*<sub>2</sub>. The extent of degradation was plotted as the percentage of protein remaining at different times. **(d)** Dependence of degradation rates (initial rates) on initiation region length for proteins with Ub<sub>4</sub> (red solid diamonds) and UbL (green solid circles) tags. Data points represent mean values and error bars show standard errors calculated from three to five repeat experiments.



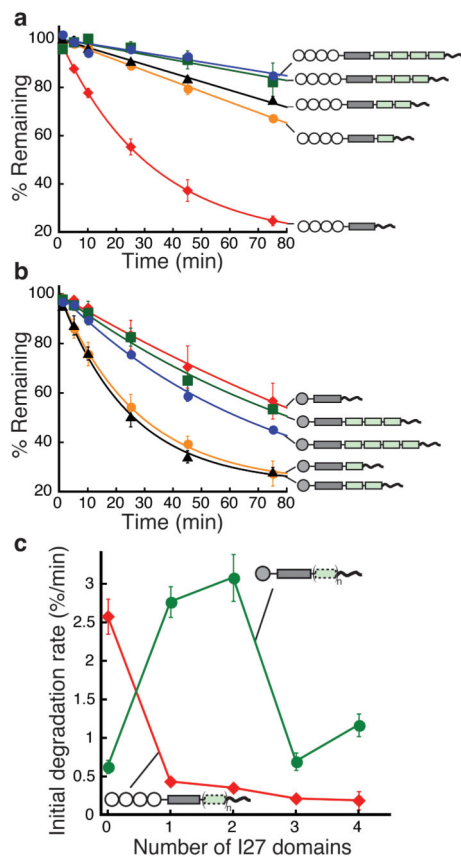
**Figure 2.** Separating the proteasome-binding tag and initiation region in degradation substrates. **(a)** Linear representation of proteasome substrates in which the proteasome-binding tag and initiation region are separated by the insertion of titin immunoglobulin I27 domains. The proteins contained a 44 amino acid long initiation region derived from cytochrome  $b_2$  at their C-termini and were targeted to the proteasome by either a Ub<sub>4</sub> or a UbL tag at their N-termini. Structures of *E. coli* DHFR [PDB ID 1DRH] **(b)** and the immunoglobulin domain I27 of titin [PDB ID 1TIT] **(c)** shown in cartoon representation using PyMol (DeLano Scientific LLC, Palo Alto, CA; [www.pymol.org](http://www.pymol.org)).



**Figure 3.**

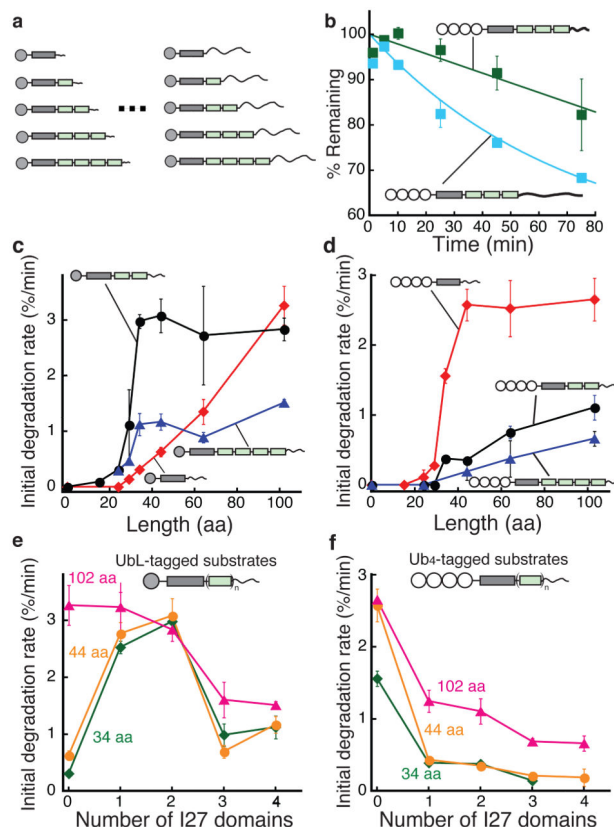
Immunoglobulin domain I27 insertions increase the distance between proteasome-binding tag and the initiation region. In these experiments the tag at the N-terminus of DHFR was replaced by eGFP and the ReAsH-binding motif (CCGPCC) was fused to the C-terminus of the initiation region as shown in a schematic representation of these proteins (a). (b) Emission spectra of FRET substrates with no (orange), one (yellow), two (light green), or four (green) I27 domain insertions excited at 470 nm (4 nm bandwidth). After digestion with proteinase K, no FRET was observed (dashed line). Inset shows emission spectra with intensities shown on a log scale. (c) FRET efficiency for eGFP and ReAsH-labeled proteins decreased with the number of I27 domains inserted. Data shown by bar graph represent mean values and error bars show standard errors calculated from three repeat experiments.





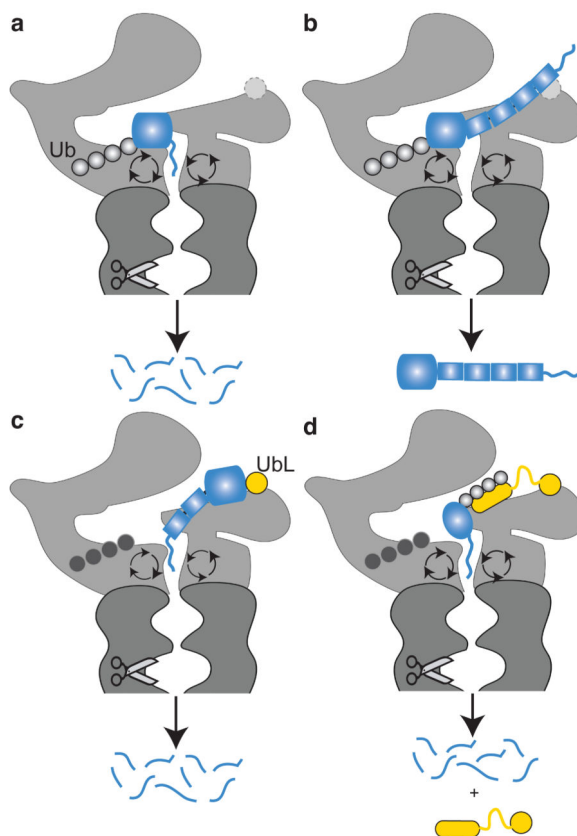
**Figure 4.**

Proteasome-mediated degradation depends on the spacing between proteasome-binding tag and initiation region. **(a, b)** Degradation kinetics for substrates with no (red solid diamonds), one (orange solid circles), two (black solid triangles), three (green solid squares), and four (blue solid square) I27 domain insertions targeted to the proteasome by either a Ub<sub>4</sub> **(a)** or a UbL tag **(b)**. **(c)** Plots of initial degradation rates as a function of the separation of the proteasome-binding tag and initiation region induced by the insertion of I27 domains show different relationships for substrates with Ub<sub>4</sub> (red solid diamonds) and UbL tags (green solid circles). For proteins with ubiquitin tags, degradation rates were highest at short separations and then decreased at larger separations; for proteins with UbL tags, degradation was slowest at short separations and then accelerated with greater separation before decreasing again at the greatest separations. Data points represent mean values and error bars show standard errors calculated from three to five repeat experiments.



**Figure 5.**

Spacing requirements for the two degron parts as initiation region length varies. To determine how the spacing of the proteasome-binding tag and initiation region modulates the effect of initiation region length on degradation, we varied proteasome-binding tag - initiation region spacing and initiation region length systematically as shown for constructs with UbL tags (**a**), Ub<sub>4</sub>-tagged substrates were constructed analogously. (**b**) Degradation kinetics for substrates in which three I27 domains separate a Ub<sub>4</sub> tag from a 44 amino acid (green squares) or a 102 amino acid (light blue squares) initiation region. (**c,d**) Results for the degradation of other proteins in the array of constructs are shown as plots of initial degradation rates against initiation region lengths for constructs with UbL (**c**) and Ub<sub>4</sub> (**d**) tags and zero (red solid diamonds), two (black solid circles), or four (blue solid triangles) I27 domain insertions. Degradation rates increase with the length of the initiation region for both tags but the strength of the response depends on the spacing. (**e,f**) Plots of initial degradation rates as a function of the distance between the proteasome-binding tag and initiation region for substrates with UbL (**e**) and Ub<sub>4</sub> (**f**) tags with 34 amino acid (green solid diamonds), 44 amino acid (orange solid circles), or 102 amino acid (purple solid triangles) initiation regions. The longest initiation regions dampen the effect of degron spacing on degradation rates. Data points represent mean values and error bars show standard errors calculated from three to five repeat experiments.



**Figure 6.** Schematic representation of the length and spacing requirement for the initiation region. Proteasome caps contain ubiquitin and UbL receptors and these occupy different locations in the caps as indicated schematically, however, the placement of these locations in the sketch is arbitrary. Shapes and colors indicated different proteins. Light grey: proteasome caps; dark grey: proteasome core; circular arrows: ATPase subunits; scissors: proteolytic sites; grey circles: ubiquitin; yellow circle: UbL domain; blue shapes: substrates; yellow shapes: UbL-UBA adaptor protein; grey circles: ubiquitin binding sites; grey dotted circles: UbL binding sites. **(a)** Substrates in which both degron parts are close to each other are degraded efficiently when the receptors for ubiquitin tag and initiation region are close to each other on the proteasome. **(b)** When the ubiquitin tag and initiation region are farther apart in the substrate, the proteasome can no longer engage both degron parts at the same time and degradation is inefficient. **(c)** The spacing requirements for degrons with UbL tags may be explained if the UbL receptor on the proteasome is located at some distance from the initiation region receptor. **(d)** The figure shows a proteasome substrate bound to a UbL-UBA adaptor. In this representation the adaptor fits between the UbL and initiation site receptors on the proteasome. This spatial arrangement positions the substrate for effective degradation and keeps the unstructured regions on the adaptor away from the initiation region receptor.