ACS OMEGA

http://pubs.acs.org/journal/acsodf

Mini-Review

# Toward Understanding the Biochemical Determinants of Protein Degradation Rates

Miguel Correa Marrero* and Inigo Barrio-Hernandez*

Cite This: *ACS Omega* 2021, 6, 5091−5100

Read Online

ACCESS | Metrics & More | Article Recommendations

**ABSTRACT:** Protein degradation is a key component of the regulation of gene expression and is at the center of several pathogenic processes. Proteins are regularly degraded, but there is large variation in their lifetimes, and the kinetics of protein degradation are not well understood. Many different factors can influence protein degradation rates, painting a highly complex picture. This has been partially unravelled in recent years thanks to invaluable advances in proteomics techniques. In this Mini-Review, we give a global vision of the determinants of protein degradation rates with the backdrop of the current understanding of proteolytic systems to give a contemporary view of the field.

## INTRODUCTION

The protein content of a cell is an important component of its identity and the functions it performs. The balance between protein synthesis and degradation controls protein abundance. This continuous turnover is necessary for cells to adapt their proteomes to internal and external perturbations and maintain an adequate amino acid pool. Furthermore, protein degradation, together with chaperone-mediated protein folding, is one of the key mechanisms to maintain quality control of proteins.[1] Thus, protein degradation is a key component of cellular homeostasis. This is shown by the fact that it impacts virtually all cellular processes, including the cell cycle, circadian rhythms, and development. Additionally, anomalies in protein degradation have been linked to impaired development,[2] cancer, and a whole set of disorders, called proteinopathies, which include many neurodegenerative diseases.[3] The activities of proteolytic systems decline with age, suggesting that loss of proteostasis is an important feature of aging.[4] Hence, understanding protein degradation will help us understand what drives the need to recycle proteins and to clarify the molecular mechanisms involved in multiple diseases, as well as aging. This will also have an impact on protein production technologies.

Proteins are regularly degraded by an elaborate machinery that needs to be highly promiscuous in order to degrade the large variety of cellular proteins. Yet, protein lifetimes vary greatly, ranging from minutes to days, which indicates that this process also has a strongly selective component. This raises a long-standing, but understudied, question in protein biochemistry: what intrinsic features determine protein degradation rate in the cell? This question has been studied since the 1970s, although early studies were severely hampered by their limited sample size. Advances in proteomics technologies have allowed us to carry out large-scale studies to revisit this question, challenge old hypotheses and test new ones. Pulsed SILAC-based methods, like the one described in ref 5, rely on the incorporation of nonradioactive isotope-labeled amino acids to cell culture media for a short period of time, allowing the measurement of degradation rates for thousands of proteins. Furthermore, our knowledge of the fraction of the ubiquitome involved in the ubiquitin-proteasome system is improving thanks to novel antibodies to detect ubiquitinated peptides, as well as inhibitors of proteasome function such as b-AP15 DUB and bortezomib.[6] Combined with downstream bioinformatics analyses, these advances have uncovered the importance of different intrinsic characteristics, such as the presence of long disordered segments or protein topology, but also of protein−protein interactions. In this Mini-Review, we summarize recent insights into proteome-wide determinants of protein degradation in eukaryotes in the context of our knowledge of the workings of proteolytic systems, as well as provide suggestions for the direction of future studies.

## PROTEIN DEGRADATION PATHWAYS

Intracellular protein degradation is mainly controlled by two pathways: the ubiquitin-proteasome system (UPS) and the autophagy-lysosome system (autophagy). The relative weight of each in total protein degradation is known to change depending on the cellular context, but in physiological conditions, the UPS predominates and carries out ∼80% of total degradation.[7] Here, we briefly review them to put the determinants of protein degradation in the context of these proteolytic pathways (Figure 1).

**Ubiquitin-Proteasome System.** The 26S proteasome, a large cylindrical complex with proteolytic activity located in the cytosol and nucleus of eukaryotic cells, is at the core of the UPS. It consists of two subcomponents: a 20S core particle with proteolytic activity and a 19S regulatory particle. The 19S

Received: October 31, 2020
Accepted: February 5, 2021
Published: February 17, 2021





https://dx.doi.org/10.1021/acsomega.0c05318
*ACS Omega* 2021, 6, 5091−5100

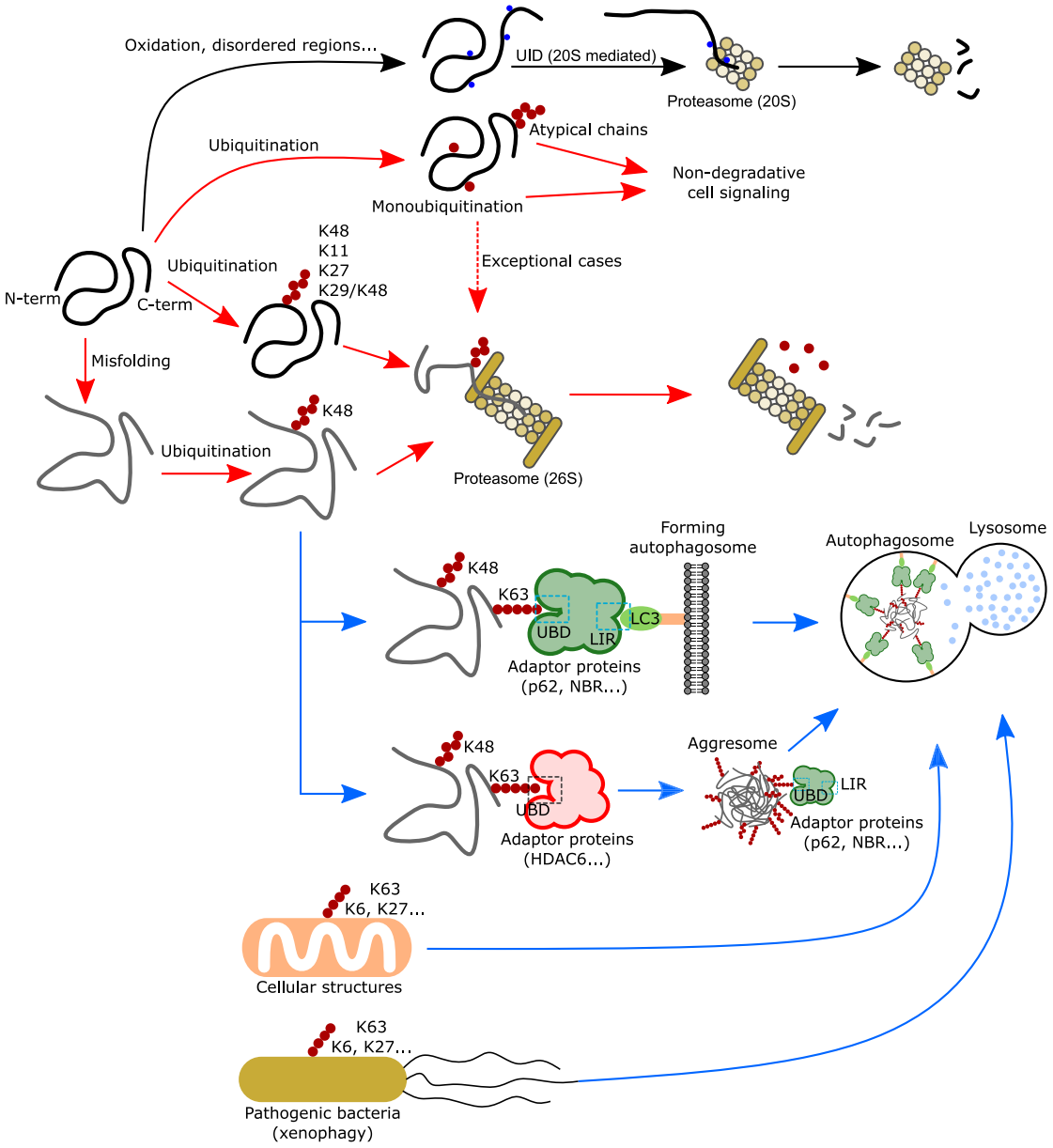


**Figure 1.** Role of ubiquitin in degradation pathways (black arrows for UID, red arrows for UPS, blue arrows for autophagy). Oxidized proteins are mainly degraded by UID in the 20S proteasome, a process independent from both ubiquitin and ATP. Intrinsically disordered proteins as well as proteins with disordered regions can also be degraded by UID. Ubiquitination can have different outcomes depending on the architecture of the chains. Monoubiquitination and polyubiquitination with atypical chains (K27, K29, K33....) are usually involved in nondegradative processes. K48 polyubiquitination is the canonical signal for protein degradation mediated by UPS. Nevertheless, proteins tagged with other chains (K11, K27, K29....), as well as monoubiquitinated proteins and even proteins without ubiquitin, can also be proteasome targets. A misfolded protein will be K48 polyubiquitinated and degraded via UPS; if this is not possible, autophagy is activated. K63 polyubiquitination is able to attract adaptor proteins like p62 or NBR with UBD domains, as well as LIR domains which mediate the interaction with the forming autophagosome protein LC3. When autophagosomes cannot be formed, other groups of adaptors (e.g., HDAC6) are able to accumulate misfolded proteins in aggresomes to be degraded when possible. Autophagy is able to degrade large intracellular structures such as obsolete or damaged organelles (e.g., mitophagy) as well as pathogens (xenophagy).

proteasome recognizes proteins tagged with ubiquitin. After recognition, a second step where the proteasome engages the substrate at an initiation site is necessary to begin degradation. Initiation sites consist of sufficiently long disordered regions. Subsequently, the 19S proteasome deubiquitinates and unfolds the protein, a process needed for the translocation to the 20S proteasome and dependent on energy provided by the AAA+ ATPase motor.[8] The recognition of ubiquitin tags, the presence of efficient initiation regions, and protein unfolding and translocation are all important components of the degradation rate.[2,9−12]

The 20S proteasome contains the proteolytic active sites. The active sites exhibit triple activity to ensure protein digestion: tryptic sites targeting the C-terminal end of basic residues, chymotryptic sites able to cleave after hydrophobic residues, and caspase-like sites that cleave after acidic residues.[13] Finally, the resulting short peptides are released from the proteasome and are further processed by associated proteases. In some cases, the products of proteasomal degradation have biological functions.


https://dx.doi.org/10.1021/acsomega.0c05318
ACS Omega 2021, 6, 5091−5100

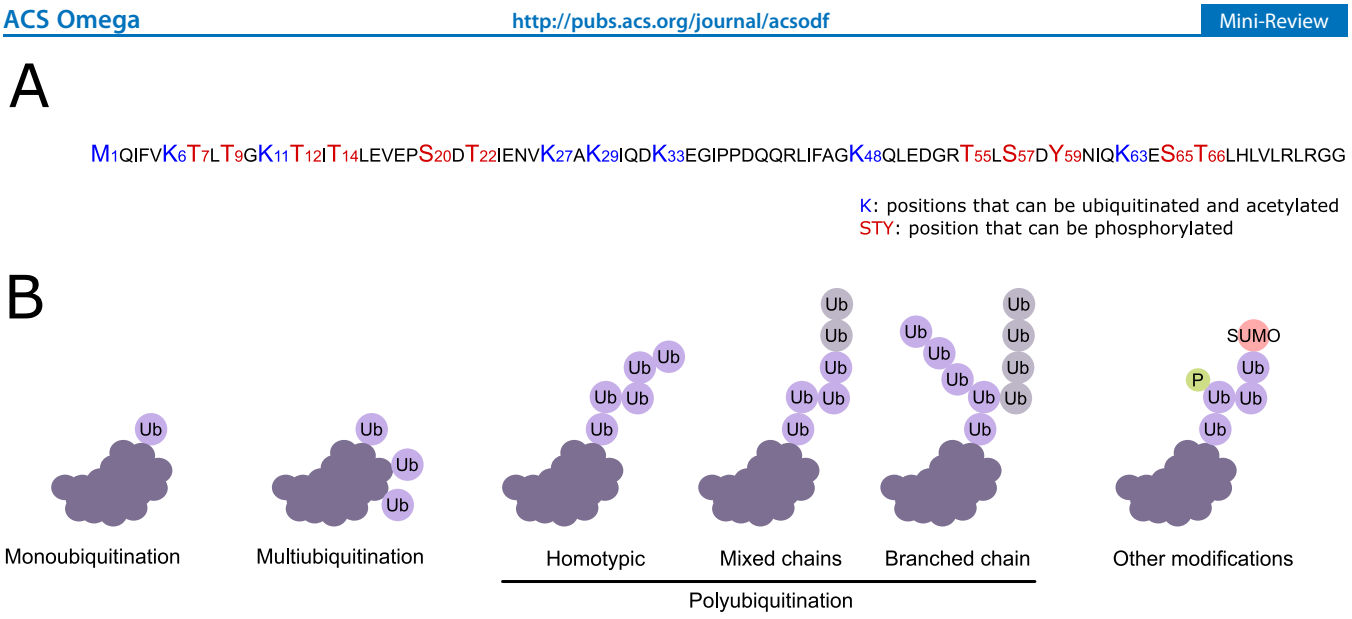


**Figure 2.** The ubiquitin code. (A) Ubiquitin sequence. In blue, residues susceptible to ubiquitination and acetylation. In red, residues susceptible to phosphorylation (based on mass spectrometry experiments). (B) Ubiquitination code, a protein can be ubiquitinated in one position (monoubiquitination) or several positions (multiubiquitination). Ubiquitin can be ubiquitinated in seven internal lysines plus the N terminus (polyubiquitination), forming homotypic (always the same lysine), mixed (different lysines in the same sequence), or branched (one ubiquitin is ubiquitinated in at least two positions at the same time) chains. Ubiquitin is also susceptible to other PTMs, such as SUMOylation, phosphorylation, or acetylation.

**Autophagy.** Targeted protein degradation mediated by autophagy in normal physiological conditions is well studied in the context of protein quality control mechanisms. If misfolded proteins tagged for degradation manage to escape the UPS, they are degraded by autophagy.[14] This is a bulk degradation system that requires the formation of an autophagosome compartment and subsequent fusion with a lysosome, which contains the required proteolytic enzymes.[15]

Autophagy to degrade misfolded proteins can be triggered in two fashions. In the first case, tagged proteins are recognized by adaptor proteins such as p62 or NBR1, which bind through their LIC domains to LC3, located in the membrane of forming unmatured autophagosomes (phagophores). In the second case, misfolded proteins that cannot be degraded directly are stored in aggresomes. HDAC6 is responsible for detecting and regulating this controlled aggregation. Eventually, aggregates are recognized by adaptors such as p62 and are linked to phagophores in an LC3-dependent way.[15]

Autophagy has other important functions in the cell worth mentioning, such as degradation of obsolete or malfunctioning large intracellular structures (e.g., organelles). Its role in degrading foreign bodies such as bacteria that have invaded the cell, referred to as xenophagy, is especially interesting. Other stress situations, such as starvation, can also trigger autophagy, in this case to supply the cell with nutrients.

## DETERMINANTS OF PROTEIN DEGRADATION

**Post-translational Modifications.** Post-translational modifications (PTMs) are covalent and often reversible modifications of proteins that are able to change and regulate their function. Virtually all biological processes are at some point regulated by PTMs, and protein degradation is no exception. Here, we cover ubiquitination and phosphorylation, the best-studied PTMs in the context of protein degradation.

*Ubiquitination.* Ubiquitination is the quintessential PTM for protein degradation. This modification consists in the addition of ubiquitin, a small 76 amino acids long protein, to amino groups in lysine side chains, as well as the $\alpha$-amino group in a protein's N-terminal end. Ubiquitination depends on the activity of three enzymes: E1 activating enzymes, which prime ubiquitin; E2 conjugating enzymes, which covalently attach ubiquitin; and E3 ligases enzyme, responsible for target recognition. Since E3 enzymes recognize ubiquitination targets, they are also the most diverse, with more than 600 genes coding for them in humans.[16] The diversity of E3 enzymes highlights the complexity and importance of this process. A minimal portion of a protein recognized by the UPS-mediated degradation machinery can be defined as a degron. This can refer to both sequence and structural features. The reverse reaction is mediated by deubiquitinase enzymes (DUBs), which modulate and attenuate the signal. It should be noted, however, that protein ubiquitination is not only involved in degradation. It also participates in nondegradative signaling, from intracellular trafficking to transcriptional regulation.[6] Here, we focus on how protein ubiquitination affects intracellular protein degradation, mediated by the UPS and autophagy.

Proteins can be ubiquitinated in one position (monoubiquitination) or in several positions (multiubiquitination). Furthermore, ubiquitin itself has seven lysines (K6, K11, K27, K29, K33, K48, and K63), as well as an N-terminal end susceptible to ubiquitination (Met1). This allows the creation of chains with different configurations (polyubiquitination), depending on the modified residue. On top of this, ubiquitin can be modified with other PTMs, including phosphorylation in 11 different positions, acetylation of lysines, and addition of small proteins such as Nedd8 or SUMO (Figure 2A). All this diversity allows a plethora of architectures which can be specifically recognized and distinguished by ubiquitin binding domains (UBDs), leading to a so-called ubiquitination code (Figure 2B).[17] This confers specificity toward the different processes that ubiquitination participates in.

The canonical signal for proteasomal degradation is K48 polyubiquitination, which is recognized by UBD-containing accessory proteins in the regulatory portion of the proteasome,


https://dx.doi.org/10.1021/acsomega.0c05318
ACS Omega 2021, 6, 5091−5100

such as RPN10 and RPN13.[14] The atypical chain K11 has also been described as a proteasomal degradation signal, either alone or as a branched and mixed chain together with K48 and K63 linkages. Another atypical chain linked to the UPS is K29, described in the ubiquitin fusion degradation (UFD) pathway in yeast. The signal is a mixed branched chain: a main K29 polyubiquitin chain, multimonoubiquitinated with K48 linkage, which requires two E3 ligases, Ufd4p for the initial K29 chain and Ufd2p for K48 multiubiquitination. Finally, monoubiquitination, which is usually involved in nondegradative signaling, has also been described in specific contexts as an UPS trigger.[18] As discussed above, the kinetics of protein degradation mediated by the proteasome are influenced by the ubiquitin-dependent target recognition, as well as postbinding effects, which include protein translocation through the channel. In this same context, the architecture of the chain as a degradation signal, as well as the number of ubiquitinated positions and the length of the chain, also affect the dynamics of protein degradation. It has been suggested that the minimum length of the chain needed to trigger degradation is 4 units.[19] Further evidence was provided by a high-resolution structure of the regulatory region of the proteasome, which showed a pocket able to fit a tetraubiquitin K48 chain.[8] Nevertheless, an elegant single molecule assay performed by Lu et al.[9] suggested that shorter chains (with two ubiquitin units) can be even more effective and that multiubiquitination in several sites is able to enhance protein degradation. This could be in agreement with the finding of extensive ubiquitinated proteins in a deep ubiquitinome study.[6] The authors found a subset of proteins with high density of ubiquitination sites and discussed a possible function as interaction platforms mediated by UBDs, as well as a way of enhancing protein degradation rate.

Although ubiquitination is a crucial component of the UPS, proteins can also be degraded in the proteasome without being ubiquitinated and without ATP expense. This alternative pathway is called ubiquitin-independent proteasomal degradation (UID), also known as degradation by default, and it is mainly, but not exclusively, mediated by the 20S proteasome (without the 19S regulatory portion) (Figure 1). Here, we will focus on 20S-mediated degradation. Intrinsically disordered proteins or proteins with disordered regions in one or both ends of their sequence can be degraded directly by the 20S proteasome. p53 is a good example of this: it has been described to be degraded upon K48 polyubiquitination, but it can also be degraded in an ubiquitin-independent manner in a mechanism involving NQO1, which is able to act as a gatekeeper of the 20S proteasome.[20] More importantly, 20S UID is also known to be responsible for the degradation of oxidized proteins (reviewed in ref 21). Normal cell metabolism generates reactive oxygen species (ROS) as byproducts, which are able to oxidize proteins. This oxidative stress leads to unfolding of oxidized proteins, which compromises their function and exposes buried hydrophobic patches. This increases their susceptibility to degradation by UID. Furthermore, oxidative stress favors the dissociation of 20S proteasome core particles from 19S regulatory particles, enhancing its degradation capabilities. Apart from the aforementioned mechanisms (gate-keepers like NQO1, intrinsically disordered proteins, and 26S dissociation), 20S mediated degradation can also be regulated by other ones such as *de novo* synthesis, subcellular location, presence of alternative regulatory components shared with the immunoproteasome or activators of proteolytic activity as well as PTMs like tyrosine phosphorylation or S-glutathionylation of cysteines in 20S subunits (reviewed in ref 22).

Protein ubiquitination is also a central modification in autophagy, especially in the context of protein quality control. If misfolded proteins tagged for degradation with K48 linkage manage to escape the UPS, they can be degraded by autophagy. K63 is the main chain formation from the ubiquitination code that is able to trigger autophagy in the two cases discussed above, and adaptors such as p62 or NBR and HDAC6 have UBDs that are able to recognize K63 and K48 with lower affinity. K63 chains are also responsible for targeting different organelles for degradation. This includes mitochondria (also degraded with the atypical K6 chain and phosphorylation of ubiquitin in S65), ribosomes, and the endoplasmic reticulum. It is still somewhat controversial whether these chains also target peroxisomes, which have been linked to monoubiquitination instead.[14] In xenophagy, using *Salmonella sp.* infection as a model, bacteria are tagged mainly with K63 but also other atypical chains such as K27, K6 and Met1 linear chains.

*Phosphorylation.* Although important, ubiquitination is not the only PTM involved in protein degradation. Phosphorylation at specific sites can produce linear motifs called phosphodegrons, which have largely been studied in short-lived cell cycle proteins. These are able to recruit E3 ubiquitin ligases and initiate protein degradation by the UPS. On the other hand, phosphorylation can also mask an active degron, blocking recognition by ubiquitin ligases and stabilizing the protein.[23]

The overall role of phosphorylation on protein degradation has recently been assessed in a study performed on growing HeLa cells. This showed an overall tendency for phosphorylation to significantly increase protein lifetimes, with only a minority of phosphosites leading to destabilization.[24] This effect is site-specific, rather than protein-specific, that is, phosphorylation at different sites in the same protein can lead to different changes in protein degradation rate. Interestingly, multiple closeby phosphorylations tend to stabilize the protein further. The effects of phosphorylation on protein degradation rate are associated with structural and sequential features of the phosphosite's environment. For example, phosphosites in a more hydrophobic environment, as well as those in an exposed area or loops, tend to stabilize the protein. Additionally, phosphosites in sequence regions enriched in glutamate tend to be very stabilizing. Altogether, these observations suggest that kinase activity can extensively alter protein degradation rates in order to respond to cellular conditions. Further study will be needed to understand the effects of phosphorylation on protein degradation under different conditions. It is also unclear whether (de)stabilization upon phosphorylation is mediated solely by the production of phosphodegrons and degron inhibition. Protein phosphorylation can also influence autophagy. Notably, PI3K and ULK1, which are major initiators of autophagosome formation, depend on activation by phosphorylation mediated by the kinases mTOR and AMPK.[14]

Finally, it is worth mentioning that enzymes controlling ubiquitination levels (E3 ligases and DUBs) can be regulated by phosphorylation and enzymes controlling phosphorylation levels (kinases and phosphatases) can be regulated by ubiquitination, adding an extra level of complexity by introducing crosstalk between the two modifications.[25]

**N-Degron and C-Degron Pathways.** In the 1980s, it was proposed that there was a strong relationship between the identity of the N-terminal residue of a protein and its *in vivo* degradation rate, the so-called N-end rule. It comprised a


https://dx.doi.org/10.1021/acsomega.0c05318
*ACS Omega* 2021, 6, 5091−5100

straightforward, direct mapping between a certain N-terminal amino acid (N-degron) and an approximate degradation rate.[26,27] However, this relationship has turned out to be far more complex than initially appreciated. Multiple proteome-wide studies in different species find little or no relationship between the N-terminal amino acid and protein degradation rate.[5,28,29] Furthermore, studies over multiple decades have revealed multiple systems that recognize different N-degrons in specific proteins or their fragments, often conditionally.[26,27] Rather than certain amino acids being stabilizing or destabilizing independently of context, any of the 20 proteinogenic amino acids can be an N-terminal destabilizing residue given a certain sequence context. Instead of there being an N-end rule, there are multiple N-degron pathways.

N-degron pathways can be understood as proteolytic systems that recognize proteins containing N-degrons, which lead to their degradation by the UPS or autophagy. Known eukaryotic pathways comprise the Arg/N-degron pathway, the Pro/N-degron pathway, the Ac/N-degron pathway and the fMet/N-degron pathway.[26,27] The main component of an N-degron is a destabilizing N-terminal residue and lysine(s) available for ubiquitination. Most N-degrons need post-translational modifications to become competent (acquired degrons). These include, but are not limited to, Nt-acetylation, Nt-deamidation, and Nt-arginylation of the $\alpha$-amino groups of Nt-residues, as well as proteolytic cleavage. These modifications can take place constitutively or conditionally. The mature N-degrons can then be recognized by components of the N-degron pathways, such as E3 ubiquitin ligases, leading to protein degradation.

Analogously, but more recently discovered, C-terminal degrons consist of sequence features in the C-terminal of the proteins able to recruit E3 ligases and trigger the UPS.[26,27] They can be present in full length proteins, early truncated proteins or as a result of proteolytic cleavage, and as is the case for N-terminal degrons, they can be avoided by specific folding or association with other proteins. In recent years, several C-terminal pathways have been unraveled, most of them linked with Cullin RING-E3 ligase complexes (CRLs) such as Gly/C degron pathways, the RxxG/C-degron pathway, the Arg/C-degron pathway, the R3/C-degron pathway, and the E2/C-degron pathway.[27] Furthermore, there are recent studies suggesting different new motifs, as well as other ligases involved. For example, inhibition of CRLs is not able to avoid degradation of proteins with alanine and valine residues acting as C-terminal degron. What's more, an atypical RING ligase, CHIP, is recruited by C-terminal aspartate after caspase cleavage.

Just as our understanding of the workings of these pathways has vastly improved, so has our understanding of their function. Initially, the N-end rule was understood to generally govern protein degradation rates, that is, as the main factor behind programmed protein degradation. Nowadays, N-degron and C-degron pathways are thought to fulfill other roles than regular, programmed protein degradation. Instead, they seem to play roles in protein quality control, such as scavenging for improperly folded or aberrant proteins, destroying improperly located proteins, controlling the correct stoichiometries of protein complexes, or ensuring the correct incorporation of PTMs. They are also implicated in other processes, such as the regulation of the cell cycle or circadian rhythms. Since the main purpose of this Mini-Review is to synthesize knowledge on programmed protein degradation, we direct readers to reviews that provide greater coverage on these pathways.[26,27]

**Protein Disorder and Topology.** After some debate and contradictory results in initial studies, intrinsic disorder and protein topology have clearly emerged in recent years as important determinants of protein lifetimes.

It has been shown in multiple species that the presence of long disordered segments, either at one of the protein ends (≥30 residues) or internally (≥40 residues), have an effect on protein degradation rate;[2] longer disordered segments do not necessarily have a larger effect (Figure 3). Proteins that have

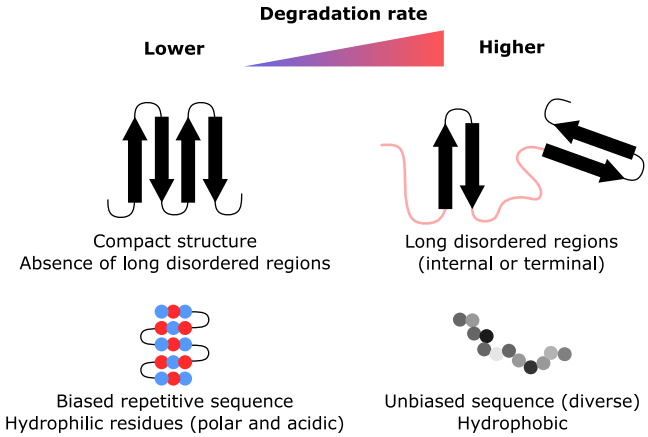


**Figure 3.** Influence of protein disorder and topology on protein degradation rates. Degradation rates tend to be higher if long disordered regions are present, either in protein termini (≥30 residues) or internally (≥40 residues). Proteasome recognition is enhanced in long disordered regions with hydrophobic and diverse sequences and limited, or even avoided, in regions with polar or acidic residues, as well as biased repetitive sequences.

both a long terminal disordered segment and an internal disordered segment tend to have the shortest degradation rates. Long disordered segments could influence degradation rate by containing specific motifs that target proteins for degradation. However, these segments are not enriched in either experimentally determined ubiquitination sites or predicted degradation motifs. In fact, most experimentally determined ubiquitination sites were found to be in structured regions. Furthermore, analysis of possible, unknown motifs in long disordered segments in short-lived proteins showed that they are actually not enriched in any motif that may serve as a specific degradation signal. Altogether, this indicates that long disordered segments do not affect protein degradation rate indirectly by embedding degradation signals.[2,28] Rather, this suggests that disordered segments have a direct effect. In fact, these critical minimum lengths correspond to the experimentally observed minimum lengths that allow the disordered terminus of a ubiquitinated substrate to efficiently initiate degradation and translocate the polypeptide into the proteolytic core. Thus, these disordered segments serve as initiation sites for degradation, where the proteasome can engage the substrate and start its degradation. The absence of such initiation sites explains the unexpected stability of certain proteins.[10]

However, a sufficiently long disordered segment is by itself not enough to efficiently initiate degradation. Rather, the proteasome has pronounced sequence preferences at initiation regions, which contribute substantially to protein lifetimes[10,11] (Figure 3). Initiation regions with biased sequences are poor for proteasome initiation, and very biased sequences may be rejected by the proteasome and escape degradation entirely.[30]


https://dx.doi.org/10.1021/acsomega.0c05318
ACS Omega 2021, 6, 5091−5100

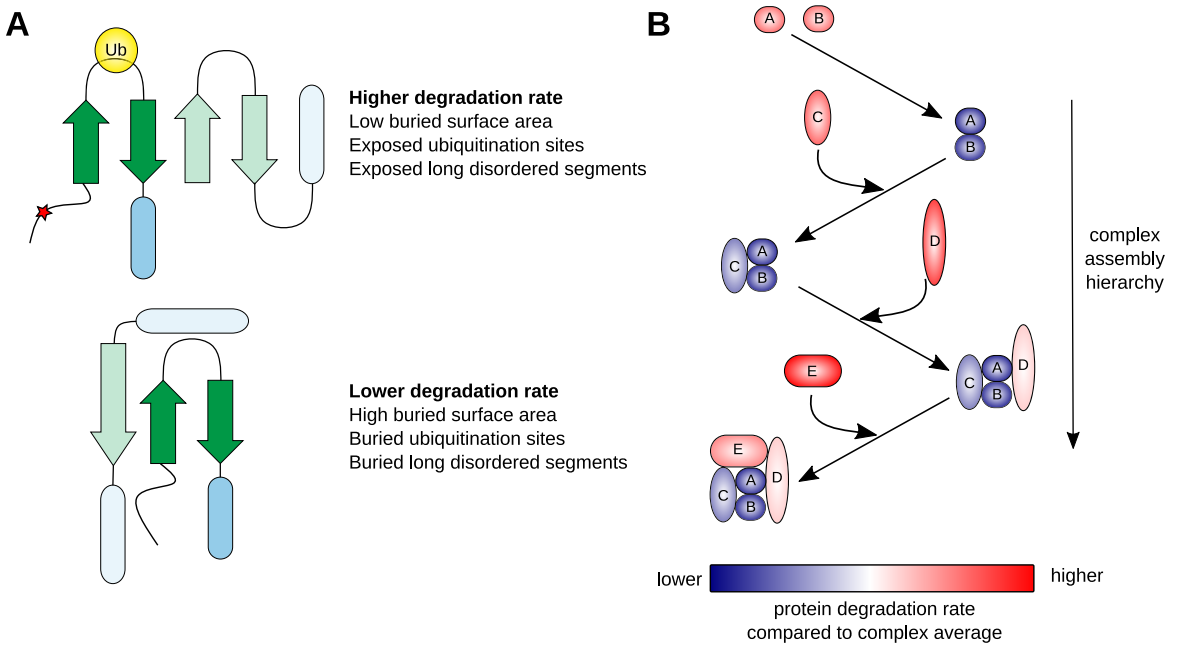


**Figure 4.** Influence of protein−protein interaction on protein degradation rate. (A) Schematic diagram of the characteristics of a protein−protein interaction that lead to higher (above) or lower (below) protein degradation rates, using a hypothetical protein and two interaction partners. Red star highlights exposed long disordered segments. (B) Influence of complex assembly order on variation in protein degradation rates within a complex. Although degradation rates within a complex tend to be homogeneous, there is a certain degree of variation. This variation is well explained by the order in which proteins assemble into the complex. Coloring of subunits in their monomeric form reflects the fact that proteins are stabilized upon complex binding.

Initiation regions rich in hydrophobic residues promote degradation better, while those rich in polar and acidic residues lead to greater stability. Finally, sequence dynamics also play a role, with stiffer sequences leading to more efficient degradation. Although it is not entirely clear what part of the proteasome binds to the initiation region, these preferences should reflect its physicochemical characteristics. These preferences might have implications for the accumulation of disease-associated proteins, which in many cases contain biased sequences (e.g huntingtin or $\alpha$-synuclein) and are able to escape degradation.[11]

Furthermore, overall protein disorder has a strong effect on protein degradation rate. This has an additional effect on top of the presence of terminal or internal disordered segments.[28] In fact, a machine learning model trained on protein degradation rate data using only features related to overall disorder and disordered segments performed almost as well as a model using a full set of sequence-derived features, highlighting the importance of intrinsic disorder.[28] However, there are different types of intrinsic disorder with distinct structural, biophysical, and functional properties.[31] For example, the net charge of intrinsically disordered regions can make them more or less compact. These characteristics may influence protein lifetimes in different ways. Furthermore, the topology of a protein, regardless of protein disorder, also influences protein degradation rate. Mallik and Kundu[12] studied this by correlating protein degradation rate with the absolute contact order. This parameter captures the overall topology of a protein: it is small for proteins that are mostly stabilized by local interactions; it is large for proteins that contain many long-range contacts; and it correlates to the force necessary to unfold a protein in molecular pulling experiments. Strikingly, the absolute contact order captures ∼50% of variation in degradation rate among monomeric proteins; the effect is smaller for oligomeric proteins, indicating that protein−protein interaction also plays a role. Both the effects of overall disorder and topology can be attributed to the mechanical resistance of the protein to being unfolded by the AAA+ ATPase molecular motor of the proteasome after engagement. The residue−residue contacts of the protein pose mechanical resistance that hampers unfolding, meaning highly disordered proteins or proteins stabilized by few contacts can be degraded more easily. Indeed, mechanical unfolding has been shown to be the rate-limiting step for degradation once a protein has been engaged by the proteasome.[30] It is unclear to what extent ubiquitin-independent proteasomal degradation plays a role in degrading highly disordered substrates *in vivo*.

Disordered regions are usually less evolutionarily constrained than those that form clearly structured domains, which has clear implications for the possibility of tuning degradation rate. Variations in protein degradation rate due to the relaxed constraints on disordered regions may represent an unappreciated source of phenotypic variability among individuals. Additionally, altering the composition or length of disordered regions of a protein could then be easily exploited across evolution to alter its degradation rate, and therefore its cellular levels, without altering its function. An analysis of yeast paralogs showed that gain/loss of disordered segments and divergence in the length of disordered segments or overall disorder leads to significant alterations in degradation rate among paralogs. These differences in degradation rate are substantial enough that they could have an effect in yeast physiology,[2] indicating that this might be a possible mechanism for subfunctionalization. Likewise, variations in protein topology across paralogs also lead to significant differences in degradation rate.[12] Finally, alternative splicing may also lead to isoforms with a different overall disorder, leading to differences in degradation rate.[2]

**Sequence Motifs as Destruction Signals.** Several short motifs have been proposed to act as inherent degrons. These include the destruction box (consensus sequence: RxxLxxxxN),


https://dx.doi.org/10.1021/acsomega.0c05318
*ACS Omega* 2021, 6, 5091−5100

the KEN-box (KENxxxN), and PEST sequences (stretches rich in P, E, S, and T). Several studies have identified a statistically significant, proteome-wide tendency for proteins containing such motifs to have a higher degradation rate. However, the effect is small,[5,28] and the presence of more than one such motif does not lead to faster degradation.[5] This indicates that although these sequence motifs may play important roles in specific processes (e.g., degradation of proteins containing destruction boxes or KEN-boxes in late mitosis after recognition by specific E3 ligases), they are not major players in regular protein turnover.

Interestingly, PEST regions are often disordered. Given the clear importance of intrinsic disorder for protein degradation and the fact that PEST regions were proposed at a time when the concept of intrinsic disorder itself was not mainstream, this raises the question of whether their effect may stem merely from their disorder, rather than because they act as a specific signal for degradation. An analysis using multiple features related to protein disorder found that not only do PEST regions have little explanatory power, but they actually provide redundant information to the model and are thus unnecessary to explain trends in protein degradation.[28] Although PEST regions have occasionally been linked to proteolysis by calpains, no general distinct mechanism for recognition of PEST regions or targeting by PEST regions has been found, and proteins containing them seem to be typically degraded by the UPS pathway.[32] Altogether, this suggests that the role of PEST regions as a specific degradation signal should be reevaluated.

**Protein−Protein Interactions.** Protein−protein interactions have been shown to play an important role in determining protein degradation rates. Overall, oligomeric proteins have longer degradation rates than monomeric proteins.[12] Furthermore, degradation rate is proportional to the number of complexes a protein participates in, likely because this raises the probability that the protein can be found in an oligomer. This indicates that protein−protein interaction hinders degradation. This effect is particularly strong in proteins that are core parts of the complex, in the sense that they are present in all forms of the complex and do not only form part of the complex temporarily. This effect is greater the more surface area is buried upon binding and seems to be mediated, at least partially, by the burial of ubiquitination sites and internal disordered segments, which would lead to escape from proteasomal recognition and engagement (Figure 4A). This also suggests that the higher the stability of the complex, the lower the degradation rate of its constituents. In the case of intrinsically disordered proteins, folding upon binding to its interaction partner(s) is another possible mechanism that would protect them from degradation.

Multiple studies have reported homogeneous degradation rates among the constituents of complexes, indicating coregulation of their degradation.[5,12] However, there is still some variation between complex subunits. This variation is very well explained by the order of the assembly pathway, with proteins that bind later in the assembly hierarchy having shorter degradation rates[12] (Figure 4B). Depending on several factors, such as the number of subunits in the complex, complex assembly can take up to several minutes. This indicates that the delay in binding can be long enough to affect protein degradation rates, possibly because of exposure of long disordered regions or degrons.

Paralogs can have different protein−protein interaction patterns, which should in turn affect protein degradation rates. Indeed, it has been observed that paralogs that bind to a different number of complexes significantly differ in their degradation rates.[12] Furthermore, since alternative splicing can yield variants with different protein−protein interaction patterns, it is to be expected that this should also lead to differences in degradation rates. Finally, it is possible that rewiring of interaction networks upon e.g. a change in internal or external conditions (for example, progression along the cell cycle or a change in growth conditions) will also lead to shifts in protein degradation rates, although this has not been explored.

**Subcellular Location.** It is clear that protein turnover is organized according to subcellular location to a certain degree, even though there seem to be no overall large differences between different cellular compartments. Results, however, are not always consistent between different studies. Since subcellular location serves as a layer of regulation, these differences may be due to the use of different species, cell types, and conditions.

The most clear, consistent finding among different studies is that mitochondrial proteins are clearly more long-lived than those of other subcellular compartments.[28,29,5,33] This is counterintuitive, given that mitochondrial proteins will be exposed to high levels of reactive oxygen species, which could damage them and lead to their degradation. It is unclear whether the long life of mitochondrial proteins is due to them having evolved to be sturdier, to the influence of intramitochondrial, non-UPS proteolytic pathways, or both. Finally, some studies have identified proteins located in the plasma membrane[28,33] and nuclear proteins[5,33] as being somewhat less stable.

**Global Sequence Properties.** Different studies have evaluated the importance of bulk sequence properties, such as global residue composition, for protein degradation. There is a clear tendency for proteins with an overall higher fraction of aliphatic and aromatic residues to be more stable, whereas those enriched in polar and charged amino acids tend to be short-lived.[5,28,29] It is unclear, however, what this reflects. It might be that proteins with a higher fraction of aliphatic and aromatic residues tend to have a hydrophobic core stabilized by more interactions, thereby opposing greater mechanical resistance to unfolding by the proteasome. Conversely, polar and charged residues tend to promote intrinsic disorder, which would make the protein easier to degrade. These global tendencies are opposed to those of initiation regions for degradation, where sequences rich in hydrophobic residues promote degradation more efficiently.

Amino acid composition also seems to play a role in protecting proteins against oxidative stress. A recent study shows differential amino acid usage between *Escherichia coli* and the radiation- and ROS-resistant bacterium *Deinococcus radiodurans*. For example, lysine and arginine are both positively charged at physiological pH, but lysine is more susceptible to oxidation by ROS. *D. radiodurans* has a strong preference for using arginine over lysine compared with *E. coli*. Tryptophan and tyrosine, which can act as antioxidants within a protein, are also significantly more abundant in *D. radiodurans*.[34] This suggests a relationship between amino acid composition, oxidative stress, and protein degradation rates. It is unclear, however, how these tendencies would be reflected in large-scale protein degradation rate data sets collected so far, as they are not necessarily collected under conditions with important oxidative stress. Nonetheless, this point merits attention.

Other bulk properties, such as the net charge, isoelectric point, sequence length, or molecular weight, seem to play a minor role or no role in protein degradation.[5,28,29,25]


https://dx.doi.org/10.1021/acsomega.0c05318
ACS Omega 2021, 6, 5091−5100

## CONCLUSIONS AND FUTURE PERSPECTIVES

Here, we have reviewed the main advances in our understanding of the determinants of protein degradation. This overview highlights how these determinants fit in the context of eukaryotic proteolytic systems. Most remarkably, it has been shown that intrinsic structural characteristics such as the presence of disordered segments and their composition and the overall disorder and topology of a protein, as well as oligomerization and order of complex assembly, explain much of protein degradation rates. Meanwhile, factors that have long been thought to be key to explain protein degradation globally, such as the identity of the N-terminal residue or specific motifs like the destruction box, are context dependent or do not play important roles.

Recent years have seen renewed interest in the study of the determinants of protein degradation. This has reflected in the increasing amounts of protein degradation rate data collected on different species, cell types, and conditions. This presents an opportunity to further an integrative understanding of what makes protein degradation tick. Analysis and modeling of this data as a whole can be used to put to the test established ideas, uncover new determinants of protein degradation, quantify the relative contribution of each parameter to protein degradation rate, and disentangle the interplay between different characteristics that may be correlated. These findings should be put in the light of our knowledge of proteolytic systems to obtain mechanistic understanding and will provide hypotheses for further experimentation. Furthermore, the models can be compared to obtain insight on how different characteristics influence degradation differently across species, cell types, or conditions. This also raises the question of whether it is possible to create a generic tool to predict protein degradation rates, just as we have tools to predict, for example, protein secondary structure or solubility. Since protein turnover is dependent on cell type and state,[5,28,33] this seems unlikely. However, trying to create such a tool is not necessarily a fool's errand. By doing so, we can still create reference models (for example, for protein production systems) and learn more about protein turnover.

Protein degradation rates have also begun to be seen from an evolutionary perspective, and it has begun to be perceived that variation in protein degradation rates (due to, for example, changes in the length of disordered segments) may be an unappreciated source of phenotypic variation. However, while the unveiled determinants of protein degradation reflect the preferences of proteolytic systems, the physiological reasons for why these preferences have evolved in the first place are not quite clear yet. Part of these preferences seem oriented toward controlling nonfunctional protein aggregation, which can be harmful to the cell. It has been shown that aggregation-prone proteins have shorter degradation rates. Together with other regulatory measures, such as tight translational control of aggregation-prone proteins, this contributes to keep the concentration of aggregation-prone proteins low and thus reduces the likelihood of potentially harmful aggregation.[35] Conversely, there is a certain functional organization of protein degradation rate; for example, proteins that participate in cell signaling tend to have short degradation rates,[5] which would contribute to regulate the duration of the transduced signal. This raises the reverse question of whether proteins that participate in certain processes have evolved sequential and structural characteristics in order to match a certain degradation rate. Further study is needed to understand what reasons have driven the evolution of protein degradation preferences.

Intriguingly, some authors have shown relationships in turnover between a protein and its transcript. Mandad et al. uncovered that differential use of synonymous codons alters protein degradation rate and that the codon sequence is predictive of protein degradation rate.[29] This illustrates that protein degradation can be determined by more than intrinsic protein properties. Additionally, Martin-Perez and Villén showed, in exponentially growing yeast, a strong correlation between protein degradation rate and mRNA degradation rate, which held across biological processes.[5] Although the mechanism is unclear, this indicates coregulation of protein and transcript lifetime. These thought-provoking findings show unexpected interconnections between different protein homeostasis parameters and highlights the importance of considering factors beyond intrinsic protein characteristics in the study of protein degradation. Understanding the full life cycle of proteins and the regulation of gene expression will require the study of phenomena such as these.

Impaired protein degradation is at the center of several pathogenic processes, such as cancer and proteinopathies, which include many neurodegenerative diseases. Better understanding of these processes has led to interesting and relevant therapeutic development in the last years, showing its potential. A good example is the use of proteasome inhibitors to treat cancer, such as bortezomib for myelomas and lymphomas. Also, in the treatment of cancer, arsenic in combination with other drugs is able to induce misfolding and degradation of the oncoprotein responsible for acute promyelocytic leukemia, one of the most effective treatments so far. The use of small molecules that are able to enhance proteasome activity is especially promising in the context of neurodegenerative diseases. Usp14 inhibitors are able to increase the degradation rate of the pathogenic aggregates of Tau protein (Alzheimer's disease) as well as TDP-43 (amyotrophic lateral sclerosis), opening new exciting therapeutic approaches.[3] Finally, a new set of drugs, called degraders,[16] which have the ability to trigger selective degradation of target proteins via UPS-recruiting E3 ligases, represent an invaluable strategy for basic research, as well as drug discovery and development of new therapies.

## AUTHOR INFORMATION


### Corresponding Authors

**Miguel Correa Marrero** − *European Bioinformatics Institute (EMBL-EBI), Hinxton, Cambridge CB10 1SD, United Kingdom;* orcid.org/0000-0002-3840-3273; Email: mcorrea@ebi.ac.uk

**Inigo Barrio-Hernandez** − *European Bioinformatics Institute (EMBL-EBI), Hinxton, Cambridge CB10 1SD, United Kingdom;* orcid.org/0000-0002-5686-0451; Email: ibarrioh@ebi.ac.uk

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.0c05318



### Funding

The authors received no specific funding for this work.


### Notes

The authors declare no competing financial interest.

### Biographies

Miguel Correa Marrero received his BSc in Biochemistry (2013) from the Autonomous University of Barcelona and his MSc in Bioinformatics



https://dx.doi.org/10.1021/acsomega.0c05318
ACS Omega 2021, 6, 5091−5100

(2015) from Wageningen University & Research. Until 2019, he was a Ph.D. student at the Bioinformatics Group of Wageningen University & Research, where he did research in multiple areas of protein bioinformatics, such as protein−protein interaction prediction using coevolutionary analysis and machine learning-based modeling of protein degradation rates. Currently, he is a postdoctoral fellow at the European Bioinformatics Institute (EMBL-EBI) in the group of Pedro Beltrao, where he studies the structural and functional impact of protein phosphorylation. His recent work has focused on collaborations to contribute to the understanding of the molecular basis of SARS-CoV-2 infection.

Iñigo Barrio-Hernandez received his BSc in Biology and Biochemistry (2011) from the University of Navarra. He got his Ph.D. from University of Southern Denmark (2016) under Prof. Blagoy Blagoev and Prof. Moustapha Kassem's supervision, studying human stem cell differentiation using mass spectrometry-based quantitative proteomics and phosphoproteomics. During this time, he also developed research in the field of protein ubiquitination. Since 2017, he has been a postdoctoral fellow at Pedro Beltrao's group (EMBL-EBI). His research is focused on network-based methods to identify groups of genes relevant for different diseases using GWAS outputs. He is also working in conservation and regulation of nondegradative protein ubiquitination, and in the last months, he has been collaborating in projects concerning the recent SARS-CoV2 outbreak.

## ■ ACKNOWLEDGMENTS

We apologize to authors whose publications could not be cited due to space constraints.

## ■ REFERENCES

(1) Amm, I.; Sommer, T.; Wolf, D. H. Protein Quality Control and Elimination of Protein Waste: The Role of the Ubiquitin-Proteasome System. *Biochim. Biophys. Acta, Mol. Cell Res.* **2014**, *1843* (1), 182−196.
(2) van der Lee, R.; Lang, B.; Kruse, K.; Gsponer, J.; Sánchez de Groot, N.; Huynen, M. A.; Matouschek, A.; Fuxreiter, M.; Babu, M. M. Intrinsically Disordered Segments Affect Protein Half-Life in the Cell and during Evolution. *Cell Rep.* **2014**, *8* (6), 1832−1844.
(3) Hanna, J.; Guerra-Moreno, A.; Ang, J.; Micoogullari, Y. Protein Degradation and the Pathologic Basis of Disease. *Am. J. Pathol.* **2019**, *189* (1), 94−103.
(4) Taylor, R. C.; Dillin, A. Aging as an Event of Proteostasis Collapse. *Cold Spring Harbor Perspect. Biol.* **2011**, *3* (5), a004440.
(5) Martin-Perez, M.; Villén, J. Determinants and Regulation of Protein Turnover in Yeast. *Cell Systems.* **2017**, *5*, 283−294.
(6) Akimov, V.; Barrio-Hernandez, I.; Hansen, S. V. F.; Hallenborg, P.; Pedersen, A.-K.; Bekker-Jensen, D. B.; Puglia, M.; Christensen, S. D. K.; Vanselow, J. T.; Nielsen, M. M.; Kratchmarova, I.; Kelstrup, C. D.; Olsen, J. V.; Blagoev, B. UbiSite Approach for Comprehensive Mapping of Lysine and N-Terminal Ubiquitination Sites. *Nat. Struct. Mol. Biol.* **2018**, *25* (7), 631−640.
(7) Lee, D. H.; Goldberg, A. L. Proteasome Inhibitors: Valuable New Tools for Cell Biologists. *Trends Cell Biol.* **1998**, *8* (10), 397−403.
(8) Chen, S.; Wu, J.; Lu, Y.; Ma, Y.-B.; Lee, B.-H.; Yu, Z.; Ouyang, Q.; Finley, D. J.; Kirschner, M. W.; Mao, Y. Structural Basis for Dynamic Regulation of the Human 26S Proteasome. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (46), 12991−12996.
(9) Lu, Y.; Lee, B.-H.; King, R. W.; Finley, D.; Kirschner, M. W. Substrate Degradation by the Proteasome: A Single-Molecule Kinetic Analysis. *Science* **2015**, *348* (6231), 1250834.
(10) Fishbain, S.; Inobe, T.; Israeli, E.; Chavali, S.; Yu, H.; Kago, G.; Babu, M. M.; Matouschek, A. Sequence Composition of Disordered Regions Fine-Tunes Protein Half-Life. *Nat. Struct. Mol. Biol.* **2015**, *22* (3), 214−221.
(11) Yu, H.; Singh Gautam, A. K.; Wilmington, S. R.; Wylie, D.; Martinez-Fonts, K.; Kago, G.; Warburton, M.; Chavali, S.; Inobe, T.; Finkelstein, I. J.; Babu, M. M.; Matouschek, A. Conserved Sequence Preferences Contribute to Substrate Recognition by the Proteasome. *J. Biol. Chem.* **2016**, *291* (28), 14526−14539.
(12) Mallik, S.; Kundu, S. Topology and Oligomerization of Mono- and Oligomeric Proteins Regulate Their Half-Lives in the Cell. *Structure* **2018**, *26* (6), 869−878.
(13) Tanaka, K. The Proteasome: Overview of Structure and Functions. *Proc. Jpn. Acad., Ser. B* **2009**, *85* (1), 12−36.
(14) Kwon, Y. T.; Ciechanover, A. The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy. *Trends Biochem. Sci.* **2017**, *42* (11), 873−886.
(15) Ji, C. H.; Kwon, Y. T. Crosstalk and Interplay between the Ubiquitin-Proteasome System and Autophagy. *Mol. Cells* **2017**, *40* (7), 441−449.
(16) Fisher, S. L.; Phillips, A. J. Targeted Protein Degradation and the Enzymology of Degraders. *Curr. Opin. Chem. Biol.* **2018**, *44*, 47−55.
(17) Swatek, K. N.; Komander, D. Ubiquitin Modifications. *Cell Res.* **2016**, *26* (4), 399−422.
(18) Braten, O.; Livneh, I.; Ziv, T.; Admon, A.; Kehat, I.; Caspi, L. H.; Gonen, H.; Bercovich, B.; Godzik, A.; Jahandideh, S.; Jaroszewski, L.; Sommer, T.; Kwon, Y. T.; Guharoy, M.; Tompa, P.; Ciechanover, A. Numerous Proteins with Unique Characteristics Are Degraded by the 26S Proteasome Following Monoubiquitination. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (32), E4639−E4647.
(19) Thrower, J. S.; Hoffman, L.; Rechsteiner, M.; Pickart, C. M. Recognition of the Polyubiquitin Proteolytic Signal. *EMBO J.* **2000**, *19* (1), 94−102.
(20) Tsvetkov, P.; Reuven, N.; Shaul, Y. Ubiquitin-Independent p53 Proteasomal Degradation. *Cell Death Differ.* **2010**, *17* (1), 103−108.
(21) Jung, T.; Höhn, A.; Grune, T. The Proteasome and the Degradation of Oxidized Proteins: Part II - Protein Oxidation and Proteasomal Degradation. *Redox Biol.* **2014**, *2*, 99−104.
(22) Ben-Nissan, G.; Sharon, M. Regulating the 20S Proteasome Ubiquitin-Independent Degradation Pathway. *Biomolecules* **2014**, *4* (3), 862−884.
(23) Holt, L. J. Regulatory Modules: Coupling Protein Stability to Phopshoregulation during Cell Division. *FEBS Lett.* **2012**, *586* (17), 2773−2777.
(24) Wu, C.; Ba, Q.; Lu, D.; Li, W.; Salovska, B.; Hou, P.; Mueller, T.; Rosenberger, G.; Gao, E.; Di, Y.; Zhou, H.; Fornasiero, E. F.; Liu, Y. Global and Site-Specific Effect of Phosphorylation on Protein Turnover. *Dev. Cell* **2021**, *56*, 111.
(25) Filipčík, P.; Curry, J. R.; Mace, P. D. When Worlds Collide—Mechanisms at the Interface between Phosphorylation and Ubiquitination. *J. Mol. Biol.* **2017**, *429*, 1097−1113.
(26) Varshavsky, A. N-Degron and C-Degron Pathways of Protein Degradation. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 358−366.
(27) Timms, R. T.; Koren, I. Tying up Loose Ends: The N-Degron and C-Degron Pathways of Protein Degradation. *Biochem. Soc. Trans.* **2020**, *48* (4), 1557−1567.
(28) Correa Marrero, M.; van Dijk, A. D. J.; de Ridder, D. Sequence-Based Analysis of Protein Degradation Rates. *Proteins: Struct., Funct., Genet.* **2017**, *85* (9), 1593−1601.
(29) Mandad, S.; Rahman, R.-U.; Centeno, T. P.; Vidal, R. O.; Wildhagen, H.; Rammner, B.; Keihani, S.; Opazo, F.; Urban, I.; Ischebeck, T.; Kirli, K.; Benito, E.; Fischer, A.; Yousefi, R. Y.; Dennerlein, S.; Rehling, P.; Feussner, I.; Urlaub, H.; Bonn, S.; Rizzoli, S. O.; Fornasiero, E. F. The Codon Sequences Predict Protein Lifetimes and Other Parameters of the Protein Life Cycle in the Mouse Brain. *Sci. Rep.* **2018**, *8* (1), 16913.
(30) Bard, J. A. M.; Bashore, C.; Dong, K. C.; Martin, A. The 26S Proteasome Utilizes a Kinetic Gateway to Prioritize Substrate Degradation. *Cell* **2019**, *177* (2), 286−298 e15..
(31) van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; Kim, P. M.; Kriwacki, R. W.; Oldfield, C. J.; Pappu, V. P.; Tompa, P.; Uversky, V. N.; Wright, P. E.; Babu, M. M. Classification of Intrinsically Disordered Regions and Proteins. *Chem. Rev.* **2014**, *114* (13), 6589−6631.


https://dx.doi.org/10.1021/acsomega.0c05318
*ACS Omega* 2021, 6, 5091−5100

(32) Rechsteiner, M.; Rogers, S. W. PEST Sequences and Regulation by Proteolysis. *Trends Biochem. Sci.* **1996**, *21* (7), 267−271.
(33) Dörrbaum, A. R.; Kochen, L.; Langer, J. D.; Schuman, E. M. Local and Global Influences on Protein Turnover in Neurons and Glia. *eLife* **2018**, *7*, e34202.
(34) Chang, R. L.; Stanley, J. A.; Robinson, M. C.; Sher, J. W.; Li, Z.; Chan, Y. A.; Omdahl, A. R.; Wattiez, R.; Godzik, A.; Matallana-Surget, S. Protein Structure, Amino Acid Composition and Sequence Determine Proteome Vulnerability to Oxidation-induced Damage. *EMBO J.* **2020**, *39*, e104523.
(35) Gsponer, J.; Babu, M. M. Cellular Strategies for Regulating Functional and Nonfunctional Protein Aggregation. *Cell Rep.* **2012**, *2* (5), 1425−1437.



https://dx.doi.org/10.1021/acsomega.0c05318
*ACS Omega* 2021, 6, 5091−5100